אלגברה ליניארית בדרך הישר
Linear Algebra Done Right

מהדורה רביעית

07 בספטמבר 2026

© 2024 שלדון אקסלר – Sheldon Axler

שלדון אקסלר – Sheldon Axler

הערות, תיקונים, והצעות
בנוגע לספר יתקבלו בברכה רבה.
אנא שלחו אותם לכתובת linear@axler.net.

–

תרגם לעברית: צורי בר יוחאי
Hebrew translation by Züri Bar Yochay

גרסה מודפסת של המהדורה האנגלית של ספר זה יצאה לאור בהוצאת שפרינגר.

גישה פתוחה ספר זה מופץ תחת רישיון בהתאם לתנאי רישיון Creative Commons Attribution-NonCommercial 4.0 International (https://creativecommons.org/licenses/by-nc/4.0), המתיר כל שימוש לא-מסחרי, שיתוף, עיבוד, הפצה ושכפול בכל אמצעי או פורמט, כל עוד יינתן קרדיט מתאים למחבר ולמקור המקוריים, יסופק קישור לרישיון Creative Commons, ויצוין אם נעשו שינויים. התמונות או חומרים אחרים של צד שלישי ספר זה כלולים ברישיון Creative Commons של הספר, אלא אם צוין אחרת בשורת הקרדיט לחומר. אם חומר כלשהו אינו כלול ברישיון Creative Commons של הספר, והשימוש המיועד בו אינו מותר על-פי דין או חורג מן השימוש המותר, יהיה עליך לקבל רשות ישירות מבעל זכויות היוצרים.

$$F_n = \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right]$$

About the Author

Sheldon Axler received his undergraduate degree from Princeton University, followed by a PhD in mathematics from the University of California at Berkeley.

As a postdoctoral Moore Instructor at MIT, Axler received a university-wide teaching award. He was then an assistant professor, associate professor, and professor at Michigan State University, where he received the first J. Sutherland Frame Teaching Award and the Distinguished Faculty Award.

Axler received the Lester R. Ford Award for expository writing from the Mathematical Association of America in 1996, for a paper that eventually expanded into this book. In addition to publishing numerous research papers, he is the author of six mathematics textbooks, ranging from freshman to graduate level. Previous editions of this book have been adopted as a textbook at over 375 universities and colleges and have been translated into three languages.

Axler has served as Editor-in-Chief of the *Mathematical Intelligencer* and Associate Editor of the *American Mathematical Monthly*. He has been a member of the Council of the American Mathematical Society and of the Board of Trustees of the Mathematical Sciences Research Institute. He has also served on the editorial board of Springer's series Undergraduate Texts in Mathematics, Graduate Texts in Mathematics, Universitext, and Springer Monographs in Mathematics.

Axler is a Fellow of the American Mathematical Society and has been a recipient of numerous grants from the National Science Foundation.

Axler joined San Francisco State University as chair of the Mathematics Department in 1997. He served as dean of the College of Science & Engineering from 2002 to 2015, when he returned to a regular faculty appointment as a professor in the Mathematics Department.

Carrie Heeter, Bishnu Sarangi

The author and his cat Moon.

Cover equation: Formula for the nth Fibonacci number. Exercise 21 in Section 5D uses linear algebra to derive this formula.

על המחבר

שלדון אקסלר קיבל את תואר הבוגר שלו מאוניברסיטת פרינסטון, ולאחריו תואר דוקטור במתמטיקה מאוניברסיטת קליפורניה בברקלי.

במהלך תקופת הפוסט־דוקטורט שלו כשהוא מועסק כמנחה (Moore Instructor) במכון הטכנולוגי של מסצ'וסטס (MIT), זכה אקסלר בפרס הצטיינות בהוראה מטעם האוניברסיטה. לאחר מכן כיהן כמרצה בכיר, כפרופסור חבר, וכפרופסור מן המניין באוניברסיטת מדינת מישיגן (Michigan State University), שם הוענקו לו פרס ההוראה הראשון ע"ש J. Sutherland Frame ופרס הסגל המצטיין (Distinguished Faculty Award).

בשנת 1996 זכה אקסלר בפרס לסטר פורד (Lester R. Ford) לכתיבה עיונית מטעם האיגוד האמריקאי למתמטיקה (MAA), עבור מאמר שהתרחב ברבות הימים לספר זה. בנוסף לפרסום מאמרי מחקר רבים, הוא חיבר שישה ספרי לימוד במתמטיקה, החל מרמת שנה א' באוניברסיטה ועד לרמת תארים מתקדמים. מהדורות קודמות של ספר זה אומצו כספר לימוד בלמעלה מ־ 375 אוניברסיטאות ומכללות, ותורגמו לשלוש שפות.

אקסלר שימש כעורך הראשי של כתב העת Mathematical Intelligencer וכעורך משנה של ה־ American Mathematical Monthly. הוא היה חבר במועצת האגודה האמריקאית למתמטיקה (AMS) ובחבר הנאמנים של המכון למחקר במדעי המתמטיקה (MSRI). כמו כן, כיהן במערכת העריכה של סדרות הספרים של הוצאת שפרינגר: Undergraduate Texts in Mathematics, Graduate Texts in Mathematics, Universitext, ו־Springer Monographs in Mathematics.

אקסלר הוא עמית של האגודה האמריקאית למתמטיקה, וזכה במענקי מחקר רבים מטעם הקרן הלאומית למדע (NSF).

בשנת 1997 הצטרף אקסלר לאוניברסיטת סן פרנסיסקו סטייט כראש המחלקה למתמטיקה. הוא שימש כדיקן הפקולטה למדעים והנדסה בין השנים 2002 ל־ 2015, ולאחר מכן חזר למשרת סגל רגילה כפרופסור במחלקה למתמטיקה.

Carrie Heeter, Bishnu Sarangi

המחבר וחתולו מון

המשוואה שעל עטיפת הספר: הנוסחה למספר פיבונאצ'י ה־ n^{th}. תרגיל 21 ביחידה ה4 משתמש באלגברה ליניארית לגזירת נוסחה זו.

Contents

About the Author **v**

Preface for Students **xii**

Preface for Instructors **xiii**

Acknowledgments **xvii**

Chapter 1
Vector Spaces **1**

1A $\mathbf{R}^n$ and $\mathbf{C}^n$ 2
Complex Numbers 2
Lists 5
$\mathbf{F}^n$ 6
Digression on Fields 10
Exercises 1A 10

1B Definition of Vector Space 12
Exercises 1B 16

1C Subspaces 18
Sums of Subspaces 19
Direct Sums 21
Exercises 1C 24

Chapter 2
Finite-Dimensional Vector Spaces **27**

2A Span and Linear Independence 28
Linear Combinations and Span 28
Linear Independence 31
Exercises 2A 37

תוכן העניינים

על המחבר v

דברי פתיחה לסטודנטים vii

דברי פתיחה למרצים viii

דברי תודה xii

פרק א

מרחבים וקטוריים 1

א1 $\mathbf{R}^n$ ו־ $\mathbf{C}^n$ 2

מספרים מרוכבים 2

רשימות 5

$\mathbf{F}^n$ 6

אתנחתא בנושא שדות 10

תרגילים א1 10

א2 הגדרת מרחב וקטורי 12

תרגילים א2 16

א3 תתי־מרחב 18

סכומי תתי־מרחב 19

סכומים ישרים 21

תרגילים א3 24

פרק ב

מרחבים וקטוריים ממד סופי 27

ב1 פְּרִיסָה ואי־תלות ליניארית 28

צירופים ליניאריים ופריסה 28

אי־תלות ליניארית 31

תרגילים ב1 37

אלגברה ליניארית בדרך הישר, מהדורה רביעית, מחבר: שלדון אקסלר
Linear Algebra Done Right, fourth edition, by Sheldon Axler

2B Bases 39
Exercises 2B 42

2C Dimension 44
Exercises 2C 48

Chapter 3
Linear Maps **51**

3A Vector Space of Linear Maps 52
Definition and Examples of Linear Maps 52
Algebraic Operations on $\mathcal{L}(V, W)$ 55
Exercises 3A 57

3B Null Spaces and Ranges 59
Null Space and Injectivity 59
Range and Surjectivity 61
Fundamental Theorem of Linear Maps 62
Exercises 3B 66

3C Matrices 69
Representing a Linear Map by a Matrix 69
Addition and Scalar Multiplication of Matrices 71
Matrix Multiplication 72
Column–Row Factorization and Rank of a Matrix 77
Exercises 3C 79

3D Invertibility and Isomorphisms 82
Invertible Linear Maps 82
Isomorphic Vector Spaces 86
Linear Maps Thought of as Matrix Multiplication 88
Change of Basis 90
Exercises 3D 93

3E Products and Quotients of Vector Spaces 96
Products of Vector Spaces 96
Quotient Spaces 98
Exercises 3E 103

3F Duality 105
Dual Space and Dual Map 105
Null Space and Range of Dual of Linear Map 109

ב2 בסיסים 39
תרגילים ב2 42
ב3 ממד 44
תרגילים ב3 48

פרק ג
העתקות ליניאריות 51

ג1 מרחב וקטורי של העתקות ליניאריות 52
הגדרה ודוגמאות של העתקות ליניאריות 52
פעולות אלגבריות מעל $\mathcal{L}(V,W)$ 55
תרגילים ג1 57
ג2 מרחבי אפס וטווחים 59
מרחב אפס וחד־חד־ערכיות 59
טווח ותכונת העל 61
המשפט היסודי של ההעתקות הליניאריות 62
תרגילים ג2 66
ג3 מטריצות 69
ייצוג העתקה ליניארית באמצעות מטריצה 69
חיבור וכפל בסקלר של מטריצות 71
כפל מטריצות 72
פירוק עמודה־שורה ודרגת מטריצה 77
תרגילים ג3 79
ג4 הֲפִיכוּת ואיזומורפיזמים 82
העתקות ליניאריות הֲפִיכוֹת 82
מרחבים וקטוריים איזומורפיים 86
העתקות ליניאריות ככפל מטריצות 88
שינוי בסיס 90
תרגילים ג4 93
ג5 מכפלות ומנות של מרחבים וקטוריים 96
מכפלות של מרחבים וקטוריים 96
מרחבי מנה 98
תרגילים ג5 103
ג6 דואליות 105
מרחב דואלי והעתקה דואלית 105
מרחב האפס והטווח של העתקה דואלית 109

Matrix of Dual of Linear Map 113
Exercises 3F 115

Chapter 4
Polynomials **119**
Zeros of Polynomials 122
Division Algorithm for Polynomials 123
Factorization of Polynomials over **C** 124
Factorization of Polynomials over **R** 127
Exercises 4 129

Chapter 5
Eigenvalues and Eigenvectors **132**

5A Invariant Subspaces 133
Eigenvalues 133
Polynomials Applied to Operators 137
Exercises 5A 139

5B The Minimal Polynomial 143
Existence of Eigenvalues on Complex Vector Spaces 143
Eigenvalues and the Minimal Polynomial 144
Eigenvalues on Odd-Dimensional Real Vector Spaces 149
Exercises 5B 150

5C Upper-Triangular Matrices 154
Exercises 5C 160

5D Diagonalizable Operators 163
Diagonal Matrices 163
Conditions for Diagonalizability 165
Gershgorin Disk Theorem 170
Exercises 5D 172

5E Commuting Operators 175
Exercises 5E 179

Chapter 6
Inner Product Spaces **181**

6A Inner Products and Norms 182
Inner Products 182

מטריצה של העתקה דואלית 113
תרגילים ג6 115

פרק ד
פולינומים 119
שורשי פולינומים 122
אלגוריתם חלוקת פולינומים 123
פירוק לגורמים של פולינומים מעל **C** 124
פירוק לגורמים של פולינומים מעל **R** 127
תרגילים ד 129

פרק ה
ערכים עצמיים ווקטורים עצמיים 132
ה1 תתי־מרחב שמורים 133
ערכים עצמיים 133
יישום פולינומים על אופרטורים 137
תרגילים ה1 139
ה2 הפולינום המינימלי 143
קיומם של ערכים עצמיים מעל מרחב וקטורי מרוכב 143
ערכים עצמיים והפולינום המינימלי 144
ערכים עצמיים מעל מרחבים וקטוריים ממשיים מממד אי־זוגי 149
תרגילים ה2 150
ה3 מטריצות משולשיות עליונות 154
תרגילים ה3 160
ה4 אופרטורים לכסינים 163
מטריצות אלכסוניות 163
תנאים ללכסינות 165
משפט עיגולי גרשגורן 170
תרגילים ה4 172
ה5 אופרטורים מתחלפים 175
תרגילים ה5 179

פרק ו
מרחבי מכפלה פנימית 181
ו1 מכפלות פנימיות ונורמות 182
מכפלות פנימיות 182

Norms 186
Exercises 6A 191

6B Orthonormal Bases 197
Orthonormal Lists and the Gram–Schmidt Procedure 197
Linear Functionals on Inner Product Spaces 204
Exercises 6B 207

6C Orthogonal Complements and Minimization Problems 211
Orthogonal Complements 211
Minimization Problems 217
Pseudoinverse 220
Exercises 6C 224

Chapter 7
Operators on Inner Product Spaces **227**

7A Self-Adjoint and Normal Operators 228
Adjoints 228
Self-Adjoint Operators 233
Normal Operators 235
Exercises 7A 239

7B Spectral Theorem 243
Real Spectral Theorem 243
Complex Spectral Theorem 246
Exercises 7B 247

7C Positive Operators 251
Exercises 7C 255

7D Isometries, Unitary Operators, and Matrix Factorization 258
Isometries 258
Unitary Operators 260
QR Factorization 263
Cholesky Factorization 266
Exercises 7D 268

7E Singular Value Decomposition 270
Singular Values 270
SVD for Linear Maps and for Matrices 273
Exercises 7E 278

נורמות 186
תרגילים ו1 191

ו2 בסיסים אורתונורמליים 197
רשימות אורתונורמליות ותהליך גרם־שמידט 197
פונקציונלים ליניאריים מעל מרחבי מכפלה פנימית 204
תרגילים ו2 207

ו3 משלימים אורתוגונליים ובעיות מזעור 211
משלימים אורתוגונליים 211
בעיות מזעור 217
פְּסֵידוֹ־הופכי 220
תרגילים ו3 224

פרק ז
אופרטורים מעל מרחבי מכפלה פנימית 227

ז1 אופרטורים צמודים לעצמם ואופרטורים נורמליים 228
צמודים 228
אופרטורים צמודים לעצמם 233
אופרטורים נורמליים 235
תרגילים ז1 239

ז2 משפט הפירוק הספקטרלי 243
משפט הפירוק הספקטרלי הממשי 243
משפט הפירוק הספקטרלי המרוכב 246
תרגילים ז2 247

ז3 אופרטורים חיוביים 251
תרגילים ז3 255

ז4 איזומטריות, אופרטורים אוניטריים, ופירוק מטריצה לגורמים 258
איזומטריות 258
אופרטורים אוניטריים 260
פירוק QR 263
פירוק שולסקי 266
תרגילים ז4 268

ז5 פירוק לערכים סינגולריים 270
ערכים סינגולריים 270
פירוק לערכים סינגולריים עבור העתקות ליניאריות ומטריצות 273
תרגילים ז5 278

7F Consequences of Singular Value Decomposition 280
Norms of Linear Maps 280
Approximation by Linear Maps with Lower-Dimensional Range 283
Polar Decomposition 285
Operators Applied to Ellipsoids and Parallelepipeds 287
Volume via Singular Values 291
Properties of an Operator as Determined by Its Eigenvalues 293
Exercises 7F 294

Chapter 8
Operators on Complex Vector Spaces **297**

8A Generalized Eigenvectors and Nilpotent Operators 298
Null Spaces of Powers of an Operator 298
Generalized Eigenvectors 300
Nilpotent Operators 303
Exercises 8A 306

8B Generalized Eigenspace Decomposition 308
Generalized Eigenspaces 308
Multiplicity of an Eigenvalue 310
Block Diagonal Matrices 314
Exercises 8B 316

8C Consequences of Generalized Eigenspace Decomposition 319
Square Roots of Operators 319
Jordan Form 321
Exercises 8C 324

8D Trace: A Connection Between Matrices and Operators 326
Exercises 8D 330

Chapter 9
Multilinear Algebra and Determinants **332**

9A Bilinear Forms and Quadratic Forms 333
Bilinear Forms 333
Symmetric Bilinear Forms 337
Quadratic Forms 341
Exercises 9A 344

ז6 מסקנות מהפירוק לערכים סינגולריים 280
נורמות של העתקות ליניאריות 280
קירובים באמצעות העתקות ליניאריות עם ממד טווח נמוך יותר 283
פירוק פולרי 285
יישום אופרטורים על אליפסואידים ומקבילונים 287
נפח באמצעות ערכים סינגולריים 291
תכונות של אופרטור כפי שנקבעות על ידי הערכים העצמיים שלו 293
תרגילים ז6 294

פרק ח
אופרטורים מעל מרחבים וקטוריים מרוכבים 297

ח1 וקטורים עצמיים מוכללים ואופרטורים נילפוטנטיים 298
מרחבי האפס של חזקות אופרטור 298
וקטורים עצמיים מוכללים 300
אופרטורים נילפוטנטיים 303
תרגילים ח1 306

ח2 פירוק למרחבים עצמיים מוכללים 308
מרחבים עצמיים מוכללים 308
ריבוי של ערך עצמי 310
מטריצות בלוקים אלכסוניות 314
תרגילים ח2 316

ח3 מסקנות מהפירוק למרחבים עצמיים מוכללים 319
שורשים ריבועיים של אופרטורים 319
צורת ז'ורדן 321
תרגילים ח3 324

ח4 עקבה: הקשר בין מטריצות ואופרטורים 326
תרגילים ח4 330

פרק ט
אלגברה מולטי־ליניארית ודטרמיננטות 332

ט1 תבניות בילינאריות ותבניות ריבועיות 333
תבניות בילינאריות 333
תבניות בילינאריות סימטריות 337
תבניות ריבועיות 341
תרגילים ט1 344

9B Alternating Multilinear Forms 346
Multilinear Forms 346
Alternating Multilinear Forms and Permutations 348
Exercises 9B 352

9C Determinants 354
Defining the Determinant 354
Properties of Determinants 357
Exercises 9C 367

9D Tensor Products 370
Tensor Product of Two Vector Spaces 370
Tensor Product of Inner Product Spaces 376
Tensor Product of Multiple Vector Spaces 378
Exercises 9D 380

Photo Credits **383**

Symbol Index **384**

Index **385**

Colophon: Notes on Typesetting **390**

ט2 תבניות מולטי־ליניאריות מתחלפות 346
תבניות מולטי־ליניאריות 346
תבניות מולטי־ליניאריות מתחלפות ותמורות 348
תרגילים ט2 352

ט3 דטרמיננטות 354
הגדרת הדטרמיננטה 354
תכונות של דטרמיננטות 357
תרגילים ט3 367

ט4 מכפלות טנזוריות 370
מכפלה טנזורית של שני מרחבים וקטוריים 370
מכפלה טנזורית של מרחבי מכפלה פנימית 376
מכפלה טנזורית של מספר מרחבים וקטוריים 378
תרגילים ט4 380

קרדיטים לתמונות 383

אינדקס סמלים וסימנים 384

אינדקס 385

קולופון: הערות בנוגע לסְדָר 391

Preface for Students

You are probably about to begin your second exposure to linear algebra. Unlike your first brush with the subject, which probably emphasized Euclidean spaces and matrices, this encounter will focus on abstract vector spaces and linear maps. These terms will be defined later, so don't worry if you do not know what they mean. This book starts from the beginning of the subject, assuming no knowledge of linear algebra. The key point is that you are about to immerse yourself in serious mathematics, with an emphasis on attaining a deep understanding of the definitions, theorems, and proofs.

You cannot read mathematics the way you read a novel. If you zip through a page in less than an hour, you are probably going too fast. When you encounter the phrase "as you should verify", you should indeed do the verification, which will usually require some writing on your part. When steps are left out, you need to supply the missing pieces. You should ponder and internalize each definition. For each theorem, you should seek examples to show why each hypothesis is necessary. Discussions with other students should help.

As a visual aid, definitions and notations are in yellow boxes (and have a title beginning with the word "definition" or "notation") and theorems are in blue boxes (in color versions of the book). Each theorem has an informal descriptive name.

Please check the website below for additional information about the book, including a link to videos that are freely available to accompany the book.

Your suggestions, comments, and corrections are most welcome.

Best wishes for success and enjoyment in learning linear algebra!

Sheldon Axler
San Francisco State University

website: https://linear.axler.net
e-mail: linear@axler.net

דברי פתיחה לסטודנטים

סביר להניח שאתם עומדים בפני המפגש השני שלכם עם אלגברה ליניארית. שלא כמו בהיכרות הראשונה שלכם עם הנושא, שככל הנראה שמה דגש על מרחבים אוקלידיים ומטריצות, הפעם המוקד יהיה על מרחבים וקטוריים מופשטים ועל העתקות ליניאריות. מונחים אלו יוגדרו בהמשך, כך שאין צורך לדאוג אם אינכם יודעים את משמעותם כעת. ספר זה מתחיל מראשית הנושא, ואינו מניח שום ידע מוקדם באלגברה ליניארית. הנקודה המרכזית היא שאתם עומדים לשקוע במתמטיקה רצינית, תוך שימת דגש על השגת הבנה מעמיקה של ההגדרות, המשפטים, וההוכחות.

אי אפשר לקרוא מתמטיקה כפי שקוראים רומן. אם אתם עוברים על עמוד בפחות משעה, סביר להניח שאתם מתקדמים מהר מדי. כאשר אתם נתקלים בביטוי "כפי שעליכם לוודא", עליכם אכן לבצע את הבדיקה, משימה שלרוב תדרוש מכם כתיבה כלשהי. כאשר שלבים הושמטו, עליכם להשלים את החסר. עליכם להתעמק בכל הגדרה ולהפנים אותה. עבור כל משפט, עליכם לחפש דוגמאות שימחישו מדוע כל הנחה בו היא הכרחית. דיונים עם סטודנטים אחרים עשויים לעזור.

כעזר חזותי, הגדרות וסימונים מופיעים בתיבות צהובות (וכותרתם מתחילה במילה "הגדרה" או "סימון") והמשפטים בתיבות כחולות (בגרסאות הצבעוניות של הספר). לכל משפט יש שם תיאורי לא־פורמלי.

אנא בקרו באתר האינטרנט שלהלן למידע נוסף על הספר, לרבות קישור לסרטונים הזמינים בחינם והמלווים את החומר.

הצעות, הערות, ותיקונים יתקבלו בברכה רבה.

איחוליי להצלחה והנאה בלימוד אלגברה ליניארית!

שלדון אקסלר
אוניברסיטת סן פרנסיסקו סטייט

אתר: https://linear.axler.net
מייל: linear@axler.net

Preface for Instructors

You are about to teach a course that will probably give students their second exposure to linear algebra. During their first brush with the subject, your students probably worked with Euclidean spaces and matrices. In contrast, this course will emphasize abstract vector spaces and linear maps.

The title of this book deserves an explanation. Most linear algebra textbooks use determinants to prove that every linear operator on a finite-dimensional complex vector space has an eigenvalue. Determinants are difficult, nonintuitive, and often defined without motivation. To prove the theorem about existence of eigenvalues on complex vector spaces, most books must define determinants, prove that a linear operator is not invertible if and only if its determinant equals 0, and then define the characteristic polynomial. This tortuous (torturous?) path gives students little feeling for why eigenvalues exist.

In contrast, the simple determinant-free proofs presented here (for example, see 5.19) offer more insight. Once determinants have been moved to the end of the book, a new route opens to the main goal of linear algebra—understanding the structure of linear operators.

This book starts at the beginning of the subject, with no prerequisites other than the usual demand for suitable mathematical maturity. A few examples and exercises involve calculus concepts such as continuity, differentiation, and integration. You can easily skip those examples and exercises if your students have not had calculus. If your students have had calculus, then those examples and exercises can enrich their experience by showing connections between different parts of mathematics.

Even if your students have already seen some of the material in the first few chapters, they may be unaccustomed to working exercises of the type presented here, most of which require an understanding of proofs.

Here is a chapter-by-chapter summary of the highlights of the book:

- Chapter 1: Vector spaces are defined in this chapter, and their basic properties are developed.
- Chapter 2: Linear independence, span, basis, and dimension are defined in this chapter, which presents the basic theory of finite-dimensional vector spaces.
- Chapter 3: This chapter introduces linear maps. The key result here is the fundamental theorem of linear maps: if T is a linear map on V, then $\dim V = \dim \text{null } T + \dim \text{range } T$. Quotient spaces and duality are topics in this chapter at a higher level of abstraction than most of the book; these topics can be skipped (except that duality is needed for tensor products in Section 9D).

דברי פתיחה למרצים

אתם עומדים ללמד קורס שיעניק לסטודנטים, ככל הנראה, את חשיפתם השנייה לאלגברה ליניארית. במפגשם הראשון עם הנושא, סביר להניח שהסטודנטים שלכם עבדו עם מרחבים אוקלידיים ועם מטריצות. לעומת זאת, בקורס זה נשים את הדגש על מרחבים וקטוריים מופשטים ועל העתקות ליניאריות.

שם הספר ראוי להסבר. רוב ספרי הלימוד באלגברה ליניארית משתמשים בדטרמיננטות כדי להוכיח שלכל אופרטור ליניארי מעל מרחב וקטורי מרוכב מממד סופי יש ערך עצמי. נושא הדטרמיננטות אינו קל, אינו אינטואיטיבי, ולעיתים קרובות מוגדר ללא מוטיבציה. כדי להוכיח את המשפט על קיום ערכים עצמיים במרחבים וקטוריים מרוכבים, רוב הספרים נאלצים להגדיר דטרמיננטות, להוכיח שאופרטור ליניארי אינו הפיך אם־ורק־אם הדטרמיננטה שלו שווה 0, ורק אז להגדיר את הפולינום האופייני. דרך חתחתים (מייסרת?) זו מעניקה לסטודנטים תפיסה מועטה להצדקת קיומם של ערכים עצמיים.

לעומת זאת, ההוכחות הפשוטות ונטולות־הדטרמיננטות המוצגות כאן (ראו למשל 5.19) מעניקות תובנות עמוקות יותר. משעה שנושא הדטרמיננטות הועבר לסוף הספר, אנו סוללים דרך חדשה אל המטרה המרכזית של האלגברה הליניארית — הבנת המבנה של אופרטורים ליניאריים.

ספר זה מתחיל מראשית הנושא, ללא דרישות קדם מלבד הדרישה המקובלת לבשלות מתמטית מתאימה. חלק קטן מהדוגמאות והתרגילים כולל מושגים מחשבון אינפיניטסימלי, כגון רציפות, גזירה, ואינטגרציה. ניתן לדלג בקלות על דוגמאות ותרגילים אלה אם הסטודנטים שלכם עדיין לא למדו חשבון אינפיניטסימלי. אם הסטודנטים שלכם כבר נחשפו לחשבון אינפיניטסימלי, הדוגמאות והתרגילים הללו אמורים להעשיר את חווייתם בכך שהם מצביעים על קשרים בין חלקים שונים של המתמטיקה.

גם אם הסטודנטים שלכם כבר נחשפו לחלק מהחומר שבפרקים הראשונים, ייתכן שאינם מורגלים בפתרון תרגילים מהסוג המוצג כאן, אשר רובם דורשים הבנה של הוכחות.

להלן סיכום של עיקרי הספר פרק־אחר־פרק:

- פרק א: בפרק זה מוגדרים מרחבים וקטוריים, ומפותחות תכונותיהם הבסיסיות.
- פרק ב: בפרק זה מוגדרים אי־תלות ליניארית, פריסה, בסיס, וממד, ובו מוצגת התאוריה הבסיסית של מרחבים וקטוריים מממד סופי.
- פרק ג: בפרק זה אנו עורכים היכרות עם העתקות ליניאריות. המסקנה המרכזית בפרק היא המשפט היסודי של ההעתקות הליניאריות: אם T העתקה ליניארית מעל V, אזי $\dim V = \dim \operatorname{null} T + \dim \operatorname{range} T$. רמת ההפשטה בנושאים מרחבי מנה ודואליות גבוהה יותר מרוב שאר הנושאים בספר; ניתן לדלג על נושאים אלו (דואליות נדרשת עבור מכפלות טנזוריות ביחידה ט4).

אלגברה ליניארית בדרך הישר, מהדורה רביעית, מחבר: שלדון אקסלר
Linear Algebra Done Right, fourth edition, by Sheldon Axler

- Chapter 4: The part of the theory of polynomials that will be needed to understand linear operators is presented in this chapter. This chapter contains no linear algebra. It can be covered quickly, especially if your students are already familiar with these results.
- Chapter 5: The idea of studying a linear operator by restricting it to small subspaces leads to eigenvectors in the early part of this chapter. The highlight of this chapter is a simple proof that on complex vector spaces, eigenvalues always exist. This result is then used to show that each linear operator on a complex vector space has an upper-triangular matrix with respect to some basis. The minimal polynomial plays an important role here and later in the book. For example, this chapter gives a characterization of the diagonalizable operators in terms of the minimal polynomial. Section 5E can be skipped if you want to save some time.
- Chapter 6: Inner product spaces are defined in this chapter, and their basic properties are developed along with tools such as orthonormal bases and the Gram–Schmidt procedure. This chapter also shows how orthogonal projections can be used to solve certain minimization problems. The pseudoinverse is then introduced as a useful tool when the inverse does not exist. The material on the pseudoinverse can be skipped if you want to save some time.
- Chapter 7: The spectral theorem, which characterizes the linear operators for which there exists an orthonormal basis consisting of eigenvectors, is one of the highlights of this book. The work in earlier chapters pays off here with especially simple proofs. This chapter also deals with positive operators, isometries, unitary operators, matrix factorizations, and especially the singular value decomposition, which leads to the polar decomposition and norms of linear maps.
- Chapter 8: This chapter shows that for each operator on a complex vector space, there is a basis of the vector space consisting of generalized eigenvectors of the operator. Then the generalized eigenspace decomposition describes a linear operator on a complex vector space. The multiplicity of an eigenvalue is defined as the dimension of the corresponding generalized eigenspace. These tools are used to prove that every invertible linear operator on a complex vector space has a square root. Then the chapter gives a proof that every linear operator on a complex vector space can be put into Jordan form. The chapter concludes with an investigation of the trace of operators.
- Chapter 9: This chapter begins by looking at bilinear forms and showing that the vector space of bilinear forms is the direct sum of the subspaces of symmetric bilinear forms and alternating bilinear forms. Then quadratic forms are diagonalized. Moving to multilinear forms, the chapter shows that the subspace of alternating n-linear forms on an n-dimensional vector space has dimension one. This result leads to a clean basis-free definition of the determinant of an operator. For complex vector spaces, the determinant turns out to equal the product of the eigenvalues, with each eigenvalue included in the product as many times as its multiplicity. The chapter concludes with an introduction to tensor products.

- פרק ד: בפרק זה אנו מציגים את החלקים בתאוריה של פולינומים אשר יידרשו להבנת אופרטורים ליניאריים. פרק זה אינו מכיל נושאים באלגברה ליניארית. אפשר לשקול לעבור עליו בזריזות, בייחוד אם לסטודנטים ישנה היכרות מוקדמת עם הנושא והמסקנות המובאות.

- פרק ה: הרעיון של חקירת אופרטור ליניארי באמצעות צמצומו לתתי־מרחב קטנים יותר מוביל לווקטורים עצמיים בחלקו המוקדם של פרק זה. שיאו של הפרק הוא הוכחה פשוטה לכך שבמרחבים וקטוריים מרוכבים תמיד קיימים ערכים עצמיים. מסקנה זו משמשת לאחר מכן להראות שלכל אופרטור ליניארי על מרחב וקטורי מרוכב יש מטריצה משולשית עליונה ביחס לבסיס כלשהו. הפולינום המינימלי ממלא תפקיד חשוב כאן ובהמשך הספר. למשל, פרק זה נותן אפיון של האופרטורים הלכסינים במונחי הפולינום המינימלי. ניתן לדלג על יחידה ה5 אם ברצונכם לחסוך זמן.

- פרק ו: בפרק זה מוגדרים מרחבי מכפלה פנימית, ומפותחות תכונותיהם הבסיסיות לצד כלים כגון בסיסים אורתונורמליים ותהליך גרם־שמידט. פרק זה גם מראה כיצד ניתן להשתמש בהטלות אורתוגונליות לפתרון בעיות מזעור מסוימות. לאחר מכן מוצג הפסידו־הופכי ככלי שימושי כאשר ההופכי אינו קיים. ניתן לדלג על החומר על הפסידו־הופכי אם ברצונכם לחסוך זמן.

- פרק ז: משפט הפירוק הספקטרלי, המאפיין את האופרטורים הליניאריים שעבורם קיים בסיס אורתונורמלי המורכב מווקטורים עצמיים, הוא משיאי הספר. העבודה שנעשתה בפרקים הקודמים נושאת כאן פרי בדמות הוכחות פשוטות במיוחד. פרק זה עוסק גם באופרטורים חיוביים, באיזומטריות, באופרטורים אוניטריים, בפירוקי מטריצות, ובמיוחד בפירוק לערכים סינגולריים, המוביל לפירוק הפולרי ולנורמות של העתקות ליניאריות.

- פרק ח: פרק זה מראה שלכל אופרטור מעל מרחב וקטורי מרוכב קיים בסיס של המרחב הווקטורי המורכב מווקטורים עצמיים מוכללים של האופרטור. לאחר מכן, הפירוק למרחבים עצמיים מוכללים מתאר אופרטור ליניארי מעל מרחב וקטורי מרוכב. הריבוי של ערך עצמי מוגדר כממד של המרחב העצמי המוכלל המתאים. כלים אלה משמשים להוכחה שלכל אופרטור ליניארי הפיך מעל מרחב וקטורי מרוכב יש שורש ריבועי. בהמשך מובאת בפרק הוכחה שכל אופרטור ליניארי מעל מרחב וקטורי מרוכב ניתן להבאה לצורת ז'ורדן. הפרק מסתיים בחקירת העקבה של אופרטורים.

- פרק ט: פרק זה פותח בהתבוננות בתבניות בילינאריות ומראה שהמרחב הווקטורי של התבניות הבילינאריות הוא סכום ישר של התת־מרחב של התבניות הבילינאריות הסימטריות ושל התת־מרחב של התבניות הבילינאריות המתחלפות. לאחר מכן מלוכסנות תבניות ריבועיות. במעבר לתבניות מולטי־לינאריות, הפרק מראה שהתת־מרחב של תבניות n־ליניאריות המתחלפות על מרחב וקטורי n־ממדי הוא חד־ממדי. מסקנה זו מובילה להגדרה נקייה, ללא תלות בבסיס, של הדטרמיננטה של אופרטור. עבור מרחבים וקטוריים מרוכבים מתברר שהדטרמיננטה שווה למכפלת הערכים העצמיים, כאשר כל ערך עצמי נכלל במכפלה מספר פעמים השווה לריבויו. הפרק מסתיים במבוא למכפלות טנזוריות.

This book usually develops linear algebra simultaneously for real and complex vector spaces by letting $\mathbf{F}$ denote either the real or the complex numbers. If you and your students prefer to think of $\mathbf{F}$ as an arbitrary field, then see the comments at the end of Section 1A. I prefer avoiding arbitrary fields at this level because they introduce extra abstraction without leading to any new linear algebra. Also, students are more comfortable thinking of polynomials as functions instead of the more formal objects needed for polynomials with coefficients in finite fields. Finally, even if the beginning part of the theory were developed with arbitrary fields, inner product spaces would push consideration back to just real and complex vector spaces.

You probably cannot cover everything in this book in one semester. Going through all the material in the first seven or eight chapters during a one-semester course may require a rapid pace. If you must reach Chapter 9, then consider skipping the material on quotient spaces in Section 3E, skipping Section 3F on duality (unless you intend to cover tensor products in Section 9D), covering Chapter 4 on polynomials in a half hour, skipping Section 5E on commuting operators, and skipping the subsection in Section 6C on the pseudoinverse.

A goal more important than teaching any particular theorem is to develop in students the ability to understand and manipulate the objects of linear algebra. Mathematics can be learned only by doing. Fortunately, linear algebra has many good homework exercises. When teaching this course, during each class I usually assign as homework several of the exercises, due the next class. Going over the homework might take up significant time in a typical class.

Some of the exercises are intended to lead curious students into important topics beyond what might usually be included in a basic second course in linear algebra.

The author's top ten

Listed below are the author's ten favorite results in the book, in order of their appearance in the book. Students who leave your course with a good understanding of these crucial results will have an excellent foundation in linear algebra.

- any two bases of a vector space have the same length (2.34)
- fundamental theorem of linear maps (3.21)
- existence of eigenvalues if $\mathbf{F} = \mathbf{C}$ (5.19)
- upper-triangular form always exists if $\mathbf{F} = \mathbf{C}$ (5.47)
- Cauchy–Schwarz inequality (6.14)
- Gram–Schmidt procedure (6.32)
- spectral theorem (7.29 and 7.31)
- singular value decomposition (7.70)
- generalized eigenspace decomposition theorem when $\mathbf{F} = \mathbf{C}$ (8.22)
- dimension of alternating n-linear forms on V is 1 if $\dim V = n$ (9.37)

בספר זה לרוב אנו מפתחים את האלגברה הליניארית עבור מרחבים וקטוריים ממשיים ומרוכבים בו זמנית, בכך ש־ $\mathbf{F}$ מסמן את המספרים הממשיים או את המספרים המרוכבים. אם קיימת העדפה שלכם או של הסטודנטים לחשוב על $\mathbf{F}$ כעל שדה שרירותי, ראו את ההערות שבסוף יחידה א1. אני מעדיף להימנע משדות שרירותיים בשלב זה, משום שהם דורשים תוספת הפשטה מבלי להוביל למסקנות חדשות כלשהן באלגברה ליניארית. כמו כן, סטודנטים חשים בנוח רב יותר לחשוב על פולינומים כעל פונקציות, ולא כעל האובייקטים הפורמליים יותר הנדרשים לפולינומים עם מקדמים בשדות סופיים. לבסוף, גם אילו פותח חלקה הראשון של התאוריה עם שדות שרירותיים, מרחבי המכפלה הפנימית היו מחזירים את הדיון למרחבים וקטוריים ממשיים ומרוכבים בלבד.

קרוב לוודאי שלא תוכלו לכסות את כל הספר בסמסטר אחד. מעבר על כל החומר שבשבעת או שמונת הפרקים הראשונים בקורס בן סמסטר אחד עשוי להכתיב קצב מזורז. אם עליכם להגיע לפרק ט, שקלו לדלג על החומר בנושא מרחבי מנה ביחידה ג5, לדלג על יחידה ג6 על נושא הדואליות (אלא אם בכוונתכם לכסות מכפלות טנזוריות ביחידה ט4), לכסות את פרק ד על פולינומים בחצי שעה, לדלג על יחידה ה5 בנושא אופרטורים מתחלפים, ולדלג על הפסידו־הופכי ביחידה ו3.

מטרה חשובה יותר מהוראת משפט מסוים זה או אחר היא לפתח אצל הסטודנטים את היכולת להבין את האובייקטים של האלגברה הליניארית ולבצע בהם מניפולציות. מתמטיקה נלמדת רק מתוך עשייה. למרבה המזל, באלגברה ליניארית יש תרגילים טובים ורבים. כשאני מלמד קורס זה, אני נוהג לתת בכל שיעור כמה מהתרגילים כשיעורי בית, להגשה בשיעור הבא. מעבר על שיעורי הבית עשוי לתפוס זמן ניכר מהשיעור.

חלק מהתרגילים נועדו להוביל סטודנטים סקרנים אל נושאים חשובים מעבר למה שנכלל בדרך כלל בקורס שני בסיסי באלגברה ליניארית.

העשירייה המובילה של המחבר

להלן עשר מסקנות בספר החביבות ביותר על המחבר, לפי סדר הופעתן. סטודנטים שיסיימו את הקורס שלכם עם הבנה טובה של מסקנות מכריעות אלה ייצאו מהקורס כשבאמתחתם יסודות איתנים באלגברה ליניארית.

- לכל שני בסיסים של מרחב וקטורי מממד סופי אותו האורך (2.34)
- המשפט היסודי של ההעתקות הליניאריות (3.21)
- קיומם של ערכים עצמיים אם $\mathbf{F} = \mathbf{C}$ ראו (5.19)
- צורה משולשית עליונה תמיד קיימת אם $\mathbf{F} = \mathbf{C}$ ראו (5.47)
- אי־שוויון קושי־שוורץ (6.14)
- תהליך גרם־שמידט (6.32)
- משפט הפירוק הספקטרלי (7.29 ו־ 7.31)
- הפירוק לערכים סינגולריים (7.70)
- משפט הפירוק למרחבים עצמיים מוכללים כאשר $\mathbf{F} = \mathbf{C}$ ראו (8.22)
- ממד מרחב התבניות ה־ n־ליניאריות המתחלפות על V הוא 1 כאשר $\dim V = n$ ראו (9.37)

Major improvements and additions for the fourth edition

- Over 250 new exercises and over 70 new examples.
- Increasing use of the minimal polynomial to provide cleaner proofs of multiple results, including necessary and sufficient conditions for an operator to have an upper-triangular matrix with respect to some basis (see Section 5C), necessary and sufficient conditions for diagonalizability (see Section 5D), and the real spectral theorem (see Section 7B).
- New section on commuting operators (see Section 5E).
- New subsection on pseudoinverse (see Section 6C).
- New subsections on QR factorization/Cholesky factorization (see Section 7D).
- Singular value decomposition now done for linear maps from an inner product space to another (possibly different) inner product space, rather than only dealing with linear operators from an inner product space to itself (see Section 7E).
- Polar decomposition now proved from singular value decomposition, rather than in the opposite order; this has led to cleaner proofs of both the singular value decomposition (see Section 7E) and the polar decomposition (see Section 7F).
- New subsection on norms of linear maps on finite-dimensional inner product spaces, using the singular value decomposition to avoid even mentioning supremum in the definition of the norm of a linear map (see Section 7F).
- New subsection on approximation by linear maps with lower-dimensional range (see Section 7F).
- New elementary proof of the important result that if T is an operator on a finite-dimensional complex vector space V, then there exists a basis of V consisting of generalized eigenvectors of T (see 8.9).
- New Chapter 9 on multilinear algebra, including bilinear forms, quadratic forms, multilinear forms, and tensor products. Determinants now are defined using a basis-free approach via alternating multilinear forms.
- New formatting to improve the student-friendly appearance of the book. For example, the definition and result boxes now have rounded corners instead of right-angle corners, for a gentler look. The main font size has been reduced from 11 point to 10.5 point.

Please check the website below for additional links and information about the book. Your suggestions, comments, and corrections are most welcome.

Best wishes for teaching a successful linear algebra class!

Sheldon Axler
San Francisco State University

website: https://linear.axler.net
e-mail: linear@axler.net

Contact the author, or Springer if the author is not available, for permission for translations or other commercial reuse of the contents of this book.

שיפורים מרכזיים ותוספות במהדורה הרביעית

- נוספו מעל 250 תרגילים חדשים ומעל 70 דוגמאות חדשות.
- הגדלת השימוש בפולינום המינימלי לקבלת הוכחות נקיות יותר של מספר מסקנות, בהן תנאים הכרחיים ומספיקים לכך שלאופרטור תהיה מטריצה משולשית עליונה ביחס לבסיס כלשהו (ראו יחידה ה3), תנאים הכרחיים ומספיקים ללכסינות (ראו יחידה ה4), ומשפט הפירוק הספקטרלי הממשי (ראו יחידה ז2).
- יחידה חדשה על אופרטורים מתחלפים (ראו יחידה ה5).
- תת־יחידה חדשה על הפסידו־הופכי (ראו יחידה ו3).
- תתי־יחידות חדשות על פירוק QR ופירוק שולסקי (ראו יחידה ז4).
- הפירוק לערכים סינגולריים נעשה כעת עבור העתקות ליניאריות ממרחב מכפלה פנימית אל מרחב מכפלה פנימית אחר (העשוי להיות שונה), ולא רק עבור אופרטורים ליניאריים ממרחב מכפלה פנימית אל עצמו (ראו יחידה ז5).
- הפירוק הפולרי מוכח כעת מתוך הפירוק לערכים סינגולריים, ולא בסדר ההפוך; הדבר הוביל להוכחות נקיות יותר הן של הפירוק לערכים סינגולריים (ראו יחידה ז5) והן של הפירוק הפולרי (ראו יחידה ז6).
- תת־יחידה חדשה על נורמות של העתקות ליניאריות מעל מרחבי מכפלה פנימית מממד סופי, תוך שימוש בפירוק לערכים סינגולריים כדי להימנע אפילו מאזכור הסופרמום בהגדרת הנורמה של העתקה ליניארית (ראו יחידה ז6).
- תת־יחידה חדשה על קירוב באמצעות העתקות ליניאריות בעלות טווח מממד נמוך יותר (ראו יחידה ז6).
- הוכחה אלמנטרית חדשה למסקנה החשובה שלפיה אם T הוא אופרטור מעל מרחב וקטורי מרוכב V מממד סופי, אז קיים בסיס של V המורכב מווקטורים עצמיים מוכללים של T (ראו 8.9).
- פרק חדש ט על אלגברה מולטי־ליניארית, הכולל תבניות בילינאריות, תבניות ריבועיות, תבניות מולטי־ליניאריות, ומכפלות טנזוריות. הדטרמיננטות מוגדרות כעת בגישה ללא תלות בבסיס, באמצעות תבניות מולטי־ליניאריות מתחלפות.
- עיצוב חדש לשיפור המראה הידידותי לסטודנט של הספר. לדוגמה, לתיבות ההגדרות והמסקנות יש כעת פינות מעוגלות במקום פינות ישרות, למראה רך יותר. גודל הגופן הראשי הוקטן מ־11 נקודות ל־10.5 נקודות.

אנא בקרו באתר שלהלן לקבלת קישורים ומידע נוסף על הספר. הצעותיכם, הערותיכם, ותיקוניכם יתקבלו בברכה רבה.

איחולים לבביים להוראת קורס מוצלח באלגברה ליניארית!

שלדון אקסלר
אוניברסיטת סן פרנסיסקו סטייט

אתר: https://linear.axler.net
מייל: linear@axler.net

צרו קשר עם המחבר, או הוצאת הספרים שפרינגר אם המחבר אינו זמין, לקבלת רשות לתרגום או לכל שימוש מסחרי אחר בתכניס שבספר.

Acknowledgments

I owe a huge intellectual debt to all the mathematicians who created linear algebra over the past two centuries. The results in this book belong to the common heritage of mathematics. A special case of a theorem may first have been proved long ago, then sharpened and improved by many mathematicians in different time periods. Bestowing proper credit on all contributors would be a difficult task that I have not undertaken. In no case should the reader assume that any result presented here represents my original contribution.

Many people helped make this a better book. The three previous editions of this book were used as a textbook at over 375 universities and colleges around the world. I received thousands of suggestions and comments from faculty and students who used the book. Many of those suggestions led to improvements in this edition. The manuscript for this fourth edition was class tested at 30 universities. I am extremely grateful for the useful feedback that I received from faculty and students during this class testing.

The long list of people who should be thanked for their suggestions would fill up many pages. Lists are boring to read. Thus to represent all contributors to this edition, I will mention only Noel Hinton, a graduate student at Australian National University, who sent me more suggestions and corrections for this fourth edition than anyone else. To everyone who contributed suggestions, let me say how truly grateful I am to all of you. Many many thanks!

I thank Springer for providing me with help when I needed it and for allowing me the freedom to make the final decisions about the content and appearance of this book. Special thanks to the two terrific mathematics editors at Springer who worked with me on this project—Loretta Bartolini during the first half of my work on the fourth edition, and Elizabeth Loew during the second half of my work on the fourth edition. I am deeply indebted to David Kramer, who did a magnificent job of copyediting and prevented me from making many mistakes.

Extra special thanks to my fantastic partner Carrie Heeter. Her understanding and encouragement enabled me to work intensely on this new edition. Our wonderful cat Moon, whose picture appears on the *About the Author* page, provided sweet breaks throughout the writing process. Moon died suddenly due to a blood clot as this book was being finished. We are grateful for five precious years with him.

Sheldon Axler

דברי תודה

אני חב חוב אינטלקטואלי עצום לכל המתמטיקאים שיצרו ופיתחו את האלגברה הליניארית במהלך מאתיים השנים האחרונות. המסקנות המובאות בספר זה שייכות למורשת המשותפת של המתמטיקה. ייתכן שמקרה פרטי של משפט הוכח לראשונה לפני זמן רב, ולאחר מכן לוטש ושופר על ידי מתמטיקאים רבים בתקופות זמן שונות. מתן ההכרה הראויה לכל התורמים יהיה משימה מורכבת שלא לקחתי על עצמי. בשום מקרה אל לקורא להניח שמסקנה כלשהי המוצגת כאן מהווה תרומה מקורית שלי.

אנשים רבים עזרו להפוך את הספר הזה לטוב יותר. שלוש המהדורות הקודמות של ספר זה שימשו כספר לימוד בלמעלה מ־375 אוניברסיטאות ומכללות ברחבי העולם. קיבלתי אלפי הצעות והערות מחברי סגל ומסטודנטים שהשתמשו בספר. רבות מההצעות הללו הובילו לשיפורים במהדורה הנוכחית. כתב היד של מהדורה רביעית זו נוסה בכיתות לימוד ב־30 אוניברסיטאות. אני אסיר תודה עמוקה על המשוב המועיל שקיבלתי מחברי סגל ומסטודנטים במהלך התנסויות אלו.

הרשימה הארוכה של האנשים שיש להודות להם על הצעותיהם הייתה ממלאת עמודים רבים. רשימות הן משעממות לקריאה. לכן, כדי לייצג את כלל התורמים למהדורה זו, אזכיר רק את נואל הינטון (Noel Hinton), סטודנט לתארים מתקדמים באוניברסיטה הלאומית של אוסטרליה (Australian National University), ששלח לי יותר הצעות ותיקונים למהדורה הרביעית מכל אדם אחר. לכל מי שתרם ממרצו והציע הצעות, הרשו לי לומר עד כמה אני באמת אסיר תודה לכולכם. תודה רבה רבה!

אני מודה להוצאת שפרינגר (Springer) על העזרה שהעניקה לי כשהזדקקתי לה, ועל כך שאפשרה לי את החופש לקבל את ההחלטות הסופיות לגבי התוכן והמראה של ספר זה. תודה מיוחדת לשתי עורכות המתמטיקה הנפלאות בשפרינגר שעבדו איתי על פרויקט זה—לורטה ברטוליני (Loretta Bartolini) במהלך המחצית הראשונה של עבודתי על המהדורה הרביעית, ואליזבת לו (Elizabeth Loew) במהלך המחצית השנייה. אני חב תודה עמוקה לדיוויד קריימר (David Kramer), שעשה עבודה נהדרת בעריכת הלשון ומנע ממני לעשות טעויות רבות.

תודה מיוחדת וגדולה לבת זוגי המדהימה, קארי היטר (Carrie Heeter). ההבנה והעידוד שלה אפשרו לי לעבוד באינטנסיביות על מהדורה חדשה זו. החתול הנפלא שלנו, מון (Moon), שתמונתו מופיעה בעמוד על המחבר, סיפק הפוגות מתוקות לאורך תהליך הכתיבה. מון מת בפתאומיות כתוצאה מקריש דם עם סיום כתיבת הספר. אנו אסירי תודה על חמש שנים יקרות מפז במחיצתו.

שלדון אקסלר

דבר המתרגם

טרם אעסוק בהכרעות התרגום עצמן, חובה עלי להודות לארבעה מתמטיקאים המלווים אותי בשביליה המופלאים של אספקלריית סודות היקום, היא המתמטיקה.

ארבעה אישים נכבדים אלו השפיעו ועודם משפיעים השפעה ישירה, בהרצאותיהם, פרסומיהם ואצל שלושה מהם – עד כדי מענה אישי להתכתבויות, למרות שמעולם לא נפגשנו ולמרות תפקידיהם ועיסוקיהם הרבים.

פותחין בכבוד אכסניא ולפיכך בראש ובראשונה אודה למחבר, הפרופסור שלדון אקסלר (Sheldon Axler) מאוניברסיטת סן פרנסיסקו סטייט, אשר נתן בי אמון והסכים, אחרי היסוסים לא מעטים, להפקיד את מלאכת התרגום לעברית של ספרו הפופולרי בידיי.

את הניצוץ שהפך ללפיד, הצית בקרבי הפרופסור סטיבן סטרוגץ (Steven Strogatz) מאוניברסיטת קורנל בהרצאותיו ובספרו Infinite Powers שאותו טרם סיימתי, שכן כשהגעתי בערך למחציתו של הספר החלטתי להקדיש את זמני ללימוד מתמטיקה בצורה רצינית ומסודרת.

כשניגשתי להכין עצמי ללימוד מתמטיקה, נתקלתי במפעלו המהפכני של הדוקטור אביב צנזור מהטכניון, הוא אשר בעזרת טושים, לוח, כריזמה המשולבת בצניעות וענווה ראויות להערצה, מקצוענות מרשימה וצלם, פתח את השערים וסלל את השבילים והנגיש לכל מבקשי דעת, מקורסי המבוא ועד לקורסי המשך מתקדמים בעברית ובאנגלית. מול הרצאות מצולמות אלו ישבתי כתלמיד ממושמע, כתבתי סיכומים ותרגלתי, תוך שאני מנכס לעצמי את תואר תלמידו למרות שמעולם לא נכחתי בהרצאותיו.

ייבדלו המוזכרים לעיל לחיים ארוכים בבריאות, שלווה ונחת, אציין את המורה הדגול והנערץ, נעים ההליכות אף הוא מהטכניון, הפרופסור דוד צילג, זכרו לברכה, ששני הקורסים שלו, אלגברה א׳ ואלגברה 1מ׳, היו לחם חוקי במשך כשנה. בסגנונו השנון וההומוריסטי המשולב בפשטות, צניעות ואנושיות טבעית ונטולת מסכות, הרצה את חומר הלימוד באופן שריתק ומרתק אלפי סטודנטים, גם כשחלפו שנים מאז ההקלטות. חדי העין יבחינו פה ושם בבחירות תרגום התואמות את בחירותיו.

מלאכת התרגום דורשת הכרעות, לרוב מנומקות אך לעיתים שרירותיות, בייחוד במקרים שבהם אין תמימות דעים בקרב המתמטיקאים דוברי העברית (העתקה או טרנספורמציה, אסוציאטיביות לעומת קיבוציות), או כאשר האקדמיה ללשון הותירה את ההכרעה בידי הדובר (תתי־מרחב לעומת תתי־מרחבים, ליניארית לעומת לינארית, אינסופי לעומת אין־סופי ודומיהם).

במקומות לא מעטים בחר המחבר במודע, בשפת המקור, במונחים שאינם הנפוצים ביותר (חלקם מצוינים בתיבות אפורות לאורכו של הספר). בחירות אלו כובדו במלואן גם כאשר בז׳רגון המקצועי המקומי נפוצה בחירה שונה (רשימת וקטורים לעומת n־יה או קבוצה, סימטריית צמוד לעומת הרמיטיות ודומיהם), זאת מתוך נאמנות למקור ושאיפה כי התרגום העברי ייתן ביטוי מלא לרוח הספר והסופר.

השתדלתי ככל שידי משגת כי התרגום יהיה חד־חד־ערכי, כלומר, ככלל, מונח נתון יתורגם באופן אחיד לאורכו של הספר גם כאשר האפשרויות השונות שקולות. אני ממתין ומצפה להערות, תיקוני שגיאות והמלצות שאותם נשלב במהדורה הדיגיטלית המתעדכנת תכופות ואף במהדורה המודפסת בבוא העת.

מוקדש באהבה ליובל, אביגיל, רון, ניקולאי רפאל ואימהותיהם המסורות.

צורי בר יוחאי

Chapter 1

Vector Spaces

Linear algebra is the study of linear maps on finite-dimensional vector spaces. Eventually we will learn what all these terms mean. In this chapter we will define vector spaces and discuss their elementary properties.

In linear algebra, better theorems and more insight emerge if complex numbers are investigated along with real numbers. Thus we will begin by introducing the complex numbers and their basic properties.

We will generalize the examples of a plane and of ordinary space to $\mathbf{R}^n$ and $\mathbf{C}^n$, which we then will generalize to the notion of a vector space. As we will see, a vector space is a set with operations of addition and scalar multiplication that satisfy natural algebraic properties.

Then our next topic will be subspaces, which play a role for vector spaces analogous to the role played by subsets for sets. Finally, we will look at sums of subspaces (analogous to unions of subsets) and direct sums of subspaces (analogous to unions of disjoint sets).

Pierre Louis Dumesnil, Nils Forsberg

René Descartes explaining his work to Queen Christina of Sweden. Vector spaces are a generalization of the description of a plane using two coordinates, as published by Descartes in 1637.

פרק א

מרחבים וקטוריים

אלגברה ליניארית היא תורת ההעתקות הליניאריות מעל מרחבים וקטוריים מממד סופי. עם הזמן נכיר את המשמעות של כל אחד מהמונחים הללו. בפרק זה נגדיר מרחבים וקטוריים ונדון במאפייניהם היסודיים.

באלגברה ליניארית, חקירת המספרים המרוכבים לצד המספרים הממשיים מובילה להבנה מעמיקה יותר, ולניסוח משפטים באופן מיטבי. לפיכך נפתח בהיכרות עם המספרים המרוכבים ומאפייניהם הבסיסיים.

אנו נכליל את הדוגמאות של מישור ומרחב רגיל ל־ $\mathbf{R}^n$ ו־ $\mathbf{C}^n$, שאותם נכליל בהמשך לרעיון של מרחב וקטורי. כפי שנראה, מרחב וקטורי הוא קבוצה עם פעולות חיבור וכפל בסקלר המקיימות תכונות אלגבריות טבעיות.

לאחר מכן הנושא שבו נעסוק יהיה תתי־מרחב, שתפקידם ביחס למרחבים וקטוריים דומה לתפקיד שממלאות תתי־קבוצה עבור קבוצות. לבסוף, נלמד על סכום תתי־מרחב (באנלוגיה לאיחוד תתי־קבוצה), וסכום ישר של תתי־מרחב (באנלוגיה לאיחוד קבוצות זרות).

Pierre Louis Dumesnil, Nils Forsberg

רנה דקארט מתאר את עבודתו לכריסטינה מלכת שוודיה.
מרחבים וקטוריים הם הכללה של תיאור המישור באמצעות שתי קואורדינטות,
כפי שפרסם דקארט בשנת 1637.

1A $\mathbf{R}^n$ and $\mathbf{C}^n$

Complex Numbers

You should already be familiar with basic properties of the set $\mathbf{R}$ of real numbers. Complex numbers were invented so that we can take square roots of negative numbers. The idea is to assume we have a square root of -1, denoted by i, that obeys the usual rules of arithmetic. Here are the formal definitions.

1.1 definition: *complex numbers*, $\mathbf{C}$

- A *complex number* is an ordered pair (a, b), where $a, b \in \mathbf{R}$, but we will write this as $a + bi$.
- The set of all complex numbers is denoted by $\mathbf{C}$:

$$\mathbf{C} = \{a + bi : a, b \in \mathbf{R}\}.$$

- *Addition* and *multiplication* on $\mathbf{C}$ are defined by

$$(a + bi) + (c + di) = (a + c) + (b + d)i,$$
$$(a + bi)(c + di) = (ac - bd) + (ad + bc)i;$$

here $a, b, c, d \in \mathbf{R}$.

If $a \in \mathbf{R}$, we identify $a + 0i$ with the real number a. Thus we think of $\mathbf{R}$ as a subset of $\mathbf{C}$. We usually write $0 + bi$ as just bi, and we usually write $0 + 1i$ as just i.

The symbol i was first used to denote $\sqrt{-1}$ by Leonhard Euler in 1777.

To motivate the definition of complex multiplication given above, pretend that we knew that $i^2 = -1$ and then use the usual rules of arithmetic to derive the formula above for the product of two complex numbers. Then use that formula to verify that we indeed have

$$i^2 = -1.$$

Do not memorize the formula for the product of two complex numbers—you can always rederive it by recalling that $i^2 = -1$ and then using the usual rules of arithmetic (as given by 1.3). The next example illustrates this procedure.

1.2 example: *complex arithmetic*

The product $(2 + 3i)(4 + 5i)$ can be evaluated by applying the distributive and commutative properties from 1.3:

$$\begin{aligned}(2 + 3i)(4 + 5i) &= 2 \cdot (4 + 5i) + (3i)(4 + 5i)\\ &= 2 \cdot 4 + 2 \cdot 5i + 3i \cdot 4 + (3i)(5i)\\ &= 8 + 10i + 12i - 15\\ &= -7 + 22i.\end{aligned}$$

או $\mathbf{R}^n$ ו־ $\mathbf{C}^n$

מספרים מרוכבים

אתם כבר אמורים להכיר את המאפיינים הבסיסיים של קבוצת המספרים הממשיים $\mathbf{R}$. מספרים מרוכבים הומצאו כדי שנוכל להוציא שורשים ריבועיים של מספרים שליליים. הרעיון הוא להניח שיש שורש ריבועי של -1, המסומן ב־ i, והוא מציית לכללי האריתמטיקה הרגילים. להלן ההגדרות הרשמיות.

1.1 הגדרה: מספרים מרוכבים, $\mathbf{C}$

- מספר מרוכב הוא זוג סדור (a, b) כאשר $a, b \in \mathbf{R}$, אך אנו נרשום אותו באופן הבא: $a + bi$.
- קבוצת כל המספרים המרוכבים מסומנת ב־ $\mathbf{C}$:

$$\mathbf{C} = \{a + bi : a, b \in \mathbf{R}\}.$$

- חיבור וכפל מעל $\mathbf{C}$ מוגדרים על ידי

$$(a + bi) + (c + di) = (a + c) + (b + d)i,$$
$$(a + bi)(c + di) = (ac - bd) + (ad + bc)i;$$

כאשר $a, b, c, d \in \mathbf{R}$.

אם $a \in \mathbf{R}$, אנו מזהים את $a + 0i$ כמספר הממשי a. לפיכך אנו מחשיבים את $\mathbf{R}$ כתת־קבוצה של $\mathbf{C}$. בדרך כלל נכתוב bi רק כ־ $0 + bi$, ואת $0 + 1i$ רק כ־ i.

השימוש בסמל i כסימון $\sqrt{-1}$ נעשה לראשונה על ידי לאונרד אוילר בשנת 1777.

להצדקת ההגדרה של כפל מרוכבים שהבאנו לעיל, העמידו פנים שידוע לנו מכבר כי $i^2 = -1$ ואז השתמשו בכללי האריתמטיקה הרגילים כדי לגזור את הנוסחה שלעיל לכפל שני מספרים מרוכבים. כעת השתמשו בנוסחה זו כדי לוודא שאכן קיבלנו

$$i^2 = -1.$$

אין צורך לשנן את הנוסחה לכפל שני מספרים מרוכבים—תמיד נוכל להגיע אליה כשנזכור ש־ $i^2 = -1$ ואז נשתמש בכללי האריתמטיקה הרגילים (כפי שמוצג ב־ 1.3). הדוגמה הבאה ממחישה תהליך זה.

1.2 דוגמה: אריתמטיקה של מרוכבים

את המכפלה $(2 + 3i)(4 + 5i)$ נוכל לחשב תוך הפעלת תכונות הדיסטריבוטיביות והקומוטטיביות המפורטות ב־ 1.3:

$$\begin{aligned}(2 + 3i)(4 + 5i) &= 2 \cdot (4 + 5i) + (3i)(4 + 5i)\\ &= 2 \cdot 4 + 2 \cdot 5i + 3i \cdot 4 + (3i)(5i)\\ &= 8 + 10i + 12i - 15\\ &= -7 + 22i.\end{aligned}$$

Our first result states that complex addition and complex multiplication have the familiar properties that we expect.

1.3 *properties of complex arithmetic*

commutativity
$\alpha + \beta = \beta + \alpha$ and $\alpha\beta = \beta\alpha$ for all $\alpha, \beta \in \mathbf{C}$.

associativity
$(\alpha + \beta) + \lambda = \alpha + (\beta + \lambda)$ and $(\alpha\beta)\lambda = \alpha(\beta\lambda)$ for all $\alpha, \beta, \lambda \in \mathbf{C}$.

identities
$\lambda + 0 = \lambda$ and $\lambda 1 = \lambda$ for all $\lambda \in \mathbf{C}$.

additive inverse
For every $\alpha \in \mathbf{C}$, there exists a unique $\beta \in \mathbf{C}$ such that $\alpha + \beta = 0$.

multiplicative inverse
For every $\alpha \in \mathbf{C}$ with $\alpha \neq 0$, there exists a unique $\beta \in \mathbf{C}$ such that $\alpha\beta = 1$.

distributive property
$\lambda(\alpha + \beta) = \lambda\alpha + \lambda\beta$ for all $\lambda, \alpha, \beta \in \mathbf{C}$.

The properties above are proved using the familiar properties of real numbers and the definitions of complex addition and multiplication. The next example shows how commutativity of complex multiplication is proved. Proofs of the other properties above are left as exercises.

1.4 example: *commutativity of complex multiplication*

To show that $\alpha\beta = \beta\alpha$ for all $\alpha, \beta \in \mathbf{C}$, suppose

$$\alpha = a + bi \quad \text{and} \quad \beta = c + di,$$

where $a, b, c, d \in \mathbf{R}$. Then the definition of multiplication of complex numbers shows that

$$\begin{aligned}\alpha\beta &= (a + bi)(c + di)\\ &= (ac - bd) + (ad + bc)i\end{aligned}$$

and

$$\begin{aligned}\beta\alpha &= (c + di)(a + bi)\\ &= (ca - db) + (cb + da)i.\end{aligned}$$

The equations above and the commutativity of multiplication and addition of real numbers show that $\alpha\beta = \beta\alpha$.

המסקנה הראשונה שלנו קובעת כי לפעולת החיבור במרוכבים ולפעולת הכפל במרוכבים יש את התכונות המוכרות שלהן אנו מצפים.

1.3 תכונות של אריתמטיקת מרוכבים

קומוטטיביות (חילופיות)
$\alpha + \beta = \beta + \alpha$ ו- $\alpha\beta = \beta\alpha$ לכל $\alpha, \beta \in \mathbf{C}$.

אסוציאטיביות (קיבוציות)
$(\alpha + \beta) + \lambda = \alpha + (\beta + \lambda)$ ו- $(\alpha\beta)\lambda = \alpha(\beta\lambda)$ לכל $\alpha, \beta, \lambda \in \mathbf{C}$.

איברי זהות (יחידה)
$\lambda + 0 = \lambda$ ו- $\lambda 1 = \lambda$ לכל $\lambda \in \mathbf{C}$.

איבר נגדי לחיבור
לכל $\alpha \in \mathbf{C}$, קיים איבר יחיד $\beta \in \mathbf{C}$ כך ש- $\alpha + \beta = 0$.

איבר הופכי
לכל $\alpha \in \mathbf{C}$ כאשר $\alpha \neq 0$, קיים איבר יחיד $\beta \in \mathbf{C}$ כך ש- $\alpha\beta = 1$.

תכונת הדיסטריבוטיביות (הפילוג)
$\lambda(\alpha + \beta) = \lambda\alpha + \lambda\beta$ לכל $\lambda, \alpha, \beta \in \mathbf{C}$.

התכונות לעיל מוכחות תוך שימוש בתכונות המוכרות של מספרים ממשיים ובהגדרות של פעולות הכפל והחיבור של מספרים מרוכבים. הדוגמה הבאה מראה איך מוכיחים את הקומוטטיביות של כפל מרוכבים. ההוכחות של יתר התכונות לעיל מושארות כתרגילים.

1.4 דוגמה: קומוטטיביות של כפל מרוכבים

כדי להראות כי $\alpha\beta = \beta\alpha$ לכל $\alpha, \beta \in \mathbf{C}$, נניח ש-

$$\alpha = a + bi \quad \text{וגם} \quad \beta = c + di,$$

כאשר $a, b, c, d \in \mathbf{R}$. אזי, הגדרת פעולת הכפל של מספרים מרוכבים מראה ש-

$$\begin{aligned}\alpha\beta &= (a + bi)(c + di)\\ &= (ac - bd) + (ad + bc)i\end{aligned}$$

וגם

$$\begin{aligned}\beta\alpha &= (c + di)(a + bi)\\ &= (ca - db) + (cb + da)i.\end{aligned}$$

המשוואות לעיל והקומוטטיביות של כפל וחיבור של מספרים ממשיים מראות ש- $\alpha\beta = \beta\alpha$.

Next, we define the additive and multiplicative inverses of complex numbers, and then use those inverses to define subtraction and division operations with complex numbers.

1.5 definition: $-\alpha$, *subtraction*, $1/\alpha$, *division*

Suppose $\alpha, \beta \in \mathbf{C}$.

- Let $-\alpha$ denote the additive inverse of α. Thus $-\alpha$ is the unique complex number such that
$$\alpha + (-\alpha) = 0.$$
- *Subtraction* on $\mathbf{C}$ is defined by
$$\beta - \alpha = \beta + (-\alpha).$$
- For $\alpha \neq 0$, let $1/\alpha$ and $\frac{1}{\alpha}$ denote the multiplicative inverse of α. Thus $1/\alpha$ is the unique complex number such that
$$\alpha(1/\alpha) = 1.$$
- For $\alpha \neq 0$, *division* by α is defined by
$$\beta/\alpha = \beta(1/\alpha).$$

So that we can conveniently make definitions and prove theorems that apply to both real and complex numbers, we adopt the following notation.

1.6 notation: $\mathbf{F}$

Throughout this book, $\mathbf{F}$ stands for either $\mathbf{R}$ or $\mathbf{C}$.

Thus if we prove a theorem involving $\mathbf{F}$, we will know that it holds when $\mathbf{F}$ is replaced with $\mathbf{R}$ and when $\mathbf{F}$ is replaced with $\mathbf{C}$.

*The letter $\mathbf{F}$ is used because $\mathbf{R}$ and $\mathbf{C}$ are examples of what are called **fields**.*

Elements of $\mathbf{F}$ are called *scalars*. The word "scalar" (which is just a fancy word for "number") is often used when we want to emphasize that an object is a number, as opposed to a vector (vectors will be defined soon).

For $\alpha \in \mathbf{F}$ and m a positive integer, we define α^m to denote the product of α with itself m times:
$$\alpha^m = \underbrace{\alpha \cdots \alpha}_{m \text{ times}}.$$
This definition implies that
$$(\alpha^m)^n = \alpha^{mn} \quad \text{and} \quad (\alpha\beta)^m = \alpha^m\beta^m$$
for all $\alpha, \beta \in \mathbf{F}$ and all positive integers m, n.

בשלב הבא, אנו מגדירים את האיבר הנגדי לחיבור והאיבר ההופכי לכפל של מספרים מרוכבים, ואז משתמשים באיברים ההופכי והנגדי הללו כדי להגדיר פעולות חיסור וחילוק של מספרים מרוכבים.

1.5 הגדרה: $-\alpha$, חיסור, $1/\alpha$, חילוק

נניח ש־ $\alpha, \beta \in \mathbf{C}$.

- נסמן ב־ $-\alpha$ את האיבר הנגדי לחיבור של α. לפיכך $-\alpha$ הוא המספר המרוכב היחיד כך ש־
$$\alpha + (-\alpha) = 0.$$
- פעולת החיסור מעל $\mathbf{C}$ מוגדרת על ידי
$$\beta - \alpha = \beta + (-\alpha).$$
- עבור $\alpha \neq 0$, נסמן ב־ $1/\alpha$ ו־ $\frac{1}{\alpha}$ את האיבר ההופכי של α. לפיכך $1/\alpha$ הוא המספר המרוכב היחיד כך ש־
$$\alpha(1/\alpha) = 1.$$
- עבור $\alpha \neq 0$, פעולת החילוק ב־ α מוגדרת על ידי
$$\beta/\alpha = \beta(1/\alpha).$$

כדי שנוכל להגדיר בנוחות ולהוכיח משפטים החלים על שתי קבוצות המספרים, הממשיים והמרוכבים כאחד, נאמץ את הסימון הבא:

1.6 סימון: $\mathbf{F}$

לאורכו של ספר זה, $\mathbf{F}$ מסמן את $\mathbf{R}$ או $\mathbf{C}$.

השימוש באות $\mathbf{F}$ נובע מהעובדה ש־ $\mathbf{R}$ ו־ $\mathbf{C}$ הם דוגמאות למה שאנו קוראים **שדות**.

לפיכך, אם הוכחנו משפט העוסק ב־ $\mathbf{F}$, נדע כי הוא מתקיים בין אם נציב $\mathbf{R}$ במקום $\mathbf{F}$ ובין אם נציב $\mathbf{C}$ במקום $\mathbf{F}$.

האיברים ב־ $\mathbf{F}$ נקראים סקלרים. נשתמש במילה "סקלר" (שאינה אלא מילה מסוגננת עבור "מספר") לעיתים קרובות כאשר נרצה להדגיש כי האובייקט המדובר הוא מספר, בניגוד לווקטור (וקטורים יוגדרו בקרוב).

עבור $\alpha \in \mathbf{F}$ ו־ m מספר שלם חיובי, אנו מגדירים את α^m כמכפלת α בעצמו m פעמים:

$$\alpha^m = \underbrace{\alpha \cdots \alpha}_{m\text{־פעמים}}.$$

הגדרה זו גוררת ש־

$$(\alpha^m)^n = \alpha^{mn} \quad \text{וגם} \quad (\alpha\beta)^m = \alpha^m \beta^m$$

לכל $\alpha, \beta \in \mathbf{F}$ ולכל המספרים השלמים החיוביים m, n.

Lists

Before defining $\mathbf{R}^n$ and $\mathbf{C}^n$, we look at two important examples.

1.7 example: $\mathbf{R}^2$ *and* $\mathbf{R}^3$

- The set $\mathbf{R}^2$, which you can think of as a plane, is the set of all ordered pairs of real numbers:
$$\mathbf{R}^2 = \{(x, y) : x, y \in \mathbf{R}\}.$$
- The set $\mathbf{R}^3$, which you can think of as ordinary space, is the set of all ordered triples of real numbers:
$$\mathbf{R}^3 = \{(x, y, z) : x, y, z \in \mathbf{R}\}.$$

To generalize $\mathbf{R}^2$ and $\mathbf{R}^3$ to higher dimensions, we first need to discuss the concept of lists.

1.8 definition: *list, length*

- Suppose n is a nonnegative integer. A *list* of *length* n is an ordered collection of n elements (which might be numbers, other lists, or more abstract objects).
- Two lists are equal if and only if they have the same length and the same elements in the same order.

*Many mathematicians call a list of length n an **n-tuple**.*

Lists are often written as elements separated by commas and surrounded by parentheses. Thus a list of length two is an ordered pair that might be written as (a, b). A list of length three is an ordered triple that might be written as (x, y, z). A list of length n might look like this:

$$(z_1, \dots, z_n).$$

Sometimes we will use the word *list* without specifying its length. Remember, however, that by definition each list has a finite length that is a nonnegative integer. Thus an object that looks like $(x_1, x_2, \dots)$, which might be said to have infinite length, is not a list.

A list of length 0 looks like this: $(\,)$. We consider such an object to be a list so that some of our theorems will not have trivial exceptions.

Lists differ from finite sets in two ways: in lists, order matters and repetitions have meaning; in sets, order and repetitions are irrelevant.

1.9 example: *lists versus sets*

- The lists $(3, 5)$ and $(5, 3)$ are not equal, but the sets $\{3, 5\}$ and $\{5, 3\}$ are equal.
- The lists $(4, 4)$ and $(4, 4, 4)$ are not equal (they do not have the same length), although the sets $\{4, 4\}$ and $\{4, 4, 4\}$ both equal the set $\{4\}$.

רשימות

לפני שנגדיר את $\mathbf{R}^n$ ו־ $\mathbf{C}^n$, נתבונן בשתי דוגמאות חשובות.

1.7 דוגמה: $\mathbf{R}^2$ ו־ $\mathbf{R}^3$

- הקבוצה $\mathbf{R}^2$, שניתן לחשוב עליה כמישור, היא קבוצת כל הזוגות הסדורים של מספרים ממשיים:

$$\mathbf{R}^2 = \{(x, y) : x, y \in \mathbf{R}\}.$$

- הקבוצה $\mathbf{R}^3$, שניתן לחשוב עליה כמרחב הרגיל, היא קבוצת כל השלישיות הסדורות של מספרים ממשיים:

$$\mathbf{R}^3 = \{(x, y, z) : x, y, z \in \mathbf{R}\}.$$

לצורך הכללת $\mathbf{R}^2$ ו־ $\mathbf{R}^3$ לממדים גבוהים יותר, נצטרך תחילה לדבר על הרעיון של רשימות.

1.8 הגדרה: רשימה, אורך

- נניח ש־ n מספר שלם אי־שלילי. רשימה באורך n היא אוסף סדור של n איברים (שייתכן שיהיו מספרים, רשימות אחרות, או אובייקטים מופשטים יותר).
- שתי רשימות הן שוות אם־ורק־אם יש להן אותו אורך ואותם איברים באותו סדר.

לרשימה באורך n מתמטיקאים רבים קוראים **n־יה**.

רשימות נכתבות לעיתים קרובות כאיברים מופרדים בפסיק, ומוקפים בסוגריים. לפיכך רשימה באורך שתיים היא זוג סדור שייתכן שייכתב כ־ (a, b). רשימה באורך שלוש היא שלישייה סדורה שייתכן שתיכתב כ־ (x, y, z). רשימה באורך n עשויה להיראות כך:

$$(z_1, \dots, z_n).$$

לעיתים נשתמש במונח רשימה ללא ציון אורכה. נזכור, עם זאת, כי על פי ההגדרה, לכל רשימה יש אורך סופי שהוא מספר שלם אי־שלילי. לפיכך אובייקט הנראה כך $(x_1, x_2, \dots)$, שאפשר לומר עליו כי אורכו אינסופי, אינו רשימה.

רשימה באורך 0 מסומנת כך: $( )$. אנו מכלילים אובייקט זה בקטגוריית הרשימות כך שחלק ממשפטינו לא יכללו החרגות טריוויאליות.

הרשימות נבדלות מקבוצות סופיות בשני אופנים: ברשימות, הסדר משמעותי, ומספר המופעים של כל איבר משמעותי; בקבוצות, סדר הופעת האיברים והופעות חוזרות אינם רלוונטיים.

1.9 דוגמה: רשימה לעומת קבוצה

- הרשימות $(3, 5)$ ו־ $(5, 3)$ אינן שוות, אך הקבוצות $\{3, 5\}$ ו־ $\{5, 3\}$ שוות.
- הרשימות $(4, 4)$ ו־ $(4, 4, 4)$ אינן שוות (אין להן אותו אורך), אף שהקבוצות $\{4, 4\}$ ו־ $\{4, 4, 4\}$ שתיהן שוות לקבוצה $\{4\}$.

$\mathbf{F}^n$

To define the higher-dimensional analogues of $\mathbf{R}^2$ and $\mathbf{R}^3$, we will simply replace $\mathbf{R}$ with $\mathbf{F}$ (which equals $\mathbf{R}$ or $\mathbf{C}$) and replace the 2 or 3 with an arbitrary positive integer.

1.10 notation: n

Fix a positive integer n for the rest of this chapter.

1.11 definition: $\mathbf{F}^n$, *coordinate*

$\mathbf{F}^n$ is the set of all lists of length n of elements of $\mathbf{F}$:

$$\mathbf{F}^n = \{(x_1, \dots, x_n) : x_k \in \mathbf{F} \text{ for } k = 1, \dots, n\}.$$

For $(x_1, \dots, x_n) \in \mathbf{F}^n$ and $k \in \{1, \dots, n\}$, we say that x_k is the kth *coordinate* of $(x_1, \dots, x_n)$.

If $\mathbf{F} = \mathbf{R}$ and n equals 2 or 3, then the definition above of $\mathbf{F}^n$ agrees with our previous notions of $\mathbf{R}^2$ and $\mathbf{R}^3$.

1.12 example: $\mathbf{C}^4$

$\mathbf{C}^4$ is the set of all lists of four complex numbers:

$$\mathbf{C}^4 = \{(z_1, z_2, z_3, z_4) : z_1, z_2, z_3, z_4 \in \mathbf{C}\}.$$

If $n \geq 4$, we cannot visualize $\mathbf{R}^n$ as a physical object. Similarly, $\mathbf{C}^1$ can be thought of as a plane, but for $n \geq 2$, the human brain cannot provide a full image of $\mathbf{C}^n$. However, even if n is large, we can perform algebraic manipulations in $\mathbf{F}^n$ as easily as in $\mathbf{R}^2$ or $\mathbf{R}^3$. For example, addition in $\mathbf{F}^n$ is defined as follows.

*Read **Flatland: A Romance of Many Dimensions**, by Edwin A. Abbott, for an amusing account of how $\mathbf{R}^3$ would be perceived by creatures living in $\mathbf{R}^2$. This novel, published in 1884, may help you imagine a physical space of four or more dimensions.*

1.13 definition: *addition in* $\mathbf{F}^n$

Addition in $\mathbf{F}^n$ is defined by adding corresponding coordinates:

$$(x_1, \dots, x_n) + (y_1, \dots, y_n) = (x_1 + y_1, \dots, x_n + y_n).$$

Often the mathematics of $\mathbf{F}^n$ becomes cleaner if we use a single letter to denote a list of n numbers, without explicitly writing the coordinates. For example, the next result is stated with x and y in $\mathbf{F}^n$ even though the proof requires the more cumbersome notation of $(x_1, \dots, x_n)$ and $(y_1, \dots, y_n)$.

$\mathbf{F}^n$

כדי להגדיר את האנלוגים של $\mathbf{R}^2$ ו־ $\mathbf{R}^3$ בממדים גבוהים יותר, פשוט נחליף את $\mathbf{R}$ ב־ $\mathbf{F}$ (השווה ל־ $\mathbf{R}$ או ל־ $\mathbf{C}$) ונחליף את 2 או 3 במספר שלם חיובי שרירותי.

1.10 סימון: n

נְקַבֵּעַ מספר שלם חיובי n עד לסוף הפרק.

1.11 הגדרה: $\mathbf{F}^n$, קואורדינטה

$\mathbf{F}^n$ היא קבוצת כל הרשימות באורך n של איברי $\mathbf{F}$:

$$\mathbf{F}^n = \{(x_1, \dots, x_n) : x_k \in \mathbf{F},\ k = 1, \dots, n\}.$$

עבור $(x_1, \dots, x_n) \in \mathbf{F}^n$ ו־ $k \in \{1, \dots, n\}$, נאמר ש־ x_k היא הקואורדינטה ה־ k^{th} של $(x_1, \dots, x_n)$.

אם $\mathbf{F} = \mathbf{R}$ ו־ n שווה 2 או 3, אזי ההגדרה לעיל של $\mathbf{F}^n$ עולה בקנה אחד עם הרעיונות הקודמים של $\mathbf{R}^2$ ו־ $\mathbf{R}^3$.

1.12 דוגמה: $\mathbf{C}^4$

$\mathbf{C}^4$ היא קבוצת כל הרשימות של ארבעה מספרים מרוכבים:

$$\mathbf{C}^4 = \{(z_1, z_2, z_3, z_4) : z_1, z_2, z_3, z_4 \in \mathbf{C}\}.$$

אם $n \geq 4$, איננו יכולים לדמיין את $\mathbf{R}^n$ כאובייקט פיזי. באופן דומה, ניתן לחשוב על $\mathbf{C}^1$ כמישור, אך עבור $n \geq 2$, תפיסה מוחשית ומלאה של $\mathbf{C}^n$ היא מעבר ליכולתו של המוח האנושי. חרף מגבלה זו, ביכולתנו לבצע מניפולציות אלגבריות ב־ $\mathbf{F}^n$, אפילו עבור ערכי n גבוהים, באותה קלות שבה אנו פועלים ב־ $\mathbf{R}^2$ או ב־ $\mathbf{R}^3$. לדוגמה, חיבור ב־ $\mathbf{F}^n$ מוגדר באופן הבא.

קראו את ספרו של **אדווין אבוט אבוט** - **שטוחלנדיה** (או **שטוחארץ** בתרגומו הקדוש יותר) - לתיאור משעשע של האופן שבו $\mathbf{R}^3$ ייתפס בתודעת יצורים החיים ב־ $\mathbf{R}^2$. הספר, שיצא לאור בשנת 1884, יכול לסייע לכם לדמיין מרחב פיזי בעל ארבעה ממדים או יותר.

1.13 הגדרה: פעולת חיבור ב־ $\mathbf{F}^n$

פעולת חיבור ב־ $\mathbf{F}^n$ מוגדרת על ידי חיבור קואורדינטות בהתאמה:

$$(x_1, \dots, x_n) + (y_1, \dots, y_n) = (x_1 + y_1, \dots, x_n + y_n).$$

ככלל, במתמטיקה של $\mathbf{F}^n$ נקבל תחביר מתמטי נקי יותר אם נשתמש באות יחידה לסימון רשימה של n מספרים, מבלי לכתוב את הקואורדינטות באופן מפורש. לדוגמה, המסקנה הבאה מנוסחת עם x ו־ y ב־ $\mathbf{F}^n$, אף שההוכחה דורשת את הסימון המסורבל יותר של $(x_1, \dots, x_n)$ ו־ $(y_1, \dots, y_n)$.

1.14 *commutativity of addition in* $\mathbf{F}^n$

If $x, y \in \mathbf{F}^n$, then $x + y = y + x$.

Proof Suppose $x = (x_1, \dots, x_n) \in \mathbf{F}^n$ and $y = (y_1, \dots, y_n) \in \mathbf{F}^n$. Then

$$\begin{aligned} x + y &= (x_1, \dots, x_n) + (y_1, \dots, y_n) \\ &= (x_1 + y_1, \dots, x_n + y_n) \\ &= (y_1 + x_1, \dots, y_n + x_n) \\ &= (y_1, \dots, y_n) + (x_1, \dots, x_n) \\ &= y + x, \end{aligned}$$

where the second and fourth equalities above hold because of the definition of addition in $\mathbf{F}^n$ and the third equality holds because of the usual commutativity of addition in $\mathbf{F}$. ∎

The symbol ∎ means "end of proof".

If a single letter is used to denote an element of $\mathbf{F}^n$, then the same letter with appropriate subscripts is often used when coordinates must be displayed. For example, if $x \in \mathbf{F}^n$, then letting x equal $(x_1, \dots, x_n)$ is good notation, as shown in the proof above. Even better, work with just x and avoid explicit coordinates when possible.

1.15 notation: 0

Let 0 denote the list of length n whose coordinates are all 0:

$$0 = (0, \dots, 0).$$

Here we are using the symbol 0 in two different ways—on the left side of the equation above, the symbol 0 denotes a list of length n, which is an element of $\mathbf{F}^n$, whereas on the right side, each 0 denotes a number. This potentially confusing practice actually causes no problems because the context should always make clear which 0 is intended.

1.16 example: *context determines which* 0 *is intended*

Consider the statement that 0 is an additive identity for $\mathbf{F}^n$:

$$x + 0 = x \quad \text{for all } x \in \mathbf{F}^n.$$

Here the 0 above is the list defined in 1.15, not the number 0, because we have not defined the sum of an element of $\mathbf{F}^n$ (namely, x) and the number 0.

1.14 קומוטטיביות של חיבור ב־ $\mathbf{F}^n$

אם $x, y \in \mathbf{F}^n$, אזי $x + y = y + x$.

הוכחה נניח ש־ $x = (x_1, \dots, x_n) \in \mathbf{F}^n$ ו־ $y = (y_1, \dots, y_n) \in \mathbf{F}^n$. אזי

$$\begin{aligned} x + y &= (x_1, \dots, x_n) + (y_1, \dots, y_n) \\ &= (x_1 + y_1, \dots, x_n + y_n) \\ &= (y_1 + x_1, \dots, y_n + x_n) \\ &= (y_1, \dots, y_n) + (x_1, \dots, x_n) \\ &= y + x, \end{aligned}$$

כאשר השוויון השני והשוויון הרביעי מתקיימים לפי הגדרת החיבור ב־ $\mathbf{F}^n$ והשוויון השלישי מתקיים לפי הקומוטטיביות הרגילה של החיבור ב־ $\mathbf{F}$. ■

הסמל ■ משמעותו סוף ההוכחה (פש"ל).

אם אנו משתמשים באות יחידה כדי לסמן איבר של $\mathbf{F}^n$, אזי אותה האות עם אינדקסים מתאימים תשמש לרוב כאשר אנו נאלצים להציג את הקואורדינטות. לדוגמה, אם $x \in \mathbf{F}^n$, אזי הגדרת $x = (x_1, \dots, x_n)$ היא סימון טוב, כפי שהוצג בהוכחה לעיל. אך טוב יותר לעבוד עם x בלבד, ולהימנע משימוש בקואורדינטות מפורשות ככל הניתן.

1.15 סימון: 0

נסמן ב־ 0 את הרשימה שאורכה n ושהקואורדינטות שלה כולן 0:

$$0 = (0, \dots, 0).$$

כאן אנו משתמשים בסמל 0 בשני אופנים שונים—באגף השמאלי של המשוואה לעיל, הסמל 0 מסמן רשימה באורך n, שהיא איבר ב־ $\mathbf{F}^n$ ואילו באגף הימני, כל 0 מסמן מספר. הרגל זה, שלכאורה עלול לגרום לבלבול, אינו בעייתי כלל שכן מתוך ההקשר תמיד ברור לאיזה 0 מתכוונים.

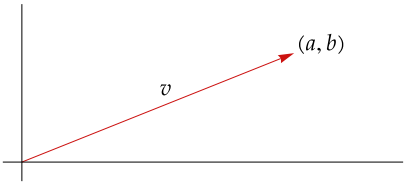
1.16 דוגמה: הֶקְשֵׁר קובע לאיזה 0 מתכוונים

נתבונן במשפט המגדיר את 0 כאיבר זהות לחיבור ב $\mathbf{F}^n$:

$$x + 0 = x \quad \text{לכל } x \in \mathbf{F}^n.$$

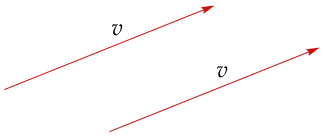
בדוגמה זו, 0 הוא הרשימה שהוגדרה ב־ 1.15, ולא המספר 0, שהרי לא הגדרנו את הסכום של איבר ב־ $\mathbf{F}^n$ (במקרה שלנו x) והמספר 0.

A picture can aid our intuition. We will draw pictures in $\mathbf{R}^2$ because we can sketch this space on two-dimensional surfaces such as paper and computer screens. A typical element of $\mathbf{R}^2$ is a point $v = (a, b)$. Sometimes we think of v not as a point but as an arrow starting at the origin and ending at (a, b), as shown here. When we think of an element of $\mathbf{R}^2$ as an arrow, we refer to it as a *vector*.



Elements of $\mathbf{R}^2$ can be thought of as points or as vectors.

When we think of vectors in $\mathbf{R}^2$ as arrows, we can move an arrow parallel to itself (not changing its length or direction) and still think of it as the same vector. With that viewpoint, you will often gain better understanding by dispensing with the coordinate axes and the explicit coordinates and just thinking of the vector, as shown in the figure here. The two arrows shown here have the same length and same direction, so we think of them as the same vector.



A vector.

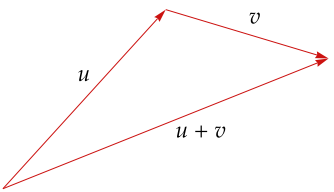
Whenever we use pictures in $\mathbf{R}^2$ or use the somewhat vague language of points and vectors, remember that these are just aids to our understanding, not substitutes for the actual mathematics that we will develop. Although we cannot draw good pictures in high-dimensional spaces, the elements of these spaces are as rigorously defined as elements of $\mathbf{R}^2$.

*Mathematical models of the economy can have thousands of variables, say $x_1, \ldots, x_{5000}$, which means that we must work in $\mathbf{R}^{5000}$. Such a space cannot be dealt with geometrically. However, the algebraic approach works well. Thus our subject is called **linear algebra**.*

For example, $\left(2, -3, 17, \pi, \sqrt{2}\right)$ is an element of $\mathbf{R}^5$, and we may casually refer to it as a point in $\mathbf{R}^5$ or a vector in $\mathbf{R}^5$ without worrying about whether the geometry of $\mathbf{R}^5$ has any physical meaning.

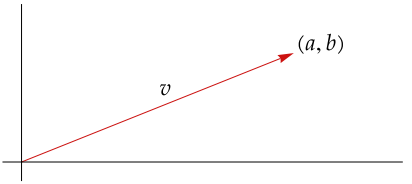
Recall that we defined the sum of two elements of $\mathbf{F}^n$ to be the element of $\mathbf{F}^n$ obtained by adding corresponding coordinates; see 1.13. As we will now see, addition has a simple geometric interpretation in the special case of $\mathbf{R}^2$.

Suppose we have two vectors u and v in $\mathbf{R}^2$ that we want to add. Move the vector v parallel to itself so that its initial point coincides with the end point of the vector u, as shown here. The sum $u + v$ then equals the vector whose initial point equals the initial point of u and whose end point equals the end point of the vector v, as shown here.



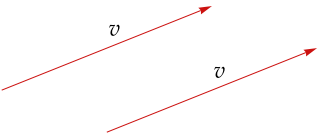
The sum of two vectors.

In the next definition, the 0 on the right side of the displayed equation is the list $0 \in \mathbf{F}^n$.



ניתן לחשוב על איברים של $\mathbf{R}^2$ כנקודות או כווקטורים.

איור יכול לסייע לאינטואיציה. נשרטט תמונות ב־ $\mathbf{R}^2$ מאחר שאנו יכולים לשרטט את המרחב הזה על גבי משטחים דו־ממדיים כגון נייר ומסכי מחשב. איבר טיפוסי של $\mathbf{R}^2$ הוא נקודה $v = (a, b)$. לעיתים אנו חושבים על v לא כעל נקודה כי אם כעל חץ שמוצאו בראשית הצירים וראשו ב־ (a, b), כפי שמוצג כאן. כאשר אנו חושבים על איבר של $\mathbf{R}^2$ כעל חץ, אנו קוראים לו וקטור.



וקטור.

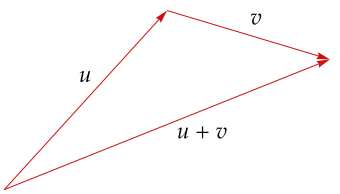
כשאנו חושבים על וקטורים כעל חיצים ב־ $\mathbf{R}^2$, אנו יכולים להזיז חץ במקביל לעצמו (מבלי לשנות כיוון או אורך) ועדיין להחשיב אותו כאותו הווקטור. מנקודת מבט זו, לעיתים קרובות תוכלו להגיע להבנה טובה יותר בכך שתתעלמו מצירי הקואורדינטות ומהקואורדינטות המפורשות, ותחשבו פשוט על הווקטור כפי שמופיע בשרטוט כאן. לשני החיצים המוצגים כאן אותו האורך ואותו הכיוון, ולכן אנו מחשיבים אותם כאותו הווקטור.

בכל פעם שאנו משתמשים בשרטוטים ב־ $\mathbf{R}^2$ או משתמשים בשפה מעורפלת במקצת של נקודות ווקטורים, זכרו שאלו הם רק עזרים להבנתנו, ואינם תחליף למתמטיקה שנְפַתֵּחַ בפועל. אף שאיננו יכולים לשרטט תמונה גרפית טובה במרחבים מממד גבוה, האיברים במרחבים אלו מוגדרים באותה מידת קפדנות שבה אנו מגדירים איברים של $\mathbf{R}^2$.

מודלים מתמטיים בכלכלה יכולים לכלול אלפי משתנים, לדוגמה, $x_1, \ldots, x_{5000}$, ומשמעות הדבר היא שעלינו לעבוד ב־ $\mathbf{R}^{5000}$. לא ניתן לעסוק במרחב מממד כזה באופן גאומטרי ובכל זאת הגישה האלגברית עובדת ללא דופי. לפיכך הנושא שלנו נקרא **אלגברה ליניארית**.

לדוגמה, $\big(2, -3, 17, \pi, \sqrt{2}\big)$ הוא איבר של $\mathbf{R}^5$, ואנו עשויים להתייחס אליו כלאחר יד כנקודה ב־ $\mathbf{R}^5$ או כווקטור ב־ $\mathbf{R}^5$ מבלי להיות מוטרדים מהשאלה האם לגאומטריה של $\mathbf{R}^5$ יש משמעות פיזית כלשהי.

נזכיר כי הגדרנו את הסכום של שני איברים של $\mathbf{F}^n$ כאיבר ב־ $\mathbf{F}^n$ המושג באמצעות חיבור הקואורדינטות בהתאמה; ראו 1.13. כפי שנראה כעת, לפעולת החיבור יש פרשנות גאומטרית פשוטה במקרה הפרטי של $\mathbf{R}^2$.



הסכום של שני וקטורים.

נניח שיש לנו שני וקטורים u ו־ v ב־ $\mathbf{R}^2$ שאותם נרצה לחבר. נמקם את הווקטור v במקביל לעצמו, כך שנקודת מוצאו חופפת לנקודת סיומו של הווקטור u, כפי שמוצג בשרטוט. הסכום $u + v$ שווה אפוא לווקטור שנקודת מוצאו שווה לנקודת המוצא של u ונקודת סיומו שווה לנקודת הסיום של v, כפי שמוצג כאן.

בהגדרה הבאה, ה־ 0 באגף הימני של המשוואה הוא הרשימה $0 \in \mathbf{F}^n$.

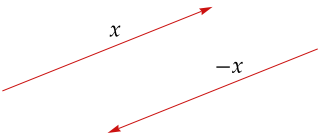
1.17 definition: *additive inverse in* $\mathbf{F}^n$, $-x$

For $x \in \mathbf{F}^n$, the *additive inverse* of x, denoted by $-x$, is the vector $-x \in \mathbf{F}^n$ such that

$$x + (-x) = 0.$$

Thus if $x = (x_1, \dots, x_n)$, then $-x = (-x_1, \dots, -x_n)$.

The additive inverse of a vector in $\mathbf{R}^2$ is the vector with the same length but pointing in the opposite direction. The figure here illustrates this way of thinking about the additive inverse in $\mathbf{R}^2$. As you can see, the vector labeled $-x$ has the same length as the vector labeled x but points in the opposite direction.



A vector and its additive inverse.

Having dealt with addition in $\mathbf{F}^n$, we now turn to multiplication. We could define a multiplication in $\mathbf{F}^n$ in a similar fashion, starting with two elements of $\mathbf{F}^n$ and getting another element of $\mathbf{F}^n$ by multiplying corresponding coordinates. Experience shows that this definition is not useful for our purposes. Another type of multiplication, called scalar multiplication, will be central to our subject. Specifically, we need to define what it means to multiply an element of $\mathbf{F}^n$ by an element of $\mathbf{F}$.

1.18 definition: *scalar multiplication in* $\mathbf{F}^n$

The *product* of a number λ and a vector in $\mathbf{F}^n$ is computed by multiplying each coordinate of the vector by λ:

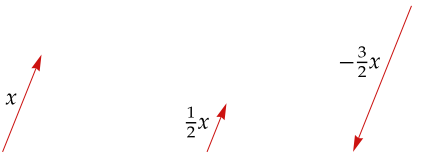
$$\lambda(x_1, \dots, x_n) = (\lambda x_1, \dots, \lambda x_n);$$

here $\lambda \in \mathbf{F}$ and $(x_1, \dots, x_n) \in \mathbf{F}^n$.

Scalar multiplication has a nice geometric interpretation in $\mathbf{R}^2$. If $\lambda > 0$ and $x \in \mathbf{R}^2$, then λx is the vector that points in the same direction as x and whose length is λ times the length of x. In other words, to get λx, we shrink or stretch x by a factor of λ, depending on whether $\lambda < 1$ or $\lambda > 1$.

If $\lambda < 0$ and $x \in \mathbf{R}^2$, then λx is the vector that points in the direction opposite to that of x and whose length is $|\lambda|$ times the length of x, as shown here.

Scalar multiplication in $\mathbf{F}^n$ multiplies together a scalar and a vector, getting a vector. In contrast, the dot product in $\mathbf{R}^2$ or $\mathbf{R}^3$ multiplies together two vectors and gets a scalar. Generalizations of the dot product will become important in Chapter 6.



Scalar multiplication.

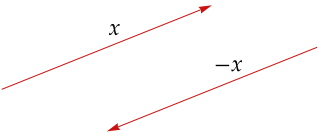
1.17 הגדרה: איבר נגדי לחיבור ב־ $\mathbf{F}^n$, סימון $-x$

עבור $x \in \mathbf{F}^n$, האיבר הנגדי לחיבור של x, המסומן ב־ $-x$, הוא הווקטור $-x \in \mathbf{F}^n$ כך ש־

$$x + (-x) = 0.$$

לפיכך, אם $x = (x_1, \dots, x_n)$, אזי $-x = (-x_1, \dots, -x_n)$.

האיבר הנגדי לחיבור של וקטור ב־ $\mathbf{R}^2$ הוא הווקטור שאורכו זהה אך מצביע בכיוון ההפוך. השרטוט המובא כאן ממחיש את אופן חשיבה זה בנוגע לאיבר נגדי לחיבור ב־ $\mathbf{R}^2$. כפי שאתם רואים, הווקטור המסומן ב־ $-x$ אורכו זהה לאורכו של הווקטור המסומן ב־ x אך הוא מצביע בכיוון ההפוך.



וקטור ואיברו הנגדי לחיבור.

אחרי שעסקנו בפעולת החיבור ב־ $\mathbf{F}^n$, נעבור כעת לעסוק בכפל. ביכולתנו להגדיר את פעולת הכפל ב־ $\mathbf{F}^n$ באופן דומה לזה שבו הגדרנו את פעולת החיבור, כלומר, באמצעות כפל קואורדינטות בהתאמה של שני איברים ב־ $\mathbf{F}^n$ וקבלת איבר נוסף של $\mathbf{F}^n$. הניסיון מראה שהגדרה זו אינה שימושית לצרכינו. סוג אחר של כפל, ששמו כפל בסקלר, יהיה מרכיב מרכזי בנושאים שבהם נעסוק. בפרט, עלינו להגדיר מהי משמעות פעולת הכפלת איבר ב־ $\mathbf{F}^n$ באיבר ב־ $\mathbf{F}$.

1.18 הגדרה: כפל בסקלר ב־ $\mathbf{F}^n$

המכפלה של מספר λ ווקטור ב־ $\mathbf{F}^n$ מחושבת באמצעות הכפלת כל קואורדינטה של הווקטור ב־ λ:

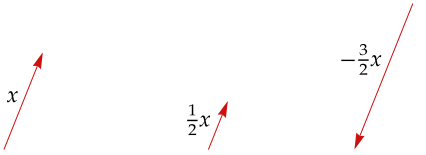
$$\lambda(x_1, \dots, x_n) = (\lambda x_1, \dots, \lambda x_n);$$

כאן $\lambda \in \mathbf{F}$ ו־ $(x_1, \dots, x_n) \in \mathbf{F}^n$.

להכפלת וקטור בסקלר יש פרשנות גאומטרית נחמדה ב־ $\mathbf{R}^2$. אם $\lambda > 0$ ו־ $x \in \mathbf{R}^2$, אזי λx הוא הווקטור המצביע לאותו הכיוון שאליו מצביע x ואורכו הוא λ פעמים אורכו של x. כלומר, כדי ליצור את λx, אנו מכווצים או מותחים את x פי λ, בהתאמה למקרים $\lambda < 1$ או $\lambda > 1$.

אם $\lambda < 0$ ו־ $x \in \mathbf{R}^2$, אזי λx הוא הווקטור הפונה לכיוון ההפוך מזה שאליו פונה x ואורכו $|\lambda|$ פעמים אורכו של x, כפי שאנו רואים בשרטוט.

בכפל בסקלר ב־ $\mathbf{F}^n$ אנו מכפילים סקלר בווקטור ומקבלים וקטור. בניגוד לכך, המכפלה הסקלרית ב־ $\mathbf{R}^2$ או $\mathbf{R}^3$ מכפילה יחד שני וקטורים ומפיקה סקלר. ההכללות של המכפלה הסקלרית יהיו חשובות בפרק ו.



כפל בסקלר.

Digression on Fields

A *field* is a set containing at least two distinct elements called 0 and 1, along with operations of addition and multiplication satisfying all properties listed in 1.3. Thus $\mathbf{R}$ and $\mathbf{C}$ are fields, as is the set of rational numbers along with the usual operations of addition and multiplication. Another example of a field is the set $\{0, 1\}$ with the usual operations of addition and multiplication except that $1 + 1$ is defined to equal 0.

In this book we will not deal with fields other than $\mathbf{R}$ and $\mathbf{C}$. However, many of the definitions, theorems, and proofs in linear algebra that work for the fields $\mathbf{R}$ and $\mathbf{C}$ also work without change for arbitrary fields. If you prefer to do so, throughout much of this book (except for Chapters 6 and 7, which deal with inner product spaces) you can think of $\mathbf{F}$ as denoting an arbitrary field instead of $\mathbf{R}$ or $\mathbf{C}$. For results (except in the inner product chapters) that have as a hypothesis that $\mathbf{F}$ is $\mathbf{C}$, you can probably replace that hypothesis with the hypothesis that $\mathbf{F}$ is an algebraically closed field, which means that every nonconstant polynomial with coefficients in $\mathbf{F}$ has a zero. A few results, such as Exercise 13 in Section 1C, require the hypothesis on $\mathbf{F}$ that $1 + 1 \neq 0$.

Exercises 1A

1 Show that $\alpha + \beta = \beta + \alpha$ for all $\alpha, \beta \in \mathbf{C}$.

2 Show that $(\alpha + \beta) + \lambda = \alpha + (\beta + \lambda)$ for all $\alpha, \beta, \lambda \in \mathbf{C}$.

3 Show that $(\alpha\beta)\lambda = \alpha(\beta\lambda)$ for all $\alpha, \beta, \lambda \in \mathbf{C}$.

4 Show that $\lambda(\alpha + \beta) = \lambda\alpha + \lambda\beta$ for all $\lambda, \alpha, \beta \in \mathbf{C}$.

5 Show that for every $\alpha \in \mathbf{C}$, there exists a unique $\beta \in \mathbf{C}$ such that $\alpha + \beta = 0$.

6 Show that for every $\alpha \in \mathbf{C}$ with $\alpha \neq 0$, there exists a unique $\beta \in \mathbf{C}$ such that $\alpha\beta = 1$.

7 Show that

$$\frac{-1 + \sqrt{3}\,i}{2}$$

is a cube root of 1 (meaning that its cube equals 1).

8 Find two distinct square roots of i.

9 Find $x \in \mathbf{R}^4$ such that

$$(4, -3, 1, 7) + 2x = (5, 9, -6, 8).$$

10 Explain why there does not exist $\lambda \in \mathbf{C}$ such that

$$\lambda(2 - 3i, 5 + 4i, -6 + 7i) = (12 - 5i, 7 + 22i, -32 - 9i).$$

אתנחתא בנושא שדות

שדה הוא קבוצה המכילה לכל הפחות שני איברים שונים הנקראים 0 ו־ 1, יחד עם פעולות חיבור וכפל המקיימות את כל התכונות שפורטו ב־ 1.3. לפיכך, $\mathbf{R}$ ו־ $\mathbf{C}$ הם שדות, כמו גם קבוצת המספרים הרציונליים יחד עם פעולות החיבור והכפל הרגילות. דוגמה נוספת לשדה היא הקבוצה $\{0, 1\}$ עם פעולות החיבור והכפל הרגילות פרט לכך ש־ $1 + 1$ מוגדר כשווה ל־ 0.

בספר זה לא נעסוק בשדות שאינם $\mathbf{R}$ ו־ $\mathbf{C}$. אף על פי כן, הגדרות, משפטים והוכחות רבים באלגברה ליניארית המתקיימים בשדות $\mathbf{R}$ ו־ $\mathbf{C}$ יתקיימו גם בכל שדה שרירותי ללא שינוי או התאמה. אם תרצו, לאורכו של ספר זה (למעט פרקים ו ו־ 8, הדנים במרחבי מכפלה פנימית) תוכלו לחשוב על $\mathbf{F}$ כעל סימון לשדה שרירותי במקום $\mathbf{R}$ או $\mathbf{C}$. עבור מסקנות (מלבד פרקי המכפלה הפנימית) המבוססות על ההנחה ש־ $\mathbf{F}$ הוא $\mathbf{C}$, תוכלו ככל הנראה להחליף הנחה זו בהנחה ש־ $\mathbf{F}$ הוא שדה סגור מבחינה אלגברית, שמשמעותו כי לכל פולינום שאינו קבוע עם מקדמים ב־ $\mathbf{F}$ יש שורש. במספר מסקנות כמו זו שבתרגיל 13 ביחידה א3 נדרשת ההנחה שב־ $\mathbf{F}$ מתקיים $0 \neq 1 + 1$.

תרגילים או

1 הראו ש־ $\alpha + \beta = \beta + \alpha$ לכל $\alpha, \beta \in \mathbf{C}$.

2 הראו ש־ $(\alpha + \beta) + \lambda = \alpha + (\beta + \lambda)$ לכל $\alpha, \beta, \lambda \in \mathbf{C}$.

3 הראו ש־ $(\alpha\beta)\lambda = \alpha(\beta\lambda)$ לכל $\alpha, \beta, \lambda \in \mathbf{C}$.

4 הראו ש־ $\lambda(\alpha + \beta) = \lambda\alpha + \lambda\beta$ לכל $\lambda, \alpha, \beta \in \mathbf{C}$.

5 הראו שלכל $\alpha \in \mathbf{C}$, קיים איבר יחיד $\beta \in \mathbf{C}$ כך ש־ $\alpha + \beta = 0$.

6 הראו שלכל $\alpha \in \mathbf{C}$ עם $\alpha \neq 0$, קיים איבר יחיד $\beta \in \mathbf{C}$ כך ש־ $\alpha\beta = 1$.

7 הראו ש־

$$\frac{-1 + \sqrt{3}i}{2}$$

הוא שורש שלישי של 1 (כלומר, העלאתו בחזקה שלישית נותנת 1).

8 מצאו שני שורשים ריבועיים שונים של i.

9 מצאו $x \in \mathbf{R}^4$ כך ש־

$$(4, -3, 1, 7) + 2x = (5, 9, -6, 8).$$

10 הסבירו מדוע לא קיים $\lambda \in \mathbf{C}$ כך ש־

$$\lambda(2 - 3i, 5 + 4i, -6 + 7i) = (12 - 5i, 7 + 22i, -32 - 9i).$$

11 Show that $(x + y) + z = x + (y + z)$ for all $x, y, z \in \mathbf{F}^n$.

12 Show that $(ab)x = a(bx)$ for all $x \in \mathbf{F}^n$ and all $a, b \in \mathbf{F}$.

13 Show that $1x = x$ for all $x \in \mathbf{F}^n$.

14 Show that $\lambda(x + y) = \lambda x + \lambda y$ for all $\lambda \in \mathbf{F}$ and all $x, y \in \mathbf{F}^n$.

15 Show that $(a + b)x = ax + bx$ for all $a, b \in \mathbf{F}$ and all $x \in \mathbf{F}^n$.

"Can you do addition?" the White Queen asked. "What's one and one and one and one and one and one and one and one and one and one?"
"I don't know," said Alice. "I lost count."

—*Through the Looking Glass*, Lewis Carroll

11 הראו ש־ $(x + y) + z = x + (y + z)$ לכל $x, y, z \in \mathbf{F}^n$.

12 הראו ש־ $(ab)x = a(bx)$ לכל $x \in \mathbf{F}^n$ ולכל $a, b \in \mathbf{F}$.

13 הראו ש־ $1x = x$ לכל $x \in \mathbf{F}^n$.

14 הראו ש־ $\lambda(x + y) = \lambda x + \lambda y$ לכל $\lambda \in \mathbf{F}$ ולכל $x, y \in \mathbf{F}^n$.

15 הראו ש־ $(a + b)x = ax + bx$ לכל $a, b \in \mathbf{F}$ ולכל $x \in \mathbf{F}^n$.

"חיבור את יודעת?" שאלה המלכה הלבנה. "כמה זה אחד ועוד אחד ועוד אחד ועוד אחד ועוד אחד ועוד אחד ועוד אחד ועוד אחד ועוד אחד ועוד אחד ועוד אחד ועוד אחד?"
"אינני יודעת," אמרה אליס. "הפסקתי לספור."

—מבעד למראה ומה אליס מצאה שם, לואיס קרול

1B Definition of Vector Space

The motivation for the definition of a vector space comes from properties of addition and scalar multiplication in $\mathbf{F}^n$: Addition is commutative, associative, and has an identity. Every element has an additive inverse. Scalar multiplication is associative. Scalar multiplication by 1 acts as expected. Addition and scalar multiplication are connected by distributive properties.

We will define a vector space to be a set V with an addition and a scalar multiplication on V that satisfy the properties in the paragraph above.

1.19 definition: *addition, scalar multiplication*

- An *addition* on a set V is a function that assigns an element $u + v \in V$ to each pair of elements $u, v \in V$.
- A *scalar multiplication* on a set V is a function that assigns an element $\lambda v \in V$ to each $\lambda \in \mathbf{F}$ and each $v \in V$.

Now we are ready to give the formal definition of a vector space.

1.20 definition: *vector space*

A *vector space* is a set V along with an addition on V and a scalar multiplication on V such that the following properties hold.

commutativity
$u + v = v + u$ for all $u, v \in V$.

associativity
$(u + v) + w = u + (v + w)$ and $(ab)v = a(bv)$ for all $u, v, w \in V$ and for all $a, b \in \mathbf{F}$.

additive identity
There exists an element $0 \in V$ such that $v + 0 = v$ for all $v \in V$.

additive inverse
For every $v \in V$, there exists $w \in V$ such that $v + w = 0$.

multiplicative identity
$1v = v$ for all $v \in V$.

distributive properties
$a(u + v) = au + av$ and $(a + b)v = av + bv$ for all $a, b \in \mathbf{F}$ and all $u, v \in V$.

The following geometric language sometimes aids our intuition.

1.21 definition: *vector, point*

Elements of a vector space are called *vectors* or *points*.

אג הגדרת מרחב וקטורי

המוטיבציה להגדרה של מרחב וקטורי נובעת מתכונות פעולות החיבור והכפל בסקלר ב $\mathbf{F}^n$: החיבור הוא קומוטטיבי, אסוציאטיבי, ויש לו איבר זהות. לכל איבר יש איבר נגדי לחיבור. כפל בסקלר הוא אסוציאטיבי. כפל בסקלר 1 פועל כמצופה. החיבור והכפל בסקלר קשורים זה לזה באמצעות תכונות הדיסטריבוטיביות.

נגדיר אפוא מרחב וקטורי כקבוצה V שעליה מוגדרות פעולות חיבור וכפל בסקלר, המקיימות את התכונות שפורטו לעיל.

1.19 הגדרה: חיבור, כפל בסקלר

- חיבור מעל הקבוצה V הוא פונקציה המתאימה איבר $u+v \in V$ לכל זוג איברים $u, v \in V$.
- כפל בסקלר מעל הקבוצה V הוא פונקציה המתאימה איבר $\lambda v \in V$ לכל $\lambda \in \mathbf{F}$ ולכל $v \in V$.

כעת באפשרותנו להגדיר מרחב וקטורי באופן פורמלי.

1.20 הגדרה: מרחב וקטורי

מרחב וקטורי הוא קבוצה V שעבורה הגדרנו פעולת חיבור ופעולת כפל בסקלר כך שהתכונות להלן מתקיימות.

קומוטטיביות

$u + v = v + u$ לכל $u, v \in V$.

אסוציאטיביות

$(u + v) + w = u + (v + w)$ ו־ $(ab)v = a(bv)$ לכל $u, v, w \in V$ ולכל $a, b \in \mathbf{F}$.

איבר זהות לחיבור

קיים איבר $0 \in V$ כך ש־ $v + 0 = v$ לכל $v \in V$.

איבר נגדי לחיבור

לכל $v \in V$, קיים איבר $w \in V$ כך ש־ $v + w = 0$.

איבר זהות לכפל

$1v = v$ לכל $v \in V$.

תכונות הדיסטריבוטיביות

$a(u + v) = au + av$ ו־ $(a + b)v = av + bv$ לכל $a, b \in \mathbf{F}$ ולכל $u, v \in V$.

הניסוח הגאומטרי הבא מסייע לעיתים לאינטואיציה שלנו.

1.21 הגדרה: וקטור, נקודה

האיברים במרחב הווקטורי נקראים וקטורים או נקודות.

The scalar multiplication in a vector space depends on $\mathbf{F}$. Thus when we need to be precise, we will say that V is a *vector space over* $\mathbf{F}$ instead of saying simply that V is a vector space. For example, $\mathbf{R}^n$ is a vector space over $\mathbf{R}$, and $\mathbf{C}^n$ is a vector space over $\mathbf{C}$.

1.22 definition: *real vector space, complex vector space*

- A vector space over $\mathbf{R}$ is called a *real vector space*.
- A vector space over $\mathbf{C}$ is called a *complex vector space*.

Usually the choice of $\mathbf{F}$ is either clear from the context or irrelevant. Thus we often assume that $\mathbf{F}$ is lurking in the background without specifically mentioning it.

With the usual operations of addition and scalar multiplication, $\mathbf{F}^n$ is a vector space over $\mathbf{F}$, as you should verify. The example of $\mathbf{F}^n$ motivated our definition of vector space.

The simplest vector space is $\{0\}$, which contains only one point.

1.23 example: $\mathbf{F}^\infty$

$\mathbf{F}^\infty$ is defined to be the set of all sequences of elements of $\mathbf{F}$:

$$\mathbf{F}^\infty = \{(x_1, x_2, \dots) : x_k \in \mathbf{F} \text{ for } k = 1, 2, \dots\}.$$

Addition and scalar multiplication on $\mathbf{F}^\infty$ are defined as expected:

$$(x_1, x_2, \dots) + (y_1, y_2, \dots) = (x_1 + y_1, x_2 + y_2, \dots),$$
$$\lambda(x_1, x_2, \dots) = (\lambda x_1, \lambda x_2, \dots).$$

With these definitions, $\mathbf{F}^\infty$ becomes a vector space over $\mathbf{F}$, as you should verify. The additive identity in this vector space is the sequence of all 0's.

Our next example of a vector space involves a set of functions.

1.24 notation: $\mathbf{F}^S$

- If S is a set, then $\mathbf{F}^S$ denotes the set of functions from S to $\mathbf{F}$.
- For $f, g \in \mathbf{F}^S$, the *sum* $f + g \in \mathbf{F}^S$ is the function defined by

$$(f + g)(x) = f(x) + g(x)$$

for all $x \in S$.

- For $\lambda \in \mathbf{F}$ and $f \in \mathbf{F}^S$, the *product* $\lambda f \in \mathbf{F}^S$ is the function defined by

$$(\lambda f)(x) = \lambda f(x)$$

for all $x \in S$.

הכפל בסקלר במרחב וקטורי תלוי ב־ $\mathbf{F}$. לפיכך כשנזדקק להגדרה מדויקת, נאמר כי V מרחב וקטורי מעל $\mathbf{F}$, במקום לומר בפשטות ש־ V מרחב וקטורי. לדוגמה, $\mathbf{R}^n$ הוא מרחב וקטורי מעל $\mathbf{R}$, ו־ $\mathbf{C}^n$ הוא מרחב וקטורי מעל $\mathbf{C}$.

1.22 הגדרה: מרחב וקטורי ממשי, מרחב וקטורי מרוכב

- מרחב וקטורי מעל $\mathbf{R}$ נקרא מרחב וקטורי ממשי.
- מרחב וקטורי מעל $\mathbf{C}$ נקרא מרחב וקטורי מרוכב.

בדרך כלל ה־ $\mathbf{F}$ שבחרנו ברור מההקשר או שאינו רלוונטי. לפיכך ברוב המקרים נניח ש־ $\mathbf{F}$ אורב ברקע מבלי שנציין אותו מפורשות.

בצירוף פעולות החיבור והכפל בסקלר הרגילות, $\mathbf{F}^n$ הוא מרחב וקטורי מעל $\mathbf{F}$, כפי שעליכם לוודא. הדוגמה של $\mathbf{F}^n$ היוותה מוטיבציה להגדרה של מרחב וקטורי.

1.23 דוגמה: $\mathbf{F}^\infty$

$\mathbf{F}^\infty$ מוגדרת כקבוצת כל הסדרות של איברים ב־ $\mathbf{F}$:

$$\mathbf{F}^\infty = \{(x_1, x_2, \dots) : x_k \in \mathbf{F}, k = 1, 2, \dots\}.$$

חיבור וכפל בסקלר ב־ $\mathbf{F}^\infty$ מוגדרים כצפוי:

$$(x_1, x_2, \dots) + (y_1, y_2, \dots) = (x_1 + y_1, x_2 + y_2, \dots),$$
$$\lambda(x_1, x_2, \dots) = (\lambda x_1, \lambda x_2, \dots).$$

עם הגדרות אלו, $\mathbf{F}^\infty$ הופכת למרחב וקטורי מעל $\mathbf{F}$, כפי שעליכם לוודא. במרחב זה, איבר הזהות לחיבור הוא הסדרה שכל איבריה 0.

דוגמתנו הבאה למרחב וקטורי עוסקת בקבוצת פונקציות.

1.24 סימון: $\mathbf{F}^S$

- אם S קבוצה, אזי $\mathbf{F}^S$ מסמן את קבוצת הפונקציות מ־ S ל־ $\mathbf{F}$.
- עבור $f, g \in \mathbf{F}^S$, הסכום $f + g \in \mathbf{F}^S$ הוא הפונקציה המוגדרת על ידי

$$(f + g)(x) = f(x) + g(x)$$

לכל $x \in S$.

- עבור $\lambda \in \mathbf{F}$ ו־ $f \in \mathbf{F}^S$, המכפלה $\lambda f \in \mathbf{F}^S$ היא הפונקציה המוגדרת על ידי

$$(\lambda f)(x) = \lambda f(x)$$

לכל $x \in S$.

As an example of the notation above, if S is the interval $[0, 1]$ and $\mathbf{F} = \mathbf{R}$, then $\mathbf{R}^{[0,1]}$ is the set of real-valued functions on the interval $[0, 1]$.

You should verify all three bullet points in the next example.

1.25 example: $\mathbf{F}^S$ *is a vector space*

- If S is a nonempty set, then $\mathbf{F}^S$ (with the operations of addition and scalar multiplication as defined above) is a vector space over $\mathbf{F}$.
- The additive identity of $\mathbf{F}^S$ is the function $0 : S \to \mathbf{F}$ defined by

$$0(x) = 0$$

for all $x \in S$.

- For $f \in \mathbf{F}^S$, the additive inverse of f is the function $-f : S \to \mathbf{F}$ defined by

$$(-f)(x) = -f(x)$$

for all $x \in S$.

The elements of the vector space $\mathbf{R}^{[0,1]}$ are real-valued functions on $[0, 1]$, not lists. In general, a vector space is an abstract entity whose elements might be lists, functions, or weird objects.

The vector space $\mathbf{F}^n$ is a special case of the vector space $\mathbf{F}^S$ because each $(x_1, \dots, x_n) \in \mathbf{F}^n$ can be thought of as a function x from the set $\{1, 2, \dots, n\}$ to $\mathbf{F}$ by writing $x(k)$ instead of x_k for the kth coordinate of $(x_1, \dots, x_n)$. In other words, we can think of $\mathbf{F}^n$ as $\mathbf{F}^{\{1,2,\dots,n\}}$. Similarly, we can think of $\mathbf{F}^\infty$ as $\mathbf{F}^{\{1,2,\dots\}}$.

Soon we will see further examples of vector spaces, but first we need to develop some of the elementary properties of vector spaces.

The definition of a vector space requires it to have an additive identity. The next result states that this identity is unique.

1.26 *unique additive identity*

A vector space has a unique additive identity.

Proof Suppose 0 and $0'$ are both additive identities for some vector space V. Then

$$0' = 0' + 0 = 0 + 0' = 0,$$

where the first equality holds because 0 is an additive identity, the second equality comes from commutativity, and the third equality holds because $0'$ is an additive identity. Thus $0' = 0$, proving that V has only one additive identity. ∎

Each element v in a vector space has an additive inverse, an element w in the vector space such that $v + w = 0$. The next result shows that each element in a vector space has only one additive inverse.

כדוגמה לסימון לעיל, אם S הוא הקטע $[0,1]$ ו־ $\mathbf{F} = \mathbf{R}$, אזי $\mathbf{R}^{[0,1]}$ היא הקבוצה של הפונקציות הממשיות מעל הקטע $[0,1]$. עליכם לוודא את כל שלושת הסעיפים בדוגמה הבאה.

1.25 דוגמה: $\mathbf{F}^S$ הוא מרחב וקטורי

- אם S היא קבוצה לא ריקה, אזי $\mathbf{F}^S$ (בצירוף פעולות החיבור והכפל בסקלר שתוארו לעיל) הוא מרחב וקטורי מעל $\mathbf{F}$.
- איבר הזהות לחיבור של $\mathbf{F}^S$ הוא הפונקציה $0 : S \to \mathbf{F}$ המוגדרת על ידי

$$0(x) = 0$$

לכל $x \in S$.

- עבור $f \in \mathbf{F}^S$, האיבר הנגדי לחיבור של f הוא הפונקציה $-f : S \to \mathbf{F}$ המוגדרת על ידי

$$(-f)(x) = -f(x)$$

לכל $x \in S$.

האיברים במרחב הווקטורי $\mathbf{R}^{[0,1]}$ הם הפונקציות הממשיות מעל הקטע $[0,1]$, ולא רשימות. באופן כללי, מרחב וקטורי הוא ישות מופשטת שאיבריה יכולים להיות רשימות, פונקציות או אובייקטים אחרים.

המרחב הווקטורי $\mathbf{F}^n$ הוא מקרה פרטי של המרחב הווקטורי $\mathbf{F}^S$ מכיוון שכל $(x_1, \dots, x_n) \in \mathbf{F}^n$ יכול להיחשב כפונקציה x מהקבוצה $\{1, 2, \dots, n\}$ אל $\mathbf{F}$ בכך שנכתוב $x(k)$ במקום x_k עבור הקואורדינטה ה־ k^{th} של $(x_1, \dots, x_n)$. כלומר, אפשר לחשוב על $\mathbf{F}^n$ כ־ $\mathbf{F}^{\{1,2,\dots,n\}}$. באופן דומה, אפשר לחשוב על $\mathbf{F}^\infty$ כ־ $\mathbf{F}^{\{1,2,\dots\}}$.

בהמשך נכיר דוגמאות נוספות של מרחבים וקטוריים, אך תחילה עלינו לְפַתֵּחַ חלק מהתכונות הבסיסיות של מרחבים וקטוריים.

הגדרת מרחב וקטורי דורשת קיומו של איבר זהות לחיבור. המסקנה הבאה קובעת כי איבר זה הוא יחיד.

1.26 יחידות איבר זהות לחיבור

במרחב וקטורי איבר הזהות לחיבור הוא יחיד.

הוכחה נניח ש־ 0 ו־ $0'$ שניהם איברי זהות לחיבור במרחב וקטורי כלשהו V. אזי

$$0' = 0' + 0 = 0 + 0' = 0,$$

כאשר השוויון הראשון תקף מכיוון ש־ 0 הוא איבר זהות לחיבור, השוויון השני תקף בשל תכונת הקומוטטיביות, ואילו השוויון השלישי מתקיים מכיוון ש־ $0'$ הוא איבר זהות לחיבור בעצמו. לפיכך, $0 = 0'$, ובכך הראינו כי ל־ V יש איבר זהות לחיבור יחיד. ∎

לכל איבר v במרחב וקטורי יש איבר נגדי לחיבור, איבר w במרחב הווקטורי כך ש־ $v + w = 0$. המסקנה הבאה מראה שלכל איבר במרחב וקטורי יש איבר נגדי לחיבור יחיד.

1.27 *unique additive inverse*

Every element in a vector space has a unique additive inverse.

Proof Suppose V is a vector space. Let $v \in V$. Suppose w and w' are additive inverses of v. Then

$$w = w + 0 = w + (v + w') = (w + v) + w' = 0 + w' = w'.$$

Thus $w = w'$, as desired. ∎

Because additive inverses are unique, the following notation now makes sense.

1.28 notation: $-v$, $w - v$

Let $v, w \in V$. Then

- $-v$ denotes the additive inverse of v;
- $w - v$ is defined to be $w + (-v)$.

Almost all results in this book involve some vector space. To avoid having to restate frequently that V is a vector space, we now make the necessary declaration once and for all.

1.29 notation: V

For the rest of this book, V denotes a vector space over $\mathbf{F}$.

In the next result, 0 denotes a scalar (the number $0 \in \mathbf{F}$) on the left side of the equation and a vector (the additive identity of V) on the right side of the equation.

1.30 *the number* 0 *times a vector*

$0v = 0$ for every $v \in V$.

Proof For $v \in V$, we have

$$0v = (0 + 0)v = 0v + 0v.$$

Adding the additive inverse of $0v$ to both sides of the equation above gives $0 = 0v$, as desired. ∎

The result in 1.30 involves the additive identity of V and scalar multiplication. The only part of the definition of a vector space that connects vector addition and scalar multiplication is the distributive property. Thus the distributive property must be used in the proof of 1.30.

In the next result, 0 denotes the additive identity of V. Although their proofs are similar, 1.30 and 1.31 are not identical. More precisely, 1.30 states that the product of the scalar 0 and any vector equals the vector 0, whereas 1.31 states that the product of any scalar and the vector 0 equals the vector 0.

1.27 יחידות איבר נגדי לחיבור

לכל איבר במרחב וקטורי יש איבר נגדי לחיבור יחיד.

הוכחה נניח ש־ V הוא מרחב וקטורי. יהי $v \in V$. נניח ש־ w ו־ w' איברים נגדיים לחיבור של v. אזי

$$w = w + 0 = w + (v + w') = (w + v) + w' = 0 + w' = w'.$$

לפיכך $w = w'$, כפי שרצינו. ■

מאחר שאיברים נגדיים לחיבור הם יחידים, הסימון הבא מקבל כעת משמעות.

1.28 סימון: $-v$, $w - v$

יהיו $v, w \in V$. אזי

- $-v$ מסמן את האיבר הנגדי לחיבור של v;
- $w - v$ מוגדר כ־ $w + (-v)$.

כמעט כל המסקנות בספר עוסקות במרחב וקטורי כלשהו. כדי להימנע מלהגדיר בכל פעם מחדש את V כמרחב וקטורי, ננסח כעת את ההצהרה הנדרשת אחת ולתמיד.

1.29 סימון: V

מכאן ואילך בספר זה, V מסמן מרחב וקטורי מעל $\mathbf{F}$.

במסקנה הבאה, 0 מסמן סקלר (המספר $0 \in \mathbf{F}$) באגף השמאלי של המשוואה, ואילו באגף הימני הוא מסמן וקטור (איבר הזהות לחיבור ב־ V).

1.30 המספר 0 כפול וקטור

$0v = 0$ לכל $v \in V$.

הוכחה עבור $v \in V$, מתקיים

$$0v = (0 + 0)v = 0v + 0v.$$

כשנוסיף את האיבר הנגדי לחיבור של $0v$ לכל אגף של המשוואה לעיל, נקבל $0 = 0v$, כפי שרצינו. ■

מסקנה 1.30 עוסקת באיבר הזהות לחיבור של V ובכפל בסקלר. בהגדרת מרחב וקטורי, תכונת הדיסטריבוטיביות היא הגורם היחיד המקשר בין חיבור וקטורים לבין כפל בסקלר. לפיכך בהוכחת 1.30 עלינו להשתמש בתכונה זו.

במסקנה הבאה, 0 מסמן את איבר הזהות לחיבור ב־ V. למרות ההוכחות הדומות, 1.30 ו־ 1.31 אינן זהות. ביתר דיוק, 1.30 קובעת שמכפלת הסקלר 0 בכל וקטור היא וקטור האפס, ואילו 1.31 קובעת שמכפלת כל סקלר בווקטור האפס היא וקטור האפס.

1.31 *a number times the vector 0*

$a0 = 0$ for every $a \in \mathbf{F}$.

Proof For $a \in \mathbf{F}$, we have

$$a0 = a(0 + 0) = a0 + a0.$$

Adding the additive inverse of $a0$ to both sides of the equation above gives $0 = a0$, as desired.

Now we show that if an element of V is multiplied by the scalar -1, then the result is the additive inverse of the element of V.

1.32 *the number -1 times a vector*

$(-1)v = -v$ for every $v \in V$.

Proof For $v \in V$, we have

$$v + (-1)v = 1v + (-1)v = (1 + (-1))v = 0v = 0.$$

This equation says that $(-1)v$, when added to v, gives 0. Thus $(-1)v$ is the additive inverse of v, as desired.

Exercises 1B

1 Prove that $-(-v) = v$ for every $v \in V$.

2 Suppose $a \in \mathbf{F}$, $v \in V$, and $av = 0$. Prove that $a = 0$ or $v = 0$.

3 Suppose $v, w \in V$. Explain why there exists a unique $x \in V$ such that $v + 3x = w$.

4 The empty set is not a vector space. The empty set fails to satisfy only one of the requirements listed in the definition of a vector space (1.20). Which one?

5 Show that in the definition of a vector space (1.20), the additive inverse condition can be replaced with the condition that

$$0v = 0 \text{ for all } v \in V.$$

Here the 0 on the left side is the number 0, and the 0 on the right side is the additive identity of V.

The phrase a "condition can be replaced" in a definition means that the collection of objects satisfying the definition is unchanged if the original condition is replaced with the new condition.

1.31 מספר כפול וקטור האפס

$0a = 0$ לכל $a \in \mathbf{F}$.

הוכחה לכל $a \in \mathbf{F}$, מתקיים

$$a0 = a(0 + 0) = a0 + a0.$$

כשנוסיף את האיבר הנגדי לחיבור של $a0$ לכל אגף של המשוואה לעיל, נקבל $0 = a0$, כפי שרצינו.

כעת נראה שאם איבר ב־ V מוכפל בסקלר -1, התוצאה היא האיבר הנגדי לחיבור של האיבר ב־ V.

1.32 המספר -1 כפול וקטור

$(-1)v = -v$ לכל $v \in V$.

הוכחה לכל $v \in V$, מתקיים

$$v + (-1)v = 1v + (-1)v = (1 + (-1))v = 0v = 0.$$

משוואה זו מראה שכאשר מחברים $(-1)v$ ל־ v מקבלים 0. לפיכך $(-1)v$ הוא האיבר הנגדי לחיבור של v, כפי שרצינו.

תרגילים אג

1 הוכיחו כי $-(-v) = v$ לכל $v \in V$.

2 נניח ש־ $a \in \mathbf{F}$, $v \in V$ ו־ $av = 0$. הוכיחו כי $a = 0$ או $v = 0$.

3 נניח ש־ $v, w \in V$. הסבירו מדוע קיים איבר יחיד $x \in V$ כך ש־ $v + 3x = w$.

4 הקבוצה הריקה אינה מרחב וקטורי. הקבוצה הריקה מקיימת את כל הדרישות שהוגדרו עבור מרחבים וקטוריים (1.20) מלבד אחת. מהי הדרישה?

5 הראו כי בהגדרת מרחב וקטורי (1.20), אפשר להחליף את התנאי לאיבר נגדי לחיבור בתנאי הדורש

$$0v = 0 \text{ for all } v \in V.$$

כאן ה־ 0 באגף השמאלי הוא המספר 0, ואילו באגף הימני ה־ 0 הוא איבר הזהות לחיבור ב־ V.

הביטוי "אפשר להחליף את התנאי" בהגדרה משמעותו שאוסף האובייקטים המקיימים את ההגדרה אינו משתנה אם התנאי המקורי מוחלף בתנאי החדש.

6 Let ∞ and $-\infty$ denote two distinct objects, neither of which is in $\mathbf{R}$. Define an addition and scalar multiplication on $\mathbf{R} \cup \{\infty, -\infty\}$ as you could guess from the notation. Specifically, the sum and product of two real numbers is as usual, and for $t \in \mathbf{R}$ define

$$t\infty = \begin{cases} -\infty & \text{if } t < 0, \\ 0 & \text{if } t = 0, \\ \infty & \text{if } t > 0, \end{cases} \qquad t(-\infty) = \begin{cases} \infty & \text{if } t < 0, \\ 0 & \text{if } t = 0, \\ -\infty & \text{if } t > 0, \end{cases}$$

and

$$\begin{aligned} t + \infty = \infty + t = \infty + \infty = \infty, \\ t + (-\infty) = (-\infty) + t = (-\infty) + (-\infty) = -\infty, \\ \infty + (-\infty) = (-\infty) + \infty = 0. \end{aligned}$$

With these operations of addition and scalar multiplication, is $\mathbf{R} \cup \{\infty, -\infty\}$ a vector space over $\mathbf{R}$? Explain.

7 Suppose S is a nonempty set. Let V^S denote the set of functions from S to V. Define a natural addition and scalar multiplication on V^S, and show that V^S is a vector space with these definitions.

8 Suppose V is a real vector space.

- The *complexification* of V, denoted by $V_{\mathbf{C}}$, equals $V \times V$. An element of $V_{\mathbf{C}}$ is an ordered pair (u, v), where $u, v \in V$, but we write this as $u + iv$.
- Addition on $V_{\mathbf{C}}$ is defined by

$$(u_1 + iv_1) + (u_2 + iv_2) = (u_1 + u_2) + i(v_1 + v_2)$$

for all $u_1, v_1, u_2, v_2 \in V$.

- Complex scalar multiplication on $V_{\mathbf{C}}$ is defined by

$$(a + bi)(u + iv) = (au - bv) + i(av + bu)$$

for all $a, b \in \mathbf{R}$ and all $u, v \in V$.

Prove that with the definitions of addition and scalar multiplication as above, $V_{\mathbf{C}}$ is a complex vector space.

Think of V as a subset of $V_{\mathbf{C}}$ by identifying $u \in V$ with $u + i0$. The construction of $V_{\mathbf{C}}$ from V can then be thought of as generalizing the construction of $\mathbf{C}^n$ from $\mathbf{R}^n$.

6 נסמן ב־ ∞ ו־ $-\infty$ שני אובייקטים שונים, שאף אחד מהם אינו ב־ $\mathbf{R}$. הגדירו פעולת חיבור וכפל בסקלר מעל $\mathbf{R} \cup \{\infty, -\infty\}$ כפי שניתן לשער מהסימון. בפרט הסכום והמכפלה של שני מספרים ממשיים מוגדרים כרגיל, ועבור $t \in \mathbf{R}$ נגדיר

$$t\infty = \begin{cases} -\infty & \text{if } t < 0, \\ 0 & \text{if } t = 0, \\ \infty & \text{if } t > 0, \end{cases} \qquad t(-\infty) = \begin{cases} \infty & \text{if } t < 0, \\ 0 & \text{if } t = 0, \\ -\infty & \text{if } t > 0, \end{cases}$$

כמו כן

$$\begin{aligned} t + \infty = \infty + t = \infty + \infty = \infty, \\ t + (-\infty) = (-\infty) + t = (-\infty) + (-\infty) = -\infty, \\ \infty + (-\infty) = (-\infty) + \infty = 0. \end{aligned}$$

עם פעולות חיבור וכפל בסקלר אלו, האם $\mathbf{R} \cup \{\infty, -\infty\}$ הוא מרחב וקטורי מעל $\mathbf{R}$? הסבירו.

7 נניח ש־ S קבוצה שאינה ריקה. יהי V^S סימון לקבוצת הפונקציות מ־ S ל V. הגדירו פעולות חיבור וכפל בסקלר טבעיות מעל V^S, והראו שאכן V^S הוא מרחב וקטורי לפי הגדרות אלו.

8 נניח ש־ V מרחב וקטורי ממשי.

- המירכוּב של V, המסומן ב־ $V_{\mathbf{C}}$, שווה ל־ $V \times V$. איבר ב־ $V_{\mathbf{C}}$ הוא זוג סדור (u, v), כאשר $u, v \in V$, אך אנו נרשום אותו כ־ $u + iv$.
- חיבור מעל $V_{\mathbf{C}}$ מוגדר על ידי

$$(u_1 + iv_1) + (u_2 + iv_2) = (u_1 + u_2) + i(v_1 + v_2)$$

לכל $u_1, v_1, u_2, v_2 \in V$.

- כפל בסקלר מרוכב מעל $V_{\mathbf{C}}$ מוגדר על ידי

$$(a + bi)(u + iv) = (au - bv) + i(av + bu)$$

לכל $a, b \in \mathbf{R}$ ולכל $u, v \in V$.

הוכיחו כי עם ההגדרות של החיבור והכפל בסקלר מרוכב כמובא לעיל, $V_{\mathbf{C}}$ הוא מרחב וקטורי מרוכב.

חשבו על V כתת־קבוצה של $V_{\mathbf{C}}$ על ידי כך שתזהו את $u \in V$ עם $u + i0$. על בניית $V_{\mathbf{C}}$ מתוך V נוכל לחשוב כעל הכללה של בניית $\mathbf{R}^n$ מתוך $\mathbf{C}^n$.

אלגברה ליניארית בדרך הישר, מהדורה רביעית, מחבר: שלדון אקסלר
Linear Algebra Done Right, fourth edition, by Sheldon Axler

1C Subspaces

By considering subspaces, we can greatly expand our examples of vector spaces.

1.33 definition: *subspace*

A subset U of V is called a *subspace* of V if U is also a vector space with the same additive identity, addition, and scalar multiplication as on V.

The next result gives the easiest way to check whether a subset of a vector space is a subspace.

*Some people use the terminology **linear subspace**, which means the same as subspace.*

1.34 *conditions for a subspace*

A subset U of V is a subspace of V if and only if U satisfies the following three conditions.

additive identity
$0 \in U$.

closed under addition
$u, w \in U$ implies $u + w \in U$.

closed under scalar multiplication
$a \in \mathbf{F}$ and $u \in U$ implies $au \in U$.

Proof If U is a subspace of V, then U satisfies the three conditions above by the definition of vector space.

Conversely, suppose U satisfies the three conditions above. The first condition ensures that the additive identity of V is in U. The second condition ensures that addition makes sense on U. The third condition ensures that scalar multiplication makes sense on U.

The additive identity condition above could be replaced with the condition that U is nonempty (because then taking $u \in U$ and multiplying it by 0 would imply that $0 \in U$). However, if a subset U of V is indeed a subspace, then usually the quickest way to show that U is nonempty is to show that $0 \in U$.

If $u \in U$, then $-u$ [which equals $(-1)u$ by 1.32] is also in U by the third condition above. Hence every element of U has an additive inverse in U.

The other parts of the definition of a vector space, such as associativity and commutativity, are automatically satisfied for U because they hold on the larger space V. Thus U is a vector space and hence is a subspace of V. ∎

The three conditions in the result above usually enable us to determine quickly whether a given subset of V is a subspace of V. You should verify all assertions in the next example.

אג תתי־מרחב

בכך שנתבונן בתתי־מרחב, נוכל להרחיב באופן ניכר את הדוגמאות של מרחבים וקטוריים.

1.33 הגדרה: תת־מרחב

תת־קבוצה U של V נקראת תת־מרחב של V אם U גם היא מרחב וקטורי עם אותו איבר זהות לחיבור, פעולת חיבור, וכפל בסקלר כמו אלו שמוגדרים מעל V.

המסקנה הבאה מעניקה לנו את הדרך הקלה ביותר לבדוק האם תת־קבוצה של מרחב וקטורי היא תת־מרחב.

יש המשתמשים במונח **תת־מרחב ליניארי**, שמשמעותו זהה לתת־מרחב.

1.34 התנאים לתת־מרחב

תת־קבוצה U של V היא תת־מרחב של V אם־ורק־אם U מקיימת את שלושת התנאים הבאים:

איבר זהות לחיבור
$0 \in U$.

סגירות תחת חיבור
$u, w \in U$ גורר ש־ $u + w \in U$.

סגירות תחת כפל בסקלר
$a \in \mathbf{F}$ ו־ $u \in U$ גורר ש־ $au \in U$.

הוכחה

אם U תת־מרחב של V, אזי U, כתת־מרחב, מקיים את שלושת התנאים לעיל לפי הגדרת מרחב וקטורי.

בכיוון ההפוך, נניח ש־ U, כתת־קבוצה, מקיימת את שלושת התנאים הללו. התנאי הראשון מבטיח שאיבר הזהות לחיבור של V נמצא ב־ U. התנאי השני מבטיח לנו שניתן לחבר איברים ב־ U. התנאי השלישי מבטיח שנוכל להכפיל בסקלר ב־ U.

את התנאי לקיומו של איבר זהות לחיבור לעיל נוכל להחליף בתנאי ש־ U אינה ריקה (מאחר שאם ניקח $u \in U$ ונכפיל אותו ב־ 0 נקבל ש־ $0 \in U$). בכל זאת, אם U, תת־קבוצה של V, היא אכן תת־מרחב, בדרך כלל הדרך המהירה ביותר להראות ש־ U אינה ריקה היא להראות ש־ $0 \in U$.

אם $u \in U$, אזי $-u$ [השווה ל־ $(-1)u$ לפי 1.32] גם הוא ב־ U לפי התנאי השלישי. לכן, לכל איבר ב־ U, קיים איבר נגדי לחיבור ב־ U.

שאר החלקים בהגדרת מרחב וקטורי, דוגמת האסוציאטיביות והקומוטטיביות, מתקיימים ב־ U מאליהם, מאחר שהם מתקיימים במרחב הגדול יותר V. לפיכך U מרחב וקטורי ולכן הוא גם תת־מרחב של V. ■

שלושת התנאים במסקנה לעיל בדרך כלל מאפשרים לנו לקבוע בזריזות האם תת־קבוצה נתונה של V היא תת־מרחב של V. עליכם לוודא את כל הטענות המופיעות בדוגמה הבאה.

1.35 example: *subspaces*

(a) If $b \in \mathbf{F}$, then
$$\{(x_1, x_2, x_3, x_4) \in \mathbf{F}^4 : x_3 = 5x_4 + b\}$$
is a subspace of $\mathbf{F}^4$ if and only if $b = 0$.

(b) The set of continuous real-valued functions on the interval $[0, 1]$ is a subspace of $\mathbf{R}^{[0,1]}$.

(c) The set of differentiable real-valued functions on $\mathbf{R}$ is a subspace of $\mathbf{R}^{\mathbf{R}}$.

(d) The set of differentiable real-valued functions f on the interval $(0, 3)$ such that $f'(2) = b$ is a subspace of $\mathbf{R}^{(0,3)}$ if and only if $b = 0$.

(e) The set of all sequences of complex numbers with limit 0 is a subspace of $\mathbf{C}^{\infty}$.

Verifying some of the items above shows the linear structure underlying parts of calculus. For example, (b) above requires the result that the sum of two continuous functions is continuous. As another example, (d) above requires the result that for a constant c, the derivative of cf equals c times the derivative of f.

The set $\{0\}$ is the smallest subspace of V, and V itself is the largest subspace of V. The empty set is not a subspace of V because a subspace must be a vector space and hence must contain at least one element, namely, an additive identity.

The subspaces of $\mathbf{R}^2$ are precisely $\{0\}$, all lines in $\mathbf{R}^2$ containing the origin, and $\mathbf{R}^2$. The subspaces of $\mathbf{R}^3$ are precisely $\{0\}$, all lines in $\mathbf{R}^3$ containing the origin, all planes in $\mathbf{R}^3$ containing the origin, and $\mathbf{R}^3$. To prove that all these objects are indeed subspaces is straightforward—the hard part is to show that they are the only subspaces of $\mathbf{R}^2$ and $\mathbf{R}^3$. That task will be easier after we introduce some additional tools in the next chapter.

Sums of Subspaces

When dealing with vector spaces, we are usually interested only in subspaces, as opposed to arbitrary subsets. The notion of the sum of subspaces will be useful.

The union of subspaces is rarely a subspace (see Exercise 12), which is why we usually work with sums rather than unions.

1.36 definition: *sum of subspaces*

Suppose $V_1, \dots, V_m$ are subspaces of V. The *sum* of $V_1, \dots, V_m$, denoted by $V_1 + \cdots + V_m$, is the set of all possible sums of elements of $V_1, \dots, V_m$. More precisely,
$$V_1 + \cdots + V_m = \{v_1 + \cdots + v_m : v_1 \in V_1, \dots, v_m \in V_m\}.$$

1.35 דוגמה: תתי־מרחב

א. אם $b \in \mathbf{F}$, אזי

$$\{(x_1, x_2, x_3, x_4) \in \mathbf{F}^4 : x_3 = 5x_4 + b\}$$

הוא תת־מרחב של $\mathbf{F}^4$ אם־ורק־אם $b = 0$.

ב. קבוצת הפונקציות הממשיות הרציפות מעל הקטע $[0, 1]$ היא תת־מרחב של $\mathbf{R}^{[0,1]}$.

ג. קבוצת הפונקציות הממשיות הגזירות מעל $\mathbf{R}$ היא תת־מרחב של $\mathbf{R}^{\mathbf{R}}$.

ד. קבוצת הפונקציות הממשיות הגזירות f מעל הקטע $(0, 3)$ כך ש־ $f'(2) = b$ היא תת־מרחב של $\mathbf{R}^{(0,3)}$ אם־ורק־אם $b = 0$.

ה. קבוצת כל הסדרות של מספרים מרוכבים שגבולן 0 היא תת־מרחב של $\mathbf{C}^{\infty}$.

הקבוצה $\{0\}$ היא תת־המרחב הקטן ביותר של V, ואילו V עצמו הוא תת־המרחב הגדול ביותר של V. הקבוצה הריקה אינה תת־מרחב של V מאחר שתת־מרחב חייב להיות מרחב וקטורי, נדרש שיהיה בו איבר אחד לפחות, הלא הוא איבר הזהות לחיבור.

אימות חלק מהסעיפים לעיל חושף את המבנה הליניארי המונח ביסוד חלק מהחשבון האינפיניטסימלי. לדוגמה, לאימות סעיף (ב) נדרשת המסקנה כי סכום שתי פונקציות רציפות הוא פונקציה רציפה. דוגמה נוספת, לסעיף (ד) לעיל נדרשת המסקנה כי עבור קבוע c, הנגזרת של cf שווה ל־ c פעמים הנגזרת של f.

תתי־המרחב של $\mathbf{R}^2$ הם: $\{0\}$, כל הישרים ב־ $\mathbf{R}^2$ העוברים דרך ראשית הצירים, ו־ $\mathbf{R}^2$ עצמו בלבד. תתי־המרחב של $\mathbf{R}^3$ הם: $\{0\}$, כל הישרים ב־ $\mathbf{R}^3$ העוברים דרך ראשית הצירים, כל המישורים ב־ $\mathbf{R}^3$ העוברים דרך ראשית הצירים, ו־ $\mathbf{R}^3$ עצמו בלבד. ההוכחה שאובייקטים אלו תתי־מרחב היא פשוטה—החלק המאתגר יותר הוא להראות שרק הם תתי־המרחב של $\mathbf{R}^2$ ו־ $\mathbf{R}^3$. משימה זו תהפוך לקלה יותר אחרי שבפרק הבא נכיר כלים נוספים.

סכומי תתי־מרחב

איחוד של תתי־מרחב לעיתים נדירות יהיה תת־מרחב (ראו תרגיל 12); זו הסיבה שלרוב נעסוק בסכומים ולא באיחודים.

כשאנו עוסקים במרחבים וקטוריים, לרוב נגלה עניין בתתי־מרחב ולא בתתי־קבוצה שרירותיות. הרעיון של סכום תתי־מרחב יהיה שימושי.

1.36 הגדרה: סכום תתי־מרחב

נניח ש־ $V_1, \dots, V_m$ תתי־מרחב של V. הסכום של $V_1, \dots, V_m$, המסומן ב־ $V_1 + \cdots + V_m$, הוא קבוצת כל הסכומים האפשריים של איברי $V_1, \dots, V_m$. וביתר דיוק,

$$V_1 + \cdots + V_m = \{v_1 + \cdots + v_m : v_1 \in V_1, \dots, v_m \in V_m\}.$$

Let's look at some examples of sums of subspaces.

1.37 example: *a sum of subspaces of* $\mathbf{F}^3$

Suppose U is the set of all elements of $\mathbf{F}^3$ whose second and third coordinates equal 0, and W is the set of all elements of $\mathbf{F}^3$ whose first and third coordinates equal 0:

$$U = \{(x, 0, 0) \in \mathbf{F}^3 : x \in \mathbf{F}\} \quad \text{and} \quad W = \{(0, y, 0) \in \mathbf{F}^3 : y \in \mathbf{F}\}.$$

Then

$$U + W = \{(x, y, 0) \in \mathbf{F}^3 : x, y \in \mathbf{F}\},$$

as you should verify.

1.38 example: *a sum of subspaces of* $\mathbf{F}^4$

Suppose

$$U = \{(x, x, y, y) \in \mathbf{F}^4 : x, y \in \mathbf{F}\} \quad \text{and} \quad W = \{(x, x, x, y) \in \mathbf{F}^4 : x, y \in \mathbf{F}\}.$$

Using words rather than symbols, we could say that U is the set of elements of $\mathbf{F}^4$ whose first two coordinates equal each other and whose third and fourth coordinates equal each other. Similarly, W is the set of elements of $\mathbf{F}^4$ whose first three coordinates equal each other.

To find a description of $U + W$, consider a typical element (a, a, b, b) of U and a typical element (c, c, c, d) of W, where $a, b, c, d \in \mathbf{F}$. We have

$$(a, a, b, b) + (c, c, c, d) = (a + c, a + c, b + c, b + d),$$

which shows that every element of $U + W$ has its first two coordinates equal to each other. Thus

1.39 $$U + W \subseteq \{(x, x, y, z) \in \mathbf{F}^4 : x, y, z \in \mathbf{F}\}.$$

To prove the inclusion in the other direction, suppose $x, y, z \in \mathbf{F}$. Then

$$(x, x, y, z) = (x, x, y, y) + (0, 0, 0, z - y),$$

where the first vector on the right is in U and the second vector on the right is in W. Thus $(x, x, y, z) \in U + W$, showing that the inclusion 1.39 also holds in the opposite direction. Hence

$$U + W = \{(x, x, y, z) \in \mathbf{F}^4 : x, y, z \in \mathbf{F}\},$$

which shows that $U + W$ is the set of elements of $\mathbf{F}^4$ whose first two coordinates equal each other.

The next result states that the sum of subspaces is a subspace, and is in fact the smallest subspace containing all the summands (which means that every subspace containing all the summands also contains the sum).

נכיר כעת מספר דוגמאות של סכומי תתי־מרחב.

1.37 דוגמה: סכום תתי־מרחב של $\mathbf{F}^3$

נניח ש־ U היא קבוצת כל איברי $\mathbf{F}^3$ שבהם הקואורדינטות השנייה והשלישית הן 0, ו־ W היא קבוצת כל איברי $\mathbf{F}^3$ שבהם הקואורדינטות הראשונה והשלישית הן 0:

$$U = \{(x, 0, 0) \in \mathbf{F}^3 : x \in \mathbf{F}\} \quad \text{וגם} \quad W = \{(0, y, 0) \in \mathbf{F}^3 : y \in \mathbf{F}\}.$$

אזי

$$U + W = \{(x, y, 0) \in \mathbf{F}^3 : x, y \in \mathbf{F}\},$$

כפי שעליכם לוודא.

1.38 דוגמה: סכום תתי־מרחב של $\mathbf{F}^4$

נניח ש־

$$U = \{(x, x, y, y) \in \mathbf{F}^4 : x, y \in \mathbf{F}\} \quad \text{ו־} \quad W = \{(x, x, x, y) \in \mathbf{F}^4 : x, y \in \mathbf{F}\}.$$

בתיאור מילולי נאמר ש־ U היא קבוצת האיברים ב־ $\mathbf{F}^4$ שערכיהם בצמד הקואורדינטות הראשון זהים, וערך זהה בצמד הקואורדינטות האחרון. באופן דומה נתאר את W כקבוצת האיברים ב־ $\mathbf{F}^4$ שבהם ערך זהה בשלוש הקואורדינטות הראשונות.

לתיאור $U + W$ ניקח איבר טיפוסי של U מהצורה (a, a, b, b) ואיבר טיפוסי של W מהצורה (c, c, c, d) כאשר $a, b, c, d \in \mathbf{F}$, ונקבל:

$$(a, a, b, b) + (c, c, c, d) = (a + c, a + c, b + c, b + d),$$

מכך נסיק כי ערכי צמד הקואורדינטות הראשון בכל איבר ב־ $U + W$ שווים. לפיכך

1.39 $$U + W \subseteq \{(x, x, y, z) \in \mathbf{F}^4 : x, y, z \in \mathbf{F}\}.$$

כדי להוכיח את ההכלה בכיוון השני, נניח ש־ $x, y, z \in \mathbf{F}$ ונקבל

$$(x, x, y, z) = (x, x, y, y) + (0, 0, 0, z - y),$$

כאשר הווקטור הראשון באגף הימני של המשוואה הוא ב־ U והווקטור השני הוא ב־ W. לפיכך $(x, x, y, z) \in U + W$, מה שמראה כי הכלה 1.39 מתקיימת גם בכיוון ההפוך. לכן

$$U + W = \{(x, x, y, z) \in \mathbf{F}^4 : x, y, z \in \mathbf{F}\},$$

מה שמראה ש־ $U + W$ היא קבוצת האיברים ב־ $\mathbf{F}^4$ שערכיהם בצמד הקואורדינטות הראשון שווים.

המסקנה הבאה קובעת כי סכום של תתי־מרחב הוא תת־מרחב, ולמעשה הוא תת־המרחב הקטן ביותר המכיל את כל המחוברים (משמעות הדבר היא שכל תת־מרחב המכיל את כל המחוברים מכיל גם את הסכום).

1.40 *sum of subspaces is the smallest containing subspace*

Suppose $V_1, \dots, V_m$ are subspaces of V. Then $V_1 + \cdots + V_m$ is the smallest subspace of V containing $V_1, \dots, V_m$.

Proof The reader can verify that $V_1 + \cdots + V_m$ contains the additive identity 0 and is closed under addition and scalar multiplication. Thus 1.34 implies that $V_1 + \cdots + V_m$ is a subspace of V.

The subspaces $V_1, \dots, V_m$ are all contained in $V_1 + \cdots + V_m$ (to see this, consider sums $v_1 + \cdots + v_m$ where all except one of the v_k's are 0). Conversely, every subspace of V containing $V_1, \dots, V_m$ contains $V_1 + \cdots + V_m$ (because subspaces must contain all finite sums of their elements). Thus $V_1 + \cdots + V_m$ is the smallest subspace of V containing $V_1, \dots, V_m$. ∎

Sums of subspaces in the theory of vector spaces are analogous to unions of subsets in set theory. Given two subspaces of a vector space, the smallest subspace containing them is their sum. Analogously, given two subsets of a set, the smallest subset containing them is their union.

Direct Sums

Suppose $V_1, \dots, V_m$ are subspaces of V. Every element of $V_1 + \cdots + V_m$ can be written in the form

$$v_1 + \cdots + v_m,$$

where each $v_k \in V_k$. Of special interest are cases in which each vector in $V_1 + \cdots + V_m$ can be represented in the form above in only one way. This situation is so important that it gets a special name (direct sum) and a special symbol ($\oplus$).

1.41 definition: *direct sum,* $\oplus$

Suppose $V_1, \dots, V_m$ are subspaces of V.

- The sum $V_1 + \cdots + V_m$ is called a *direct sum* if each element of $V_1 + \cdots + V_m$ can be written in only one way as a sum $v_1 + \cdots + v_m$, where each $v_k \in V_k$.
- If $V_1 + \cdots + V_m$ is a direct sum, then $V_1 \oplus \cdots \oplus V_m$ denotes $V_1 + \cdots + V_m$, with the $\oplus$ notation serving as an indication that this is a direct sum.

1.42 example: *a direct sum of two subspaces*

Suppose U is the subspace of $\mathbf{F}^3$ of those vectors whose last coordinate equals 0, and W is the subspace of $\mathbf{F}^3$ of those vectors whose first two coordinates equal 0:

$$U = \{(x, y, 0) \in \mathbf{F}^3 : x, y \in \mathbf{F}\} \quad \text{and} \quad W = \{(0, 0, z) \in \mathbf{F}^3 : z \in \mathbf{F}\}.$$

Then $\mathbf{F}^3 = U \oplus W$, as you should verify.

1.40 סכום תתי־מרחב הוא התת־מרחב הקטן ביותר המכיל את תתי־המרחב

נניח ש־ $V_1, \dots, V_m$ תתי־מרחב של V. אזי, $V_1 + \cdots + V_m$ הוא התת־מרחב הקטן ביותר של V המכיל את $V_1, \dots, V_m$.

סכום תתי־מרחב בתורת המרחבים הווקטוריים דומה לאיחוד תתי־קבוצה בתורת הקבוצות. בהינתן שני תתי־מרחב של מרחב וקטורי כלשהו, התת־מרחב הקטן ביותר המכיל אותם הוא סכומם. באופן דומה, בהינתן שתי תתי־קבוצה של קבוצה כלשהי, התת־קבוצה הקטנה ביותר המכילה אותן היא איחודן.

הוכחה ניתן לוודא כי $V_1 + \cdots + V_m$ מכיל את איבר הזהות לחיבור 0 וסגור תחת חיבור וכפל בסקלר. לפיכך מ־ 1.34 נסיק כי $V_1 + \cdots + V_m$ תת־מרחב של V.

תתי־המרחב $V_1, \dots, V_m$ כולם מוכלים ב־ $V_1 + \cdots + V_m$ (כדי להיווכח בכך, נשקול סכומים $v_1 + \cdots + v_m$ כאשר כל האיברים מלבד אחד מה־ v_k־ים הם איבר האפס). בכיוון ההפוך, כל תת־מרחב של V המכיל את $V_1, \dots, V_m$ מכיל גם את $V_1 + \cdots + V_m$ (שכן על התת־מרחב להכיל את כל הסכומים הסופיים של איבריו). לפיכך $V_1 + \cdots + V_m$ הוא התת־מרחב הקטן ביותר של V המכיל את $V_1, \dots, V_m$. ■

סכומים ישרים

נניח ש־ $V_1, \dots, V_m$ תתי־מרחב של V. כל איבר ב־ $V_1 + \cdots + V_m$ יכול להיכתב באופן הבא

$$v_1 + \cdots + v_m,$$

כאשר כל $v_k \in V_k$. קיים עניין מיוחד במקרים שבהם הייצוג של כל וקטור ב־ $V_1 + \cdots + V_m$ בצורה לעיל הוא יחיד. מצב זה חשוב עד כדי כך שהוענק לו שם מיוחד (סכום ישר) ואפילו סימון מיוחד ($\oplus$).

1.41 הגדרה: סכום ישר, $\oplus$

נניח ש־ $V_1, \dots, V_m$ תתי־מרחב של V.

- הסכום $V_1 + \cdots + V_m$ נקרא סכום ישר אם כל איבר של $V_1 + \cdots + V_m$ יכול להיכתב בדרך אחת בלבד כסכום $v_1 + \cdots + v_m$, כאשר כל $v_k \in V_k$.
- אם $V_1 + \cdots + V_m$ סכום ישר, אזי נסמן את $V_1 + \cdots + V_m$ כ־ $V_1 \oplus \cdots \oplus V_m$ כאשר הסימון $\oplus$ משמש כאות לכך שהסכום הוא סכום ישר.

1.42 דוגמה: סכום ישר של שני תתי־מרחב

נניח ש־ U הוא תת־מרחב של $\mathbf{F}^3$ המורכב מכל הווקטורים שהקואורדינטה האחרונה שלהם שווה 0, ו־ W הוא תת־מרחב של $\mathbf{F}^3$ המורכב מכל הווקטורים ששתי הקואורדינטות הראשונות שלהם שוות 0:

$$U = \{(x, y, 0) \in \mathbf{F}^3 : x, y \in \mathbf{F}\} \quad \text{וגם} \quad W = \{(0, 0, z) \in \mathbf{F}^3 : z \in \mathbf{F}\}.$$

אזי, $\mathbf{F}^3 = U \oplus W$, כפי שעליכם לוודא.

1.43 example: *a direct sum of multiple subspaces*

Suppose V_k is the subspace of $\mathbf{F}^n$ of those vectors whose coordinates are all 0, except possibly in the kth slot; for example, $V_2 = \{(0, x, 0, \dots, 0) \in \mathbf{F}^n : x \in \mathbf{F}\}$. Then

To produce $\oplus$ in TEX, type `\oplus`.

$$\mathbf{F}^n = V_1 \oplus \cdots \oplus V_n,$$

as you should verify.

Sometimes nonexamples add to our understanding as much as examples.

1.44 example: *a sum that is not a direct sum*

Suppose

$$V_1 = \{(x, y, 0) \in \mathbf{F}^3 : x, y \in \mathbf{F}\},$$

$$V_2 = \{(0, 0, z) \in \mathbf{F}^3 : z \in \mathbf{F}\},$$

$$V_3 = \{(0, y, y) \in \mathbf{F}^3 : y \in \mathbf{F}\}.$$

Then $\mathbf{F}^3 = V_1 + V_2 + V_3$ because every vector $(x, y, z) \in \mathbf{F}^3$ can be written as

$$(x, y, z) = (x, y, 0) + (0, 0, z) + (0, 0, 0),$$

where the first vector on the right side is in V_1, the second vector is in V_2, and the third vector is in V_3.

However, $\mathbf{F}^3$ does not equal the direct sum of V_1, V_2, V_3, because the vector $(0, 0, 0)$ can be written in more than one way as a sum $v_1 + v_2 + v_3$, with each $v_k \in V_k$. Specifically, we have

$$(0, 0, 0) = (0, 1, 0) + (0, 0, 1) + (0, -1, -1)$$

and, of course,

$$(0, 0, 0) = (0, 0, 0) + (0, 0, 0) + (0, 0, 0),$$

where the first vector on the right side of each equation above is in V_1, the second vector is in V_2, and the third vector is in V_3. Thus the sum $V_1 + V_2 + V_3$ is not a direct sum.

The definition of direct sum requires every vector in the sum to have a unique representation as an appropriate sum. The next result shows that when deciding whether a sum of subspaces is a direct sum, we only need to consider whether 0 can be uniquely written as an appropriate sum.

The symbol $\oplus$, which is a plus sign inside a circle, reminds us that we are dealing with a special type of sum of subspaces—each element in the direct sum can be represented in only one way as a sum of elements from the specified subspaces.

1.43 דוגמה: סכום ישר של מספר תתי־מרחב

נניח ש־ V_k הוא תת־מרחב של $\mathbf{F}^n$ המורכב מהווקטורים שכל הקואורדינטות שלהם 0, פרט אולי לרכיב ה־ k^{th}; לדוגמה, $V_2 = \{(0, x, 0, \dots, 0) \in \mathbf{F}^n : x \in \mathbf{F}\}$. אזי

$$\mathbf{F}^n = V_1 \oplus \cdots \oplus V_n,$$

כפי שעליכם לוודא.

לכתיבת $\oplus$ ב־TEX, הקלידו \oplus.

לעיתים אי־דוגמאות מסייעות להבנה באותה מידה שבה הדוגמאות מסייעות.

1.44 דוגמה: סכום שאינו סכום ישר

נניח ש־

$$V_1 = \{(x, y, 0) \in \mathbf{F}^3 : x, y \in \mathbf{F}\},$$

$$V_2 = \{(0, 0, z) \in \mathbf{F}^3 : z \in \mathbf{F}\},$$

$$V_3 = \{(0, y, y) \in \mathbf{F}^3 : y \in \mathbf{F}\}.$$

אזי $\mathbf{F}^3 = V_1 + V_2 + V_3$ מאחר שכל וקטור $(x, y, z) \in \mathbf{F}^3$ יכול להיכתב כ־

$$(x, y, z) = (x, y, 0) + (0, 0, z) + (0, 0, 0),$$

כאשר הווקטור הראשון באגף הימני של המשוואה הוא ב־ V_1, הווקטור השני הוא ב־ V_2, והווקטור השלישי הוא ב־ V_3.

אך עדיין, $\mathbf{F}^3$ אינו סכום ישר של V_1, V_2, V_3, מאחר שהווקטור $(0, 0, 0)$ יכול להיכתב ביותר מדרך אחת כסכום $v_1 + v_2 + v_3$, עבור כל $v_k \in V_k$. בפרט, מתקיים

$$(0, 0, 0) = (0, 1, 0) + (0, 0, 1) + (0, -1, -1)$$

וכמובן גם,

$$(0, 0, 0) = (0, 0, 0) + (0, 0, 0) + (0, 0, 0),$$

כאשר הווקטור הראשון באגף הימני של כל משוואה לעיל הוא ב־ V_1, הווקטור השני ב־ V_2, והווקטור השלישי ב־ V_3. לפיכך הסכום $V_1 + V_2 + V_3$ אינו סכום ישר.

הסמל $\oplus$, שהוא סימן החיבור בתוך מעגל, מזכיר לנו שאנו עוסקים במקרה מיוחד של סכום תתי־מרחב—לכל איבר בסכום ישר ייצוג יחיד כסכום איברים מתתי־המרחב המוגדרים.

הגדרת סכום ישר דורשת כי לכל וקטור בסכום יהיה ייצוג יחיד כסכום מהצורה הנדרשת. במסקנה הבאה נראה כי כאשר עלינו לקבוע האם סכום של תתי־מרחב הוא סכום ישר, כל שעלינו לבדוק הוא האם 0 יכול להיכתב באופן יחיד כסכום מהצורה הנדרשת.

1.45 *condition for a direct sum*

Suppose $V_1, \dots, V_m$ are subspaces of V. Then $V_1 + \cdots + V_m$ is a direct sum if and only if the only way to write 0 as a sum $v_1 + \cdots + v_m$, where each $v_k \in V_k$, is by taking each v_k equal to 0.

Proof First suppose $V_1 + \cdots + V_m$ is a direct sum. Then the definition of direct sum implies that the only way to write 0 as a sum $v_1 + \cdots + v_m$, where each $v_k \in V_k$, is by taking each v_k equal to 0.

Now suppose that the only way to write 0 as a sum $v_1 + \cdots + v_m$, where each $v_k \in V_k$, is by taking each v_k equal to 0. To show that $V_1 + \cdots + V_m$ is a direct sum, let $v \in V_1 + \cdots + V_m$. We can write

$$v = v_1 + \cdots + v_m$$

for some $v_1 \in V_1, \dots, v_m \in V_m$. To show that this representation is unique, suppose we also have

$$v = u_1 + \cdots + u_m,$$

where $u_1 \in V_1, \dots, u_m \in V_m$. Subtracting these two equations, we have

$$0 = (v_1 - u_1) + \cdots + (v_m - u_m).$$

Because $v_1 - u_1 \in V_1, \dots, v_m - u_m \in V_m$, the equation above implies that each $v_k - u_k$ equals 0. Thus $v_1 = u_1, \dots, v_m = u_m$, as desired. ∎

The next result gives a simple condition for testing whether a sum of two subspaces is a direct sum.

The symbol $\iff$ used below means "if and only if"; this symbol could also be read to mean "is equivalent to".

1.46 *direct sum of two subspaces*

Suppose U and W are subspaces of V. Then

$$U + W \text{ is a direct sum} \iff U \cap W = \{0\}.$$

Proof First suppose that $U + W$ is a direct sum. If $v \in U \cap W$, then $0 = v + (-v)$, where $v \in U$ and $-v \in W$. By the unique representation of 0 as the sum of a vector in U and a vector in W, we have $v = 0$. Thus $U \cap W = \{0\}$, completing the proof in one direction.

To prove the other direction, now suppose $U \cap W = \{0\}$. To prove that $U + W$ is a direct sum, suppose $u \in U$, $w \in W$, and

$$0 = u + w.$$

To complete the proof, we only need to show that $u = w = 0$ (by 1.45). The equation above implies that $u = -w \in W$. Thus $u \in U \cap W$. Hence $u = 0$, which by the equation above implies that $w = 0$, completing the proof. ∎

1.45 תנאי לסכום ישר

נניח ש־ $V_1, \dots, V_m$ תתי־מרחב של V. אזי $V_1 + \cdots + V_m$ הוא סכום ישר אם־ורק־אם הדרך היחידה לבטא את 0 כסכום $v_1 + \cdots + v_m$, עבור כל $v_k \in V_k$, היא שכל v_k שווה 0.

הוכחה תחילה נניח ש־ $V_1 + \cdots + V_m$ הוא סכום ישר. הגדרת סכום ישר מכתיבה שהאופן היחיד שבו ניתן לבטא את 0 כסכום $v_1 + \cdots + v_m$, עבור $v_k \in V_k$, הוא כשניקח את איבר האפס מכל תת־מרחב, כלומר כל $v_k = 0$.

כעת נניח שהדרך היחידה לבטא את 0 כסכום $v_1 + \cdots + v_m$, עבור $v_k \in V_k$, היא כשנבחר רק את איבר ה־ 0 עבור כל v_k. כדי להראות ש־ $V_1 + \cdots + V_m$ הוא סכום ישר, יהי $v \in V_1 + \cdots + V_m$. נוכל לכתוב

$$v = v_1 + \cdots + v_m$$

עבור $v_1 \in V_1, \dots, v_m \in V_m$ כלשהם. כדי להראות שזהו אופן ההצגה היחיד של הווקטור, נניח בנוסף שקיימת חלופה, כך ש־

$$v = u_1 + \cdots + u_m,$$

עבור $u_1 \in V_1, \dots, u_m \in V_m$ כלשהם. בחיסור שתי המשוואות הללו, נקבל

$$0 = (v_1 - u_1) + \cdots + (v_m - u_m).$$

מאחר ש־ $v_1 - u_1 \in V_1, \dots, v_m - u_m \in V_m$, מהמשוואה לעיל נסיק שכל $v_k - u_k$ שווה 0. לפיכך $v_1 = u_1, \dots, v_m = u_m$ כפי שרצינו. ■

המסקנה הבאה מספקת תנאי פשוט לבדיקה האם סכום של שני תתי־מרחב הוא סכום ישר.

הסימן $\iff$ שבשימוש להלן משמעותו "אם־ורק־אם"; אפשר לקרוא אותו גם כ־ "שקול ל...".

1.46 סכום ישר של שני תתי־מרחב

נניח ש־ U ו־ W תתי־מרחב של V. אזי

$$U + W \text{ סכום ישר } \iff U \cap W = \{0\}.$$

הוכחה תחילה נניח ש־ $U + W$ סכום ישר. אם $v \in U \cap W$, אזי $0 = v + (-v)$ כאשר $v \in U$ ו־ $-v \in W$. לפי כלל הייצוג הייחודי ל־ 0 כסכום של וקטור ב־ U ווקטור ב־ W, קיבלנו $v = 0$. לפיכך $U \cap W = \{0\}$, ובכך השלמנו כיוון אחד של ההוכחה.

להוכחת הכיוון השני, נניח כעת ש־ $U \cap W = \{0\}$. כדי להוכיח ש־ $U + W$ סכום ישר, נניח ש־ $u \in U$, $w \in W$, וגם

$$0 = u + w.$$

להשלמת ההוכחה, כל שנדרש הוא להראות ש־ $u = w = 0$ (לפי 1.45). מהמשוואה לעיל נסיק כי $u = -w \in W$. לפיכך $u \in U \cap W$. לכן $u = 0$, ולפי המשוואה לעיל זה גורר ש־ $w = 0$, ובכך הושלמה ההוכחה. ■

The result above deals only with the case of two subspaces. When asking about a possible direct sum with more than two subspaces, it is not enough to test that each pair of the subspaces intersect only at 0. To see this, consider Example 1.44. In that nonexample of a direct sum, we have $V_1 \cap V_2 = V_1 \cap V_3 = V_2 \cap V_3 = \{0\}$.

Sums of subspaces are analogous to unions of subsets. Similarly, direct sums of subspaces are analogous to disjoint unions of subsets. No two subspaces of a vector space can be disjoint, because both contain 0. So disjointness is replaced, at least in the case of two subspaces, with the requirement that the intersection equal $\{0\}$.

Exercises 1C

1 For each of the following subsets of $\mathbf{F}^3$, determine whether it is a subspace of $\mathbf{F}^3$.
(a) $\{(x_1, x_2, x_3) \in \mathbf{F}^3 : x_1 + 2x_2 + 3x_3 = 0\}$
(b) $\{(x_1, x_2, x_3) \in \mathbf{F}^3 : x_1 + 2x_2 + 3x_3 = 4\}$
(c) $\{(x_1, x_2, x_3) \in \mathbf{F}^3 : x_1 x_2 x_3 = 0\}$
(d) $\{(x_1, x_2, x_3) \in \mathbf{F}^3 : x_1 = 5x_3\}$

2 Verify all assertions about subspaces in Example 1.35.

3 Show that the set of differentiable real-valued functions f on the interval $(-4, 4)$ such that $f'(-1) = 3f(2)$ is a subspace of $\mathbf{R}^{(-4,4)}$.

4 Suppose $b \in \mathbf{R}$. Show that the set of continuous real-valued functions f on the interval $[0, 1]$ such that $\int_0^1 f = b$ is a subspace of $\mathbf{R}^{[0,1]}$ if and only if $b = 0$.

5 Is $\mathbf{R}^2$ a subspace of the complex vector space $\mathbf{C}^2$?

6 (a) Is $\{(a, b, c) \in \mathbf{R}^3 : a^3 = b^3\}$ a subspace of $\mathbf{R}^3$?
(b) Is $\{(a, b, c) \in \mathbf{C}^3 : a^3 = b^3\}$ a subspace of $\mathbf{C}^3$?

7 Prove or give a counterexample: If U is a nonempty subset of $\mathbf{R}^2$ such that U is closed under addition and under taking additive inverses (meaning $-u \in U$ whenever $u \in U$), then U is a subspace of $\mathbf{R}^2$.

8 Give an example of a nonempty subset U of $\mathbf{R}^2$ such that U is closed under scalar multiplication, but U is not a subspace of $\mathbf{R}^2$.

9 A function $f : \mathbf{R} \to \mathbf{R}$ is called *periodic* if there exists a positive number p such that $f(x) = f(x + p)$ for all $x \in \mathbf{R}$. Is the set of periodic functions from $\mathbf{R}$ to $\mathbf{R}$ a subspace of $\mathbf{R}^{\mathbf{R}}$? Explain.

10 Suppose V_1 and V_2 are subspaces of V. Prove that the intersection $V_1 \cap V_2$ is a subspace of V.

המסקנה לעיל עוסקת רק במקרה של שני תתי־מרחב. כשנבוא לדון באפשרות של סכום ישר עבור יותר משני תתי־מרחב, הבדיקה שבחיתוך כל צמד קיים רק איבר האפס אינה מספיקה. כדי להיווכח בכך נתבונן בדוגמה 1.44. באי־דוגמה זו לסכום ישר, ראינו ש־ $V_1 \cap V_2 = V_1 \cap V_3 = V_2 \cap V_3 = \{0\}$.

סכומי תתי־מרחב מקבילים לאיחודי תתי־קבוצה. באופן דומה, קיימת הקבלה בין סכום ישר של תתי־מרחב לבין איחוד זר של תתי־קבוצה. אמנם, לאור העובדה ששני תתי־מרחב לעולם אינם זרים, שהרי איבר האפס משותף לשניהם, נוכל לומר שאת זרות תתי־הקבוצה הפרנו, לפחות במקרה של שני תתי־מרחב, בדרישה שהחיתוך שווה ל־ $\{0\}$.

תרגילים אג

1 עבור כל אחת מתתי־הקבוצה של $\mathbf{F}^3$ להלן, קבעו האם היא גם תת־מרחב של $\mathbf{F}^3$.

א. $\{(x_1, x_2, x_3) \in \mathbf{F}^3 : x_1 + 2x_2 + 3x_3 = 0\}$

ב. $\{(x_1, x_2, x_3) \in \mathbf{F}^3 : x_1 + 2x_2 + 3x_3 = 4\}$

ג. $\{(x_1, x_2, x_3) \in \mathbf{F}^3 : x_1 x_2 x_3 = 0\}$

ד. $\{(x_1, x_2, x_3) \in \mathbf{F}^3 : x_1 = 5x_3\}$

2 ודאו את כל הטענות על תתי־מרחב בדוגמה 1.35.

3 הראו שקבוצת כל הפונקציות הגזירות הממשיות f מעל הקטע $(-4, 4)$ כך ש־ $f'(-1) = 3f(2)$ היא תת־מרחב של $\mathbf{R}^{(-4,4)}$.

4 נניח ש־ $b \in \mathbf{R}$. הראו שקבוצת הפונקציות הממשיות הרציפות f בקטע $[0, 1]$ כך ש־ $\int_0^1 f = b$ היא תת־מרחב של $\mathbf{R}^{[0,1]}$ אם־ורק־אם $b = 0$.

5 האם $\mathbf{R}^2$ תת־מרחב של המרחב הווקטורי המרוכב $\mathbf{C}^2$?

6 א. האם $\{(a, b, c) \in \mathbf{R}^3 : a^3 = b^3\}$ תת־מרחב של $\mathbf{R}^3$?

ב. האם $\{(a, b, c) \in \mathbf{C}^3 : a^3 = b^3\}$ תת־מרחב של $\mathbf{C}^3$?

7 הוכיחו או הביאו דוגמה נגדית: אם U היא תת־קבוצה לא ריקה של $\mathbf{R}^2$, סגורה תחת חיבור ואיבר נגדי לחיבור (כלומר $-u \in U$ עבור כל $u \in U$), אזי U תת־מרחב של $\mathbf{R}^2$.

8 הביאו דוגמה לתת־קבוצה לא ריקה U של $\mathbf{R}^2$, הסגורה תחת כפל בסקלר, ועדיין U אינה תת־מרחב של $\mathbf{R}^2$.

9 פונקציה $f: \mathbf{R} \to \mathbf{R}$ נקראת מחזורית אם קיים מספר חיובי p כך ש־ $f(x) = f(x+p)$ לכל $x \in \mathbf{R}$. האם קבוצת הפונקציות המחזוריות מ־ $\mathbf{R}$ ל־ $\mathbf{R}$ תת־מרחב של $\mathbf{R}^{\mathbf{R}}$? הסבירו.

10 נניח ש־ V_1 ו־ V_2 תתי־מרחב של V. הוכיחו כי החיתוך $V_1 \cap V_2$ הוא תת־מרחב של V.

11 Prove that the intersection of every collection of subspaces of V is a subspace of V.

12 Prove that the union of two subspaces of V is a subspace of V if and only if one of the subspaces is contained in the other.

13 Prove that the union of three subspaces of V is a subspace of V if and only if one of the subspaces contains the other two.

This exercise is surprisingly harder than Exercise 12, possibly because this exercise is not true if we replace $\mathbf{F}$ *with a field containing only two elements.*

14 Suppose

$$U = \{(x, -x, 2x) \in \mathbf{F}^3 : x \in \mathbf{F}\} \quad \text{and} \quad W = \{(x, x, 2x) \in \mathbf{F}^3 : x \in \mathbf{F}\}.$$

Describe $U + W$ using symbols, and also give a description of $U + W$ that uses no symbols.

15 Suppose U is a subspace of V. What is $U + U$?

16 Is the operation of addition on the subspaces of V commutative? In other words, if U and W are subspaces of V, is $U + W = W + U$?

17 Is the operation of addition on the subspaces of V associative? In other words, if V_1, V_2, V_3 are subspaces of V, is

$$(V_1 + V_2) + V_3 = V_1 + (V_2 + V_3)?$$

18 Does the operation of addition on the subspaces of V have an additive identity? Which subspaces have additive inverses?

19 Prove or give a counterexample: If V_1, V_2, U are subspaces of V such that

$$V_1 + U = V_2 + U,$$

then $V_1 = V_2$.

20 Suppose

$$U = \{(x, x, y, y) \in \mathbf{F}^4 : x, y \in \mathbf{F}\}.$$

Find a subspace W of $\mathbf{F}^4$ such that $\mathbf{F}^4 = U \oplus W$.

21 Suppose

$$U = \{(x, y, x + y, x - y, 2x) \in \mathbf{F}^5 : x, y \in \mathbf{F}\}.$$

Find a subspace W of $\mathbf{F}^5$ such that $\mathbf{F}^5 = U \oplus W$.

22 Suppose

$$U = \{(x, y, x + y, x - y, 2x) \in \mathbf{F}^5 : x, y \in \mathbf{F}\}.$$

Find three subspaces W_1, W_2, W_3 of $\mathbf{F}^5$, none of which equals $\{0\}$, such that $\mathbf{F}^5 = U \oplus W_1 \oplus W_2 \oplus W_3$.

11 הוכיחו כי החיתוך של כל אוסף תתי־מרחב של V הוא תת־מרחב של V.

12 הוכיחו כי איחוד שני תתי־מרחב של V הוא תת־מרחב של V אם־ורק־אם אחד מתתי־המרחב מוכל באחר.

13 הוכיחו כי איחוד של שלושה תתי־מרחב של V הוא תת־מרחב של V אם־ורק־אם תת־מרחב אחד מכיל את שני האחרים.

תרגיל זה קשה בפתיע יותר מתרגיל 12; ייתכן שהדבר נובע מכך שהטענה אינה נכונה אם מחליפים את $\mathbf{F}$ בשדה המכיל רק שני איברים.

14 נניח ש־

$$W = \{(x, x, 2x) \in \mathbf{F}^3 : x \in \mathbf{F}\} \quad \text{וגם} \quad U = \{(x, -x, 2x) \in \mathbf{F}^3 : x \in \mathbf{F}\}.$$

תארו את $U + W$ תוך שימוש בסימונים מתמטיים, ובנוסף, תנו תיאור מילולי ל־$U + W$ ללא שימוש בסימונים מתמטיים.

15 נניח ש־ U תת־מרחב של V. מהו $U + U$?

16 האם פעולת החיבור מעל תתי־המרחב של V היא קומוטטיבית? כלומר, אם U ו־ W תתי־מרחב של V, האם $U + W = W + U$?

17 האם פעולת החיבור מעל תתי־המרחב של V היא אסוציאטיבית? כלומר, אם V_1, V_2, V_3 תתי־מרחב של V, האם

$$(V_1 + V_2) + V_3 = V_1 + (V_2 + V_3)?$$

18 האם לפעולת החיבור מעל תתי־המרחב של V יש איבר זהות לחיבור? לאילו תתי־מרחב יש איבר נגדי לחיבור?

19 הוכיחו או הביאו דוגמה נגדית: אם V_1, V_2, U תתי־מרחב של V כך ש־

$$V_1 + U = V_2 + U,$$

אזי $V_1 = V_2$.

20 נניח ש־

$$U = \{(x, x, y, y) \in \mathbf{F}^4 : x, y \in \mathbf{F}\}.$$

מצאו תת־מרחב W של $\mathbf{F}^4$ כך ש־ $\mathbf{F}^4 = U \oplus W$.

21 נניח ש־

$$U = \{(x, y, x + y, x - y, 2x) \in \mathbf{F}^5 : x, y \in \mathbf{F}\}.$$

מצאו תת־מרחב W של $\mathbf{F}^5$ כך ש־ $\mathbf{F}^5 = U \oplus W$.

22 נניח ש־

$$U = \{(x, y, x + y, x - y, 2x) \in \mathbf{F}^5 : x, y \in \mathbf{F}\}.$$

מצאו שלושה תתי־מרחב W_1, W_2, W_3 של $\mathbf{F}^5$, שאף אחד מהם אינו שווה ל־ $\{0\}$, כך ש־ $\mathbf{F}^5 = U \oplus W_1 \oplus W_2 \oplus W_3$.

23 Prove or give a counterexample: If V_1, V_2, U are subspaces of V such that

$$V = V_1 \oplus U \quad \text{and} \quad V = V_2 \oplus U,$$

then $V_1 = V_2$.

Hint: When trying to discover whether a conjecture in linear algebra is true or false, it is often useful to start by experimenting in $\mathbf{F}^2$.

24 A function $f : \mathbf{R} \to \mathbf{R}$ is called *even* if

$$f(-x) = f(x)$$

for all $x \in \mathbf{R}$. A function $f : \mathbf{R} \to \mathbf{R}$ is called *odd* if

$$f(-x) = -f(x)$$

for all $x \in \mathbf{R}$. Let V_e denote the set of real-valued even functions on $\mathbf{R}$ and let V_o denote the set of real-valued odd functions on $\mathbf{R}$. Show that $\mathbf{R}^{\mathbf{R}} = V_e \oplus V_o$.

23 הוכיחו או הביאו דוגמה נגדית: אם V_1, V_2, U תתי־מרחב של V כך ש־

$$V = V_1 \oplus U \quad \text{וגם} \quad V = V_2 \oplus U,$$

אזי $V_1 = V_2$.

רמז: כאשר מנסים לגלות האם השערה מסוימת באלגברה ליניארית נכונה, לעיתים קרובות הבדיקה ב־$\mathbf{F}^2$ היא נקודת התחלה טובה.

24 פונקציה $f\colon \mathbf{R} \to \mathbf{R}$ נקראת זוגית אם:

$$f(-x) = f(x)$$

לכל $x \in \mathbf{R}$.

פונקציה $f\colon \mathbf{R} \to \mathbf{R}$ נקראת אי־זוגית אם:

$$f(-x) = -f(x)$$

לכל $x \in \mathbf{R}$. נסמן ב־V_e את קבוצת הפונקציות הממשיות הזוגיות מעל $\mathbf{R}$ וב־V_o את קבוצת הפונקציות הממשיות האי־זוגיות מעל $\mathbf{R}$. הראו ש־$\mathbf{R}^{\mathbf{R}} = V_e \oplus V_o$.

Chapter 2

Finite-Dimensional Vector Spaces

In the last chapter we learned about vector spaces. Linear algebra focuses not on arbitrary vector spaces, but on finite-dimensional vector spaces, which we introduce in this chapter.

We begin this chapter by considering linear combinations of lists of vectors. This leads us to the crucial concept of linear independence. The linear dependence lemma will become one of our most useful tools.

A list of vectors in a vector space that is small enough to be linearly independent and big enough so the linear combinations of the list fill up the vector space is called a basis of the vector space. We will see that every basis of a vector space has the same length, which will allow us to define the dimension of a vector space.

This chapter ends with a formula for the dimension of the sum of two subspaces.

standing assumptions for this chapter

- **F** denotes **R** or **C**.
- V denotes a vector space over **F**.

The main building of the Institute for Advanced Study, in Princeton, New Jersey. Paul Halmos (1916–2006) wrote the first modern linear algebra book in this building. Halmos's linear algebra book was published in 1942 (second edition published in 1958). The title of Halmos's book was the same as the title of this chapter.

פרק ב

מרחבים וקטוריים ממד סופי

בפרק הקודם למדנו על מרחבים וקטוריים. אלגברה ליניארית אינה עוסקת בכל מרחב וקטורי שרירותי, כי אם במרחבים בעלי ממד סופי, שאותם נכיר בפרק זה.

את הפרק נפתח בדיון על צירופים ליניאריים של רשימות וקטורים, מה שיוביל אותנו אל הרעיון המרכזי ששמו אי־תלות ליניארית. לָמַת התלות הליניארית תהפוך לאחד הכלים השימושיים ביותר שלנו.

רשימת וקטורים במרחב וקטורי שהיא קטנה דיה להיות בלתי תלויה ליניארית, אך עדיין גדולה דיה כך שצירופי איבריה ממלאים את כל המרחב הווקטורי, נקראת בסיס של המרחב. אנו ניווכח כי לכל הבסיסים במרחב אורך זהה, מה שיאפשר לנו להגדיר את ממד המרחב.

את הפרק נחתום בנוסחה לממד של סכום שני תתי־מרחב.

הנחות המוצא לפרק זה

- $\mathbf{F}$ מסמן $\mathbf{R}$ או $\mathbf{C}$.
- V מסמן מרחב וקטורי מעל $\mathbf{F}$.

הבניין המרכזי של המכון למחקר מתקדם, בפרינסטון, ניו־ג'רזי, ארה"ב. פול הלמוש (1916–2006) כתב את ספר האלגברה הליניארית המודרני הראשון בבניין זה. ספר זה של הלמוש יצא לאור בשנת 1942 (מהדורה שנייה הודפסה בשנת 1958). שם ספרו זה של הלמוש הוא כשמו של פרק זה, ובאנגלית — Finite-Dimensional Vector Spaces.

אלגברה ליניארית בדרך הישר, מהדורה רביעית, מחבר: שלדון אקסלר
Linear Algebra Done Right, fourth edition, by Sheldon Axler

2A Span and Linear Independence

We have been writing lists of numbers surrounded by parentheses, and we will continue to do so for elements of $\mathbf{F}^n$; for example, $(2, -7, 8) \in \mathbf{F}^3$. However, now we need to consider lists of vectors (which may be elements of $\mathbf{F}^n$ or of other vector spaces). To avoid confusion, we will usually write lists of vectors without surrounding parentheses. For example, $(4, 1, 6), (9, 5, 7)$ is a list of length two of vectors in $\mathbf{R}^3$.

2.1 notation: *list of vectors*

We will usually write lists of vectors without surrounding parentheses.

Linear Combinations and Span

A sum of scalar multiples of the vectors in a list is called a linear combination of the list. Here is the formal definition.

2.2 definition: *linear combination*

A *linear combination* of a list $v_1, \dots, v_m$ of vectors in V is a vector of the form

$$a_1 v_1 + \cdots + a_m v_m,$$

where $a_1, \dots, a_m \in \mathbf{F}$.

2.3 example: *linear combinations in* $\mathbf{R}^3$

- $(17, -4, 2)$ is a linear combination of $(2, 1, -3), (1, -2, 4)$, which is a list of length two of vectors in $\mathbf{R}^3$, because

$$(17, -4, 2) = 6(2, 1, -3) + 5(1, -2, 4).$$

- $(17, -4, 5)$ is not a linear combination of $(2, 1, -3), (1, -2, 4)$, which is a list of length two of vectors in $\mathbf{R}^3$, because there do not exist numbers $a_1, a_2 \in \mathbf{R}$ such that

$$(17, -4, 5) = a_1(2, 1, -3) + a_2(1, -2, 4).$$

  In other words, the system of equations

$$\begin{aligned} 17 &= 2a_1 + a_2 \\ -4 &= a_1 - 2a_2 \\ 5 &= -3a_1 + 4a_2 \end{aligned}$$

  has no solutions (as you should verify).

בו פְּרִיסָה ואי־תלות ליניארית

עד כה נהגנו להקיף בסוגריים רשימות מספרים, וכך נמשיך לעשות עבור איברים ב־ $\mathbf{F}^n$; לדוגמה, $(2,-7,8) \in \mathbf{F}^3$. אולם, כעת עלינו להביא בחשבון רשימות של וקטורים (שייתכן שיהיו איברים ב־ $\mathbf{F}^n$ או במרחבים וקטוריים אחרים). כדי להימנע מבלבול, ככלל, רשימות הווקטורים תיכתבנה ללא סוגריים. לדוגמה, $(4,1,6), (9,5,7)$ היא רשימה בת שני וקטורים ב־ $\mathbf{R}^3$.

2.1 סימון: רשימת וקטורים

בדרך כלל, נכתוב את רשימות הווקטורים ללא סוגריים.

צירופים ליניאריים ופריסה

סכום של כפולות סקלר של הווקטורים ברשימה נתונה נקרא צירוף ליניארי של הרשימה. להלן ההגדרה הרשמית.

2.2 הגדרה: צירוף ליניארי

צירוף ליניארי של רשימת הווקטורים $v_1, \dots, v_m$ ב V הוא וקטור מהצורה

$$a_1 v_1 + \cdots + a_m v_m,$$

כאשר $a_1, \dots, a_m \in \mathbf{F}$.

2.3 דוגמה: צירופים ליניאריים ב־ $\mathbf{R}^3$

- $(17,-4,2)$ הוא צירוף ליניארי של $(2,1,-3), (1,-2,4)$, שהיא רשימה בת שני וקטורים ב־ $\mathbf{R}^3$, שכן
$$(17,-4,2) = 6(2,1,-3) + 5(1,-2,4).$$
- $(17,-4,5)$ אינו צירוף ליניארי של $(2,1,-3), (1,-2,4)$, שהיא רשימה בת שני וקטורים ב־ $\mathbf{R}^3$, שכן לא קיימים מספרים $a_1, a_2 \in \mathbf{R}$ כך ש־
$$(17,-4,5) = a_1(2,1,-3) + a_2(1,-2,4).$$
כלומר, למערכת המשוואות
$$\begin{aligned} 17 &= 2a_1 + a_2 \\ -4 &= a_1 - 2a_2 \\ 5 &= -3a_1 + 4a_2 \end{aligned}$$
אין פתרון (כפי שעליכם לוודא).

2.4 definition: *span*

The set of all linear combinations of a list of vectors $v_1, \dots, v_m$ in V is called the *span* of $v_1, \dots, v_m$, denoted by $\operatorname{span}(v_1, \dots, v_m)$. In other words,

$$\operatorname{span}(v_1, \dots, v_m) = \{a_1 v_1 + \cdots + a_m v_m : a_1, \dots, a_m \in \mathbf{F}\}.$$

The span of the empty list ( ) is defined to be $\{0\}$.

2.5 example: *span*

The previous example shows that in $\mathbf{R}^3$,

- $(17, -4, 2) \in \operatorname{span}((2, 1, -3), (1, -2, 4))$;
- $(17, -4, 5) \notin \operatorname{span}((2, 1, -3), (1, -2, 4))$.

2.6 *span is the smallest containing subspace*

The span of a list of vectors in V is the smallest subspace of V containing all vectors in the list.

Proof Suppose $v_1, \dots, v_m$ is a list of vectors in V.

First we show that $\operatorname{span}(v_1, \dots, v_m)$ is a subspace of V. The additive identity is in $\operatorname{span}(v_1, \dots, v_m)$ because

$$0 = 0v_1 + \cdots + 0v_m.$$

*Some mathematicians use the terminology **linear span**, which means the same as span.*

Also, $\operatorname{span}(v_1, \dots, v_m)$ is closed under addition because

$$(a_1 v_1 + \cdots + a_m v_m) + (c_1 v_1 + \cdots + c_m v_m) = (a_1 + c_1)v_1 + \cdots + (a_m + c_m)v_m.$$

Furthermore, $\operatorname{span}(v_1, \dots, v_m)$ is closed under scalar multiplication because

$$\lambda(a_1 v_1 + \cdots + a_m v_m) = \lambda a_1 v_1 + \cdots + \lambda a_m v_m.$$

Thus $\operatorname{span}(v_1, \dots, v_m)$ is a subspace of V (by 1.34).

Each v_k is a linear combination of $v_1, \dots, v_m$ (to show this, set $a_k = 1$ and let the other a's in 2.2 equal 0). Thus $\operatorname{span}(v_1, \dots, v_m)$ contains each v_k. Conversely, because subspaces are closed under scalar multiplication and addition, every subspace of V that contains each v_k contains $\operatorname{span}(v_1, \dots, v_m)$. Thus $\operatorname{span}(v_1, \dots, v_m)$ is the smallest subspace of V containing all the vectors $v_1, \dots, v_m$. ∎

2.7 definition: *spans*

If $\operatorname{span}(v_1, \dots, v_m)$ equals V, we say that the list $v_1, \dots, v_m$ *spans* V.

2.4 הגדרה: פריסה, span

קבוצת כל הצירופים הליניאריים של רשימת הווקטורים $v_1, \dots, v_m$ ב־ V נקראת הפריסה של $v_1, \dots, v_m$ ומסומנת על ידי $\operatorname{span}(v_1, \dots, v_m)$. במילים אחרות,

$$\operatorname{span}(v_1, \dots, v_m) = \{a_1 v_1 + \cdots + a_m v_m : a_1, \dots, a_m \in \mathbf{F}\}.$$

הפריסה של רשימת וקטורים ריקה ( ) מוגדרת כ־ $\{0\}$.

2.5 דוגמה: span

בדוגמה הקודמת ראינו שב־ $\mathbf{R}^3$,

- $(17, -4, 2) \in \operatorname{span}((2, 1, -3), (1, -2, 4))$
- $(17, -4, 5) \notin \operatorname{span}((2, 1, -3), (1, -2, 4))$.

2.6 span הוא תת־המרחב הקטן ביותר המכיל

ה־ span של רשימת וקטורים ב־ V הוא התת־מרחב הקטן ביותר של V המכיל את כל הווקטורים שברשימה.

יש מתמטיקאים המשתמשים במונח **פריסה ליניארית**, שמשמעותו זהה לפריסה.

הוכחה נניח ש־ $v_1, \dots, v_m$ היא רשימת וקטורים ב־ V.

ראשית נראה ש־ $\operatorname{span}(v_1, \dots, v_m)$ הוא תת־מרחב של V. איבר הזהות לחיבור קיים ב־ $\operatorname{span}(v_1, \dots, v_m)$ שכן

$$0 = 0v_1 + \cdots + 0v_m.$$

בנוסף, $\operatorname{span}(v_1, \dots, v_m)$ סגור תחת פעולת החיבור, מאחר ש־

$$(a_1 v_1 + \cdots + a_m v_m) + (c_1 v_1 + \cdots + c_m v_m) = (a_1 + c_1)v_1 + \cdots + (a_m + c_m)v_m.$$

יתר על כן, $\operatorname{span}(v_1, \dots, v_m)$ סגור תחת כפל בסקלר, שהרי

$$\lambda(a_1 v_1 + \cdots + a_m v_m) = \lambda a_1 v_1 + \cdots + \lambda a_m v_m.$$

לפיכך, $\operatorname{span}(v_1, \dots, v_m)$ הוא תת־מרחב של V (לפי 1.34).

כל v_k הוא צירוף ליניארי של $v_1, \dots, v_m$ (להראות זאת, נגדיר $a_k = 1$ ובשאר ה־ a־ים ב־ 2.2 נציב 0). לפיכך, $\operatorname{span}(v_1, \dots, v_m)$ מכיל כל אחד מה־ v_k־ים. בכיוון ההפוך, מאחר שתתי־מרחב סגורים תחת כפל בסקלר וחיבור, כל תת־מרחב של V המכיל כל אחד מהווקטורים v_k מכיל את $\operatorname{span}(v_1, \dots, v_m)$. לפיכך $\operatorname{span}(v_1, \dots, v_m)$ הוא התת־מרחב הקטן ביותר של V המכיל את כל הווקטורים $v_1, \dots, v_m$. ■

2.7 הגדרה: פריסה

אם $\operatorname{span}(v_1, \dots, v_m)$ שווה ל־ V, נאמר כי הרשימה $v_1, \dots, v_m$ פורסת את V.

2.8 example: *a list that spans* $\mathbf{F}^n$

Suppose n is a positive integer. We want to show that

$$(1, 0, \dots, 0), (0, 1, 0, \dots, 0), \dots, (0, \dots, 0, 1)$$

spans $\mathbf{F}^n$. Here the kth vector in the list above has 1 in the kth slot and 0 in all other slots.

Suppose $(x_1, \dots, x_n) \in \mathbf{F}^n$. Then

$$(x_1, \dots, x_n) = x_1(1, 0, \dots, 0) + x_2(0, 1, 0, \dots, 0) + \cdots + x_n(0, \dots, 0, 1).$$

Thus $(x_1, \dots, x_n) \in \operatorname{span}((1, 0, \dots, 0), (0, 1, 0, \dots, 0), \dots, (0, \dots, 0, 1))$, as desired.

Now we can make one of the key definitions in linear algebra.

2.9 definition: *finite-dimensional vector space*

A vector space is called *finite-dimensional* if some list of vectors in it spans the space.

Example 2.8 above shows that $\mathbf{F}^n$ is a finite-dimensional vector space for every positive integer n.

Recall that by definition every list has finite length.

The definition of a polynomial is no doubt already familiar to you.

2.10 definition: *polynomial,* $\mathcal{P}(\mathbf{F})$

- A function $p\colon \mathbf{F} \to \mathbf{F}$ is called a *polynomial* with coefficients in $\mathbf{F}$ if there exist $a_0, \dots, a_m \in \mathbf{F}$ such that
$$p(z) = a_0 + a_1 z + a_2 z^2 + \cdots + a_m z^m$$
for all $z \in \mathbf{F}$.
- $\mathcal{P}(\mathbf{F})$ is the set of all polynomials with coefficients in $\mathbf{F}$.

With the usual operations of addition and scalar multiplication, $\mathcal{P}(\mathbf{F})$ is a vector space over $\mathbf{F}$, as you should verify. Hence $\mathcal{P}(\mathbf{F})$ is a subspace of $\mathbf{F}^{\mathbf{F}}$, the vector space of functions from $\mathbf{F}$ to $\mathbf{F}$.

If a polynomial (thought of as a function from $\mathbf{F}$ to $\mathbf{F}$) is represented by two sets of coefficients, then subtracting one representation of the polynomial from the other produces a polynomial that is identically zero as a function on $\mathbf{F}$ and hence has all zero coefficients (if you are unfamiliar with this fact, just believe it for now; we will prove it later—see 4.8). **Conclusion:** the coefficients of a polynomial are uniquely determined by the polynomial. Thus the next definition uniquely defines the degree of a polynomial.

2.8 דוגמה: רשימה הפורסת את $\mathbf{F}^n$

נניח ש־ n מספר שלם חיובי. אנו רוצים להראות כי

$$(1,0,\dots,0),(0,1,0,\dots,0),\dots,(0,\dots,0,1)$$

פורסת את $\mathbf{F}^n$. הווקטור ה־ k^{th} ברשימה לעיל מכיל 1 בקואורדינטה ה־ k^{th}, ו־ 0 בכל שאר הרכיבים.

נניח ש־ $(x_1,\dots,x_n) \in \mathbf{F}^n$. אזי

$$(x_1,\dots,x_n) = x_1(1,0,\dots,0) + x_2(0,1,0,\dots,0) + \cdots + x_n(0,\dots,0,1).$$

לכן $(x_1,\dots,x_n) \in \operatorname{span}((1,0,\dots,0),(0,1,0,\dots,0),\dots,(0,\dots,0,1))$, כפי שרצינו.

כעת אנו פונים לאחת מהגדרות היסוד באלגברה ליניארית.

2.9 הגדרה: מרחב וקטורי מממד סופי

מרחב וקטורי נקרא מממד סופי אם רשימת וקטורים כלשהי בתוכו פורסת את המרחב.

דוגמה 2.8 לעיל הראתה ש־ $\mathbf{F}^n$ הוא מרחב מממד סופי לכל n מספר שלם חיובי.

נזכיר כי לפי ההגדרה, כל רשימה היא באורך סופי.

אין ספק שהגדרת הפולינום כבר מוכרת לכם.

2.10 הגדרה: פולינום, $\mathcal{P}(\mathbf{F})$

- פונקציה $p : \mathbf{F} \to \mathbf{F}$ נקראת פולינום עם מקדמים ב־ $\mathbf{F}$ אם קיימים $a_0,\dots,a_m \in \mathbf{F}$ כך ש־
$$p(z) = a_0 + a_1 z + a_2 z^2 + \cdots + a_m z^m$$
לכל $z \in \mathbf{F}$.
- $\mathcal{P}(\mathbf{F})$ היא קבוצת כל הפולינומים עם מקדמים ב־ $\mathbf{F}$.

יחד עם פעולות החיבור והכפל בסקלר הרגילות, $\mathcal{P}(\mathbf{F})$ הוא מרחב וקטורי מעל $\mathbf{F}$, כפי שעליכם לוודא. לכן, $\mathcal{P}(\mathbf{F})$ הוא תת־מרחב של $\mathbf{F}^{\mathbf{F}}$, המרחב הווקטורי של פונקציות מ־ $\mathbf{F}$ ל־ $\mathbf{F}$.

אם פולינום (כשאנו חושבים עליו כפונקציה מ־ $\mathbf{F}$ ל־ $\mathbf{F}$) מיוצג על ידי שתי קבוצות מקדמים, אזי חיסור ייצוג אחד של הפולינום מהאחר יניב פולינום המתאפס זהותית כפונקציה מעל $\mathbf{F}$ ולכן כל המקדמים בפולינום זה הם אפסים (אם עובדה זו עדיין אינה ברורה, קבלו אותה לעת עתה; אנו נוכיח אותה בהמשך—ראו 4.8). **מסקנה:** מקדמי הפולינום נקבעים באופן יחיד על ידי הפולינום. לפיכך ההגדרה הבאה קובעת באופן יחיד את מעלת הפולינום.

2.11 definition: *degree of a polynomial,* $\deg p$

- A polynomial $p \in \mathcal{P}(\mathbf{F})$ is said to have *degree* m if there exist scalars $a_0, a_1, \dots, a_m \in \mathbf{F}$ with $a_m \neq 0$ such that for every $z \in \mathbf{F}$, we have
$$p(z) = a_0 + a_1 z + \cdots + a_m z^m.$$
- The polynomial that is identically 0 is said to have degree $-\infty$.
- The degree of a polynomial p is denoted by $\deg p$.

In the next definition, we use the convention that $-\infty < m$, which means that the polynomial 0 is in $\mathcal{P}_m(\mathbf{F})$.

2.12 notation: $\mathcal{P}_m(\mathbf{F})$

For m a nonnegative integer, $\mathcal{P}_m(\mathbf{F})$ denotes the set of all polynomials with coefficients in $\mathbf{F}$ and degree at most m.

If m is a nonnegative integer, then $\mathcal{P}_m(\mathbf{F}) = \operatorname{span}(1, z, \dots, z^m)$ [here we slightly abuse notation by letting z^k denote a function]. Thus $\mathcal{P}_m(\mathbf{F})$ is a finite-dimensional vector space for each nonnegative integer m.

2.13 definition: *infinite-dimensional vector space*

A vector space is called *infinite-dimensional* if it is not finite-dimensional.

2.14 example: $\mathcal{P}(\mathbf{F})$ *is infinite-dimensional*

Consider any list of elements of $\mathcal{P}(\mathbf{F})$. Let m denote the highest degree of the polynomials in this list. Then every polynomial in the span of this list has degree at most m. Thus z^{m+1} is not in the span of our list. Hence no list spans $\mathcal{P}(\mathbf{F})$. Thus $\mathcal{P}(\mathbf{F})$ is infinite-dimensional.

Linear Independence

Suppose $v_1, \dots, v_m \in V$ and $v \in \operatorname{span}(v_1, \dots, v_m)$. By the definition of span, there exist $a_1, \dots, a_m \in \mathbf{F}$ such that
$$v = a_1 v_1 + \cdots + a_m v_m.$$
Consider the question of whether the choice of scalars in the equation above is unique. Suppose $c_1, \dots, c_m$ is another set of scalars such that
$$v = c_1 v_1 + \cdots + c_m v_m.$$
Subtracting the last two equations, we have
$$0 = (a_1 - c_1)v_1 + \cdots + (a_m - c_m)v_m.$$

2.11 הגדרה: מעלת הפולינום, $\deg p$

- על פולינום $p \in \mathcal{P}(\mathbf{F})$ נאמר שמעלתו m אם קיימים סקלרים $a_0, a_1, \dots, a_m \in \mathbf{F}$ כאשר $a_m \neq 0$ כך שלכל $z \in \mathbf{F}$, מתקיים
$$p(z) = a_0 + a_1 z + \cdots + a_m z^m.$$
- על הפולינום שהוא זהותית 0 נאמר שמעלתו $-\infty$.
- את מעלת הפולינום p נסמן ב־ $\deg p$.

בהגדרה הבאה, אנו משתמשים במוסכמה ש־ $-\infty < m$, שמשמעותה היא שפולינום האפס הוא ב־ $\mathcal{P}_m(\mathbf{F})$.

2.12 סימון: $\mathcal{P}_m(\mathbf{F})$

עבור m מספר שלם אי־שלילי, $\mathcal{P}_m(\mathbf{F})$ מסמן את קבוצת כל הפולינומים עם מקדמים ב־ $\mathbf{F}$ שמעלתם m לכל היותר.

אם m מספר שלם אי־שלילי, אזי $\mathcal{P}_m(\mathbf{F}) = \operatorname{span}(1, z, \dots, z^m)$ [כאן אנו מרשים לעצמנו אי־דיוק קל בסימון, בכך שאנו משתמשים ב־ z^k לסימון פונקציה]. לפיכך, $\mathcal{P}_m(\mathbf{F})$ הוא מרחב וקטורי מממד סופי לכל מספר שלם אי־שלילי m.

2.13 הגדרה: מרחב וקטורי מממד אינסופי

מרחב וקטורי נקרא מממד אינסופי אם אינו מממד סופי.

2.14 דוגמה: $\mathcal{P}(\mathbf{F})$ הוא מממד אינסופי

נתבונן ברשימה כלשהי של איברים של $\mathcal{P}(\mathbf{F})$. נסמן ב־ m את המעלה הגבוהה ביותר של הפולינומים ברשימה. אזי, מעלת כל פולינום בפריסת רשימה זו היא לכל היותר m. לפיכך, z^{m+1} אינו בפריסה של הרשימה, ולכן אין רשימה הפורסת את $\mathcal{P}(\mathbf{F})$. לפיכך $\mathcal{P}(\mathbf{F})$ הוא מממד אינסופי.

אי־תלות ליניארית

נניח ש־ $v_1, \dots, v_m \in V$ ו־ $v \in \operatorname{span}(v_1, \dots, v_m)$. לפי הגדרת הפריסה, קיימים $a_1, \dots, a_m \in \mathbf{F}$ כך ש־
$$v = a_1 v_1 + \cdots + a_m v_m.$$
השאלה שנעסוק בה היא האם בחירת הסקלרים במשוואה לעיל יחידה. נניח ש־ $c_1, \dots, c_m$ היא קבוצה אחרת של סקלרים כך ש־
$$v = c_1 v_1 + \cdots + c_m v_m.$$
בחיסור שתי המשוואות האחרונות, נקבל
$$0 = (a_1 - c_1)v_1 + \cdots + (a_m - c_m)v_m.$$

Thus we have written 0 as a linear combination of $(v_1, \dots, v_m)$. If the only way to do this is by using 0 for all the scalars in the linear combination, then each $a_k - c_k$ equals 0, which means that each a_k equals c_k (and thus the choice of scalars was indeed unique). This situation is so important that we give it a special name—linear independence—which we now define.

2.15 definition: *linearly independent*

- A list $v_1, \dots, v_m$ of vectors in V is called *linearly independent* if the only choice of $a_1, \dots, a_m \in \mathbf{F}$ that makes

$$a_1 v_1 + \cdots + a_m v_m = 0$$

  is $a_1 = \cdots = a_m = 0$.
- The empty list $()$ is also declared to be linearly independent.

The reasoning above shows that $v_1, \dots, v_m$ is linearly independent if and only if each vector in $\operatorname{span}(v_1, \dots, v_m)$ has only one representation as a linear combination of $v_1, \dots, v_m$.

2.16 example: *linearly independent lists*

(a) To see that the list $(1, 0, 0, 0), (0, 1, 0, 0), (0, 0, 1, 0)$ is linearly independent in $\mathbf{F}^4$, suppose $a_1, a_2, a_3 \in \mathbf{F}$ and

$$a_1(1, 0, 0, 0) + a_2(0, 1, 0, 0) + a_3(0, 0, 1, 0) = (0, 0, 0, 0).$$

Thus

$$(a_1, a_2, a_3, 0) = (0, 0, 0, 0).$$

Hence $a_1 = a_2 = a_3 = 0$. Thus the list $(1, 0, 0, 0), (0, 1, 0, 0), (0, 0, 1, 0)$ is linearly independent in $\mathbf{F}^4$.

(b) Suppose m is a nonnegative integer. To see that the list $1, z, \dots, z^m$ is linearly independent in $\mathcal{P}(\mathbf{F})$, suppose $a_0, a_1, \dots, a_m \in \mathbf{F}$ and

$$a_0 + a_1 z + \cdots + a_m z^m = 0,$$

where we think of both sides as elements of $\mathcal{P}(\mathbf{F})$. Then

$$a_0 + a_1 z + \cdots + a_m z^m = 0$$

for all $z \in \mathbf{F}$. As discussed earlier (and as follows from 4.8), this implies that $a_0 = a_1 = \cdots = a_m = 0$. Thus $1, z, \dots, z^m$ is a linearly independent list in $\mathcal{P}(\mathbf{F})$.

(c) A list of length one in a vector space is linearly independent if and only if the vector in the list is not 0.

(d) A list of length two in a vector space is linearly independent if and only if neither of the two vectors in the list is a scalar multiple of the other.

לפיכך, כתבנו 0 כצירוף ליניארי של $(v_1, \dots, v_m)$. אם הדרך היחידה לעשות זאת היא על ידי כך שנבחר אפסים עבור כל הסקלרים בצירוף הליניארי, אזי כל $a_k - c_k$ שווה 0, שמשמעותו כי כל a_k שווה ל־ c_k (ולפיכך בחירת הסקלרים היתה אכן יחידה). למצב זה חשיבות כה רבה, עד כדי הענקת שם ייחודי—אי־תלות ליניארית—שכעת ניגש להגדרתה.

2.15 הגדרה: אי־תלות ליניארית

- רשימת וקטורים $v_1, \dots, v_m$ ב־ V נקראת בלתי תלויה ליניארית אם הדרך היחידה ש־ $a_1, \dots, a_m \in \mathbf{F}$ יקיימו את
$$a_1 v_1 + \cdots + a_m v_m = 0$$
היא כאשר $a_1 = \cdots = a_m = 0$.
- הרשימה הריקה ( ) מוגדרת גם היא כבלתי תלויה ליניארית.

ההנמקה לעיל מראה ש־ $v_1, \dots, v_m$ בלתי תלויה ליניארית אם־ורק־אם לכל וקטור ב־ $\operatorname{span}(v_1, \dots, v_m)$ יש ייצוג אחד בלבד כצירוף ליניארי של $v_1, \dots, v_m$.

2.16 דוגמה: רשימות בלתי תלויות ליניארית

א. כדי להיווכח שהרשימה $(1,0,0,0), (0,1,0,0), (0,0,1,0)$ בלתי תלויה ליניארית ב־ $\mathbf{F}^4$, נניח ש־ $a_1, a_2, a_3 \in \mathbf{F}$ ו־
$$a_1(1,0,0,0) + a_2(0,1,0,0) + a_3(0,0,1,0) = (0,0,0,0).$$
לפיכך
$$(a_1, a_2, a_3, 0) = (0,0,0,0).$$
לכן $a_1 = a_2 = a_3 = 0$. לפיכך הרשימה $(1,0,0,0), (0,1,0,0), (0,0,1,0)$ בלתי תלויה ליניארית ב־ $\mathbf{F}^4$.

ב. נניח ש־ m מספר שלם אי־שלילי. כדי להיווכח שהרשימה $1, z, \dots, z^m$ בלתי תלויה ליניארית ב־ $\mathcal{P}(\mathbf{F})$, נניח ש־ $a_0, a_1, \dots, a_m \in \mathbf{F}$ וגם
$$a_0 + a_1 z + \cdots + a_m z^m = 0,$$
כאשר אנו חושבים על שני האגפים כאיברים ב־ $\mathcal{P}(\mathbf{F})$. אזי
$$a_0 + a_1 z + \cdots + a_m z^m = 0$$
לכל $z \in \mathbf{F}$. לפי הנאמר לעיל (וכפי שנובע מ־ 4.8), זה גורר ש־ $a_0 = a_1 = \cdots = a_m = 0$. לפיכך $1, z, \dots, z^m$ רשימה בלתי תלויה ליניארית ב־ $\mathcal{P}(\mathbf{F})$.

ג. רשימה בת איבר יחיד במרחב וקטורי היא בלתי תלויה ליניארית אם־ורק־אם הווקטור ברשימה אינו 0.

ד. רשימה בת שני איברים במרחב וקטורי היא בלתי תלויה ליניארית אם־ורק־אם אף אחד מהווקטורים אינו כפולת סקלר של האחר.

If some vectors are removed from a linearly independent list, the remaining list is also linearly independent, as you should verify.

2.17 definition: *linearly dependent*

- A list of vectors in V is called *linearly dependent* if it is not linearly independent.
- In other words, a list $v_1, \dots, v_m$ of vectors in V is linearly dependent if there exist $a_1, \dots, a_m \in \mathbf{F}$, not all 0, such that $a_1 v_1 + \cdots + a_m v_m = 0$.

2.18 example: *linearly dependent lists*

- $(2,3,1), (1,-1,2), (7,3,8)$ is linearly dependent in $\mathbf{F}^3$ because
$$2(2,3,1) + 3(1,-1,2) + (-1)(7,3,8) = (0,0,0).$$
- The list $(2,3,1), (1,-1,2), (7,3,c)$ is linearly dependent in $\mathbf{F}^3$ if and only if $c = 8$, as you should verify.
- If some vector in a list of vectors in V is a linear combination of the other vectors, then the list is linearly dependent. (Proof: After writing one vector in the list as equal to a linear combination of the other vectors, move that vector to the other side of the equation, where it will be multiplied by -1.)
- Every list of vectors in V containing the 0 vector is linearly dependent. (This is a special case of the previous bullet point.)

The next lemma is a terrific tool. It states that given a linearly dependent list of vectors, one of the vectors is in the span of the previous ones. Furthermore, we can throw out that vector without changing the span of the original list.

2.19 *linear dependence lemma*

Suppose $v_1, \dots, v_m$ is a linearly dependent list in V. Then there exists $k \in \{1, 2, \dots, m\}$ such that
$$v_k \in \operatorname{span}(v_1, \dots, v_{k-1}).$$
Furthermore, if k satisfies the condition above and the kth term is removed from $v_1, \dots, v_m$, then the span of the remaining list equals $\operatorname{span}(v_1, \dots, v_m)$.

Proof Because the list $v_1, \dots, v_m$ is linearly dependent, there exist numbers $a_1, \dots, a_m \in \mathbf{F}$, not all 0, such that
$$a_1 v_1 + \cdots + a_m v_m = 0.$$
Let k be the largest element of $\{1, \dots, m\}$ such that $a_k \neq 0$. Then
$$v_k = -\frac{a_1}{a_k} v_1 - \cdots - \frac{a_{k-1}}{a_k} v_{k-1},$$
which proves that $v_k \in \operatorname{span}(v_1, \dots, v_{k-1})$, as desired.

אם הוסרו מספר וקטורים מרשימה בלתי תלויה ליניארית, הרשימה עם הווקטורים הנותרים אף היא בלתי תלויה ליניארית, כפי שעליכם לוודא.

2.17 הגדרה: תלות ליניארית

- רשימת וקטורים ב־ V נקראת תלויה ליניארית אם היא אינה בלתי תלויה ליניארית.
- במילים אחרות, רשימת וקטורים $v_1, \dots, v_m$ ב־ V תלויה ליניארית אם קיימים $a_1, \dots, a_m \in \mathbf{F}$, לא כולם אפסים, כך ש־ $a_1 v_1 + \cdots + a_m v_m = 0$.

2.18 דוגמה: רשימות תלויות ליניארית

- $(2,3,1), (1,-1,2), (7,3,8)$ תלויה ליניארית ב־ $\mathbf{F}^3$ מאחר ש־
$$2(2,3,1) + 3(1,-1,2) + (-1)(7,3,8) = (0,0,0).$$
- הרשימה $(2,3,1), (1,-1,2), (7,3,c)$ תלויה ליניארית ב־ $\mathbf{F}^3$ אם־ורק־אם $c = 8$, כפי שעליכם לוודא.
- אם וקטור כלשהו ברשימת וקטורים ב־ V הוא צירוף ליניארי של הווקטורים האחרים, הרשימה תלויה ליניארית. (הוכחה: אחרי כתיבת וקטור אחד ברשימה כצירוף ליניארי של הווקטורים האחרים, העבירו את הווקטור אל האגף השני של המשוואה, תוך הכפלתו בסקלר -1.)
- כל רשימת וקטורים ב־ V המכילה את וקטור האפס היא תלויה ליניארית. (זהו מקרה פרטי של הסעיף הקודם.)

הלֶמָה שלהלן היא כלי מצוין. היא קובעת כי בהינתן רשימת וקטורים תלויה ליניארית, אחד הווקטורים הוא בפריסה (span) של קודמיו. יתר על כן, נוכל להסיר וקטור זה מבלי לשנות את הפריסה של הרשימה המקורית.

2.19 לֶמַת התלות הליניארית

נניח ש־ $v_1, \dots, v_m$ היא רשימה תלויה ליניארית ב־ V. אזי קיים $k \in \{1, 2, \dots, m\}$ כך ש־
$$v_k \in \operatorname{span}(v_1, \dots, v_{k-1}).$$
יתר על כן, אם k מקיים את התנאי לעיל, אזי אחרי הסרת האיבר ה־ k^{th} מ־ $v_1, \dots, v_m$, פריסת הרשימה שנותרה שווה ל־ $\operatorname{span}(v_1, \dots, v_m)$.

הוכחה מאחר שהרשימה $v_1, \dots, v_m$ תלויה ליניארית, קיימים סקלרים $a_1, \dots, a_m \in \mathbf{F}$, לא כולם אפסים, כך ש־
$$a_1 v_1 + \cdots + a_m v_m = 0.$$
יהי k האינדקס הגדול ביותר ב־ $\{1, \dots, m\}$ כך ש־ $a_k \neq 0$. אזי
$$v_k = -\frac{a_1}{a_k} v_1 - \cdots - \frac{a_{k-1}}{a_k} v_{k-1},$$
מה שמוכיח ש־ $v_k \in \operatorname{span}(v_1, \dots, v_{k-1})$, כפי שרצינו.

Now suppose k is any element of $\{1, \dots, m\}$ such that $v_k \in \operatorname{span}(v_1, \dots, v_{k-1})$. Let $b_1, \dots, b_{k-1} \in \mathbf{F}$ such that

$$v_k = b_1 v_1 + \cdots + b_{k-1} v_{k-1}. \tag{2.20}$$

Suppose $u \in \operatorname{span}(v_1, \dots, v_m)$. Then there exist $c_1, \dots, c_m \in \mathbf{F}$ such that

$$u = c_1 v_1 + \cdots + c_m v_m.$$

In the equation above, we can replace v_k with the right side of 2.20, which shows that u is in the span of the list obtained by removing the kth term from $v_1, \dots, v_m$. Thus removing the kth term of the list $v_1, \dots, v_m$ does not change the span of the list.

If $k = 1$ in the linear dependence lemma, then $v_k \in \operatorname{span}(v_1, \dots, v_{k-1})$ means that $v_1 = 0$, because $\operatorname{span}(\,) = \{0\}$. Note also that parts of the proof of the linear dependence lemma need to be modified if $k = 1$. In general, the proofs in the rest of the book will not call attention to special cases that must be considered involving lists of length 0, the subspace $\{0\}$, or other trivial cases for which the result is true but needs a slightly different proof. Be sure to check these special cases yourself.

2.21 example: *smallest k in linear dependence lemma*

Consider the list

$$(1, 2, 3), (6, 5, 4), (15, 16, 17), (8, 9, 7)$$

in $\mathbf{R}^3$. This list of length four is linearly dependent, as we will soon see. Thus the linear dependence lemma implies that there exists $k \in \{1, 2, 3, 4\}$ such that the kth vector in this list is a linear combination of the previous vectors in the list. Let's see how to find the smallest value of k that works.

Taking $k = 1$ in the linear dependence lemma works if and only if the first vector in the list equals 0. Because $(1, 2, 3)$ is not the 0 vector, we cannot take $k = 1$ for this list.

Taking $k = 2$ in the linear dependence lemma works if and only if the second vector in the list is a scalar multiple of the first vector. However, there does not exist $c \in \mathbf{R}$ such that $(6, 5, 4) = c(1, 2, 3)$. Thus we cannot take $k = 2$ for this list.

Taking $k = 3$ in the linear dependence lemma works if and only if the third vector in the list is a linear combination of the first two vectors. Thus for the list in this example, we want to know whether there exist $a, b \in \mathbf{R}$ such that

$$(15, 16, 17) = a(1, 2, 3) + b(6, 5, 4).$$

The equation above is equivalent to a system of three linear equations in the two unknowns a, b. Using Gaussian elimination or appropriate software, we find that $a = 3, b = 2$ is a solution of the equation above, as you can verify. Thus for the list in this example, taking $k = 3$ is the smallest value of k that works in the linear dependence lemma.

כעת, נניח ש־ k הוא אינדקס כלשהו ב־ $\{1, \dots, m\}$ כך ש־ $v_k \in \operatorname{span}(v_1, \dots, v_{k-1})$. יהיו $b_1, \dots, b_{k-1} \in \mathbf{F}$ כך ש־

$$v_k = b_1 v_1 + \cdots + b_{k-1} v_{k-1}. \tag{2.20}$$

נניח ש־ $u \in \operatorname{span}(v_1, \dots, v_m)$. כלומר, קיימים סקלרים $c_1, \dots, c_m \in \mathbf{F}$ כך ש־

$$u = c_1 v_1 + \cdots + c_m v_m.$$

במשוואה לעיל נוכל להחליף את v_k בביטוי באגף הימני של 2.20, כלומר u הוא בפריסה של הרשימה, גם אחרי שהסרנו את האיבר ה־ k^{th} מ־ $v_1, \dots, v_m$. לפיכך הסרת האיבר ה־ k^{th} מהרשימה $v_1, \dots, v_m$ אינה משפיעה על פריסתה. ■

אם $k = 1$ בלֻמַה לעיל, אז המשמעות של $v_k \in \operatorname{span}(v_1, \dots, v_{k-1})$ היא ש־ $v_1 = 0$, מאחר ש־ $\operatorname{span}(\ ) = \{0\}$. נציין כי חלקים בהוכחה ידרשו שינויים אם $k = 1$. באופן כללי, בהוכחות בספר לא נסב את תשומת לבכם למקרים מיוחדים שעלינו לקחת בחשבון אם הם קשורים לרשימות באורך 0, לתת־מרחב $\{0\}$, או למקרים טריוויאליים אחרים שעבורם התוצאה עדיין נכונה, אך דורשים הוכחה שונה מעט. ודאו כי אתם בודקים מקרים מיוחדים אלו בעצמכם.

2.21 דוגמה: ה־ k הקטן ביותר בְּלֻמַת התלות הליניארית

נתבונן ברשימה

$$(1, 2, 3), (6, 5, 4), (15, 16, 17), (8, 9, 7)$$

ב־ $\mathbf{R}^3$. רשימה זו בת ארבעה וקטורים היא תלויה ליניארית, כפי שנראה. לפיכך, לפי הלמה שזה עתה הכרנו, קיים $k \in \{1, 2, 3, 4\}$ כך שהווקטור ה־ k^{th} ברשימה הוא צירוף ליניארי של קודמיו. הבה נראה כיצד נמצא את הערך הקטן ביותר עבור k זה.

עבור $k = 1$ נקבל תוצאה נכונה אם־ורק־אם הווקטור הראשון ברשימה הוא 0. מאחר ש־ $(1, 2, 3)$ אינו וקטור האפס, נפסלת האפשרות כי $k = 1$.

עבור $k = 2$ נקבל תוצאה נכונה אם־ורק־אם הווקטור השני ברשימה הוא כפולת סקלר של הווקטור הראשון. מאחר שאין בנמצא $c \in \mathbf{R}$ כך ש־ $(6, 5, 4) = c(1, 2, 3)$, לפיכך נפסלת האפשרות כי $k = 2$.

עבור $k = 3$ נקבל תוצאה נכונה אם־ורק־אם הווקטור השלישי ברשימה הוא צירוף ליניארי של שני הווקטורים הראשונים. לפיכך עבור הרשימה בדוגמה זו, אנו רוצים לברר האם קיימים $a, b \in \mathbf{R}$ כך ש־

$$(15, 16, 17) = a(1, 2, 3) + b(6, 5, 4).$$

משוואה זו שקולה למערכת של שלוש משוואות ליניאריות בשני נעלמים a, b. תוך שימוש בשיטת החילוץ של גאוס או בתוכנה מתאימה, נגלה כי $a = 3$, $b = 2$ הם פתרון למשוואה לעיל, כפי שתוכלו לוודא. לפיכך עבור דוגמה זו, $k = 3$ הוא הערך הקטן ביותר שאיתו נוכל להפעיל את הלֻמַה.

Now we come to a key result. It says that no linearly independent list in V is longer than a spanning list in V.

2.22 *length of linearly independent list ≤ length of spanning list*

In a finite-dimensional vector space, the length of every linearly independent list of vectors is less than or equal to the length of every spanning list of vectors.

Proof Suppose that $u_1, \dots, u_m$ is linearly independent in V. Suppose also that $w_1, \dots, w_n$ spans V. We need to prove that $m \le n$. We do so through the process described below with m steps; note that in each step we add one of the u's and remove one of the w's.

Step 1

Let B be the list $w_1, \dots, w_n$, which spans V. Adjoining u_1 at the beginning of this list produces a linearly dependent list (because u_1 can be written as a linear combination of $w_1, \dots, w_n$). In other words, the list

$$u_1, w_1, \dots, w_n$$

is linearly dependent.

Thus by the linear dependence lemma (2.19), one of the vectors in the list above is a linear combination of the previous vectors in the list. We know that $u_1 \neq 0$ because the list $u_1, \dots, u_m$ is linearly independent. Thus u_1 is not in the span of the previous vectors in the list above (because u_1 is not in $\{0\}$, which is the span of the empty list). Hence the linear dependence lemma implies that we can remove one of the w's so that the new list B (of length n) consisting of u_1 and the remaining w's spans V.

Step k, for $k = 2, \dots, m$

The list B (of length n) from step $k-1$ spans V. In particular, u_k is in the span of the list B. Thus the list of length $(n+1)$ obtained by adjoining u_k to B, placing it just after $u_1, \dots, u_{k-1}$, is linearly dependent. By the linear dependence lemma (2.19), one of the vectors in this list is in the span of the previous ones, and because $u_1, \dots, u_k$ is linearly independent, this vector cannot be one of the u's.

Hence there still must be at least one remaining w at this step. We can remove from our new list (after adjoining u_k in the proper place) a w that is a linear combination of the previous vectors in the list, so that the new list B (of length n) consisting of $u_1, \dots, u_k$ and the remaining w's spans V.

After step m, we have added all the u's and the process stops. At each step as we add a u to B, the linear dependence lemma implies that there is some w to remove. Thus there are at least as many w's as u's. ∎

כעת אנו מגיעים למסקנה מרכזית. במסקנה נקבע כי לא קיימת רשימה בלתי תלויה ליניארית ב־ V שמספר איבריה גדול מזה של רשימה הפורסת את V.

2.22 מספר הווקטורים ברשימה בלתי תלויה ליניארית $\leq$ מזה של רשימה פורסת

במרחב וקטורי מממד סופי, מספר הווקטורים בכל רשימה בלתי תלויה ליניארית הוא קטן או שווה לזה של כל רשימה פורסת.

הוכחה נניח ש־ $u_1, \dots, u_m$ בלתי תלויה ליניארית ב־ V. בנוסף, נניח ש־ $w_1, \dots, w_n$ פורסת את V. עלינו להוכיח ש־ $m \leq n$. נעשה זאת באמצעות תהליך בן m שלבים המפורט להלן; שימו לב, כי בכל שלב אנו מוסיפים את אחד ה־ u־ים וכנגדו מסירים את אחד ה w־ים.

שלב 1

תהי B הרשימה $w_1, \dots, w_n$ הפורסת את V. צירוף u_1 כאיבר הראשון ברשימה ייצור רשימה תלויה ליניארית (מאחר ש־ u_1 יכול להיכתב כצירוף ליניארי של $w_1, \dots, w_n$). במילים אחרות, הרשימה

$$u_1, w_1, \dots, w_n$$

תלויה ליניארית.

לפיכך לפי לֶמַת התלות הליניארית (2.19), אחד הווקטורים ברשימה הוא צירוף ליניארי של קודמיו. ידוע לנו ש־ $u_1 \neq 0$ שכן הרשימה $u_1, \dots, u_m$ היא בלתי תלויה ליניארית. לפיכך, u_1 אינו בפריסה של קודמיו (מאחר ש־ u_1 אינו בפריסה של $\{0\}$, שהיא הפריסה של הרשימה הריקה). לכן מהלמה נובע, כי נוכל להסיר את אחד ה־ w־ים כך שהרשימה החדשה B (בת n איברים) המורכבת מ־ u_1 ושאר ה־ w־ים הנותרים פורסת את V.

שלב k, עבור k = 2, ..., m

הרשימה B (עם n איברים) משלב $k-1$ פורסת את V. בפרט הווקטור u_k הוא בפריסה של הרשימה B. לפיכך הרשימה בת $(n+1)$ איברים שהושגה באמצעות הוספת הווקטור u_k ל־ B, שאותו הצבנו מייד אחרי $u_1, \dots, u_{k-1}$, תלויה ליניארית. לכן, לפי לֶמַת התלות הליניארית (2.19), אחד הווקטורים ברשימה הוא צירוף ליניארי של קודמיו. מאחר ש־ $u_1, \dots, u_k$ בלתי תלויה ליניארית, וקטור זה אינו יכול להיות אחד ה־ u־ים.

לכן בהכרח נותר לפחות w אחד בשלב זה. מהרשימה החדשה שבידינו (אחרי שצירפנו u_k למקום הנכון) נוכל להסיר w שהוא צירוף ליניארי של הווקטורים המופיעים לפניו ברשימה, כך שהרשימה החדשה B (שאורכה n) המורכבת מ־ $u_1, \dots, u_k$ וה־ w־ים הנותרים פורסת את V.

אחרי m שלבים, הוספנו את כל ה־ u־ים והתהליך נעצר. בכל שלב שבו הוספנו u ל־ B, לפי הלמה הסרנו מהרשימה w כנגדו. לפיכך מספר ה־ w־ים הוא לפחות כמספר ה־ u־ים. ■

The next two examples show how the result above can be used to show, without any computations, that certain lists are not linearly independent and that certain lists do not span a given vector space.

2.23 example: *no list of length 4 is linearly independent in* $\mathbf{R}^3$

The list $(1, 0, 0), (0, 1, 0), (0, 0, 1)$, which has length three, spans $\mathbf{R}^3$. Thus no list of length larger than three is linearly independent in $\mathbf{R}^3$.

For example, we now know that $(1, 2, 3), (4, 5, 8), (9, 6, 7), (-3, 2, 8)$, which is a list of length four, is not linearly independent in $\mathbf{R}^3$.

2.24 example: *no list of length 3 spans* $\mathbf{R}^4$

The list $(1, 0, 0, 0), (0, 1, 0, 0), (0, 0, 1, 0), (0, 0, 0, 1)$, which has length four, is linearly independent in $\mathbf{R}^4$. Thus no list of length less than four spans $\mathbf{R}^4$.

For example, we now know that $(1, 2, 3, -5), (4, 5, 8, 3), (9, 6, 7, -1)$, which is a list of length three, does not span $\mathbf{R}^4$.

Our intuition suggests that every subspace of a finite-dimensional vector space should also be finite-dimensional. We now prove that this intuition is correct.

2.25 *finite-dimensional subspaces*

Every subspace of a finite-dimensional vector space is finite-dimensional.

Proof Suppose V is finite-dimensional and U is a subspace of V. We need to prove that U is finite-dimensional. We do this through the following multistep construction.

Step 1

If $U = \{0\}$, then U is finite-dimensional and we are done. If $U \neq \{0\}$, then choose a nonzero vector $u_1 \in U$.

Step *k*

If $U = \operatorname{span}(u_1, \dots, u_{k-1})$, then U is finite-dimensional and we are done. If $U \neq \operatorname{span}(u_1, \dots, u_{k-1})$, then choose a vector $u_k \in U$ such that

$$u_k \notin \operatorname{span}(u_1, \dots, u_{k-1}).$$

After each step, as long as the process continues, we have constructed a list of vectors such that no vector in this list is in the span of the previous vectors. Thus after each step we have constructed a linearly independent list, by the linear dependence lemma (2.19). This linearly independent list cannot be longer than any spanning list of V (by 2.22). Thus the process eventually terminates, which means that U is finite-dimensional. ∎

שתי הדוגמאות הבאות מראות כיצד מהמסקנה האחרונה נוכל לקבוע, ללא פעולות חישוב, כי רשימות מסוימות אינן בלתי תלויות ליניארית, וכי רשימות מסוימות אינן פורסות מרחב וקטורי נתון.

2.23 דוגמה: ב־ $\mathbf{R}^3$ אין רשימה בלתי תלויה ליניארית בת 4 איברים

הרשימה $(1,0,0), (0,1,0), (0,0,1)$, שאורכה שלושה איברים, פורסת את $\mathbf{R}^3$. לפיכך אין רשימה בלתי תלויה ליניארית שאורכה גדול מ־ 3 ב־ $\mathbf{R}^3$.

לדוגמה, אנו יודעים כי $(1,2,3), (4,5,8), (9,6,7), (-3,2,8)$, שהיא רשימה בת ארבעה איברים, אינה בלתי תלויה ליניארית ב־ $\mathbf{R}^3$.

2.24 דוגמה: אין רשימה בת 3 איברים הפורסת את $\mathbf{R}^4$

הרשימה $(1,0,0,0), (0,1,0,0), (0,0,1,0), (0,0,0,1)$, באורך ארבע, היא בלתי תלויה ליניארית ב־ $\mathbf{R}^4$. לפיכך אין רשימה שאורכה קטן מארבעה הפורסת את $\mathbf{R}^4$.

לדוגמה, אנו יודעים כי $(1,2,3,-5), (4,5,8,3), (9,6,7,-1)$, רשימה באורך שלוש, אינה פורסת את $\mathbf{R}^4$.

באופן אינטואיטיבי, אנו מבינים כי כל תת־מרחב של מרחב מממד סופי יהיה אף הוא מממד סופי. כעת נוכיח את נכונות הטענה.

2.25 תתי־מרחב מממד סופי

כל תת־מרחב של מרחב וקטורי מממד סופי הוא מממד סופי.

הוכחה נניח ש־ V מרחב וקטורי מממד סופי, ו־ U תת־מרחב של V. עלינו להוכיח כי U מממד סופי. נעשה זאת בתהליך רב־שלבי.

שלב 1

אם $U = \{0\}$ אזי U מממד סופי, ובזאת הסתיימה ההוכחה. אם $U \neq \{0\}$, נבחר וקטור שונה מאפס $u_1 \in U$.

שלב k

אם $U = \operatorname{span}(u_1, \dots, u_{k-1})$, אזי U מממד סופי, ובזאת הסתיימה ההוכחה. אם $U \neq \operatorname{span}(u_1, \dots, u_{k-1})$, נבחר וקטור $u_k \in U$ כך ש־

$$u_k \notin \operatorname{span}(u_1, \dots, u_{k-1}).$$

בסיום כל שלב, כל עוד התהליך נמשך, הרכבנו רשימה של וקטורים כך שאף וקטור ברשימה אינו בפריסה של הווקטורים שקדמו לו. לפיכך, אחרי כל שלב יצרנו רשימה בלתי תלויה ליניארית, לפי לֶמַת התלות הליניארית (2.19). אורך רשימה זו אינו יכול להיות גדול מזה של רשימה הפורסת את V (לפי 2.22). לפיכך התהליך יסתיים בסופו של דבר, ומשמעות הדבר היא כי U מממד סופי. ■

Exercises 2A

1 Find a list of four distinct vectors in $\mathbf{F}^3$ whose span equals

$$\{(x, y, z) \in \mathbf{F}^3 : x + y + z = 0\}.$$

2 Prove or give a counterexample: If v_1, v_2, v_3, v_4 spans V, then the list

$$v_1 - v_2, v_2 - v_3, v_3 - v_4, v_4$$

also spans V.

3 Suppose $v_1, \dots, v_m$ is a list of vectors in V. For $k \in \{1, \dots, m\}$, let

$$w_k = v_1 + \cdots + v_k.$$

Show that $\operatorname{span}(v_1, \dots, v_m) = \operatorname{span}(w_1, \dots, w_m)$.

4 (a) Show that a list of length one in a vector space is linearly independent if and only if the vector in the list is not 0.

(b) Show that a list of length two in a vector space is linearly independent if and only if neither of the two vectors in the list is a scalar multiple of the other.

5 Find a number t such that

$$(3, 1, 4), (2, -3, 5), (5, 9, t)$$

is not linearly independent in $\mathbf{R}^3$.

6 Show that the list $(2, 3, 1), (1, -1, 2), (7, 3, c)$ is linearly dependent in $\mathbf{F}^3$ if and only if $c = 8$.

7 (a) Show that if we think of $\mathbf{C}$ as a vector space over $\mathbf{R}$, then the list $1 + i, 1 - i$ is linearly independent.

(b) Show that if we think of $\mathbf{C}$ as a vector space over $\mathbf{C}$, then the list $1 + i, 1 - i$ is linearly dependent.

8 Suppose v_1, v_2, v_3, v_4 is linearly independent in V. Prove that the list

$$v_1 - v_2, v_2 - v_3, v_3 - v_4, v_4$$

is also linearly independent.

9 Prove or give a counterexample: If $v_1, v_2, \dots, v_m$ is a linearly independent list of vectors in V, then

$$5v_1 - 4v_2, v_2, v_3, \dots, v_m$$

is linearly independent.

תרגילים 2ב

1 מצאו רשימה של ארבעה וקטורים שונים ב־ $\mathbf{F}^3$ שפריסתם שווה ל־
$$\{(x,y,z) \in \mathbf{F}^3 : x+y+z=0\}.$$

2 הוכיחו או הביאו דוגמה נגדית: אם v_1, v_2, v_3, v_4 פורסת את V, אזי
$$v_1 - v_2, v_2 - v_3, v_3 - v_4, v_4$$
גם היא פורסת את V.

3 נניח ש־ $v_1, \dots, v_m$ רשימת וקטורים ב־ V. עבור $k \in \{1, \dots, m\}$, יהי
$$w_k = v_1 + \cdots + v_k.$$
הראו ש־ $\operatorname{span}(v_1, \dots, v_m) = \operatorname{span}(w_1, \dots, w_m)$.

4 א. הראו כי במרחב וקטורי, רשימה בת וקטור יחיד היא בלתי תלויה ליניארית אם־ורק־אם הווקטור הזה אינו 0.

ב. הראו כי במרחב וקטורי, רשימה בת שני וקטורים היא בלתי תלויה ליניארית אם־ורק־אם אף אחד מהווקטורים אינו כפולת סקלר של האחר.

5 מצאו מספר t כך ש־
$$(3,1,4), (2,-3,5), (5,9,t)$$
אינה בלתי תלויה ליניארית ב־ $\mathbf{R}^3$.

6 הראו כי הרשימה $(2,3,1), (1,-1,2), (7,3,c)$ תלויה ליניארית ב־ $\mathbf{F}^3$ אם־ורק־אם $c = 8$.

7 א. הראו שאם אנו חושבים על $\mathbf{C}$ כעל מרחב וקטורי מעל $\mathbf{R}$, אזי הרשימה $1+i, 1-i$ בלתי תלויה ליניארית.

ב. הראו שאם אנו חושבים על $\mathbf{C}$ כמרחב וקטורי מעל $\mathbf{C}$, אזי הרשימה $1+i, 1-i$ תלויה ליניארית.

8 נניח ש־ v_1, v_2, v_3, v_4 בלתי תלויה ליניארית ב־ V. הוכיחו כי הרשימה
$$v_1 - v_2, v_2 - v_3, v_3 - v_4, v_4$$
גם היא בלתי תלויה ליניארית.

9 הוכיחו או הביאו דוגמה נגדית: אם $v_1, v_2, \dots, v_m$ רשימת וקטורים בלתי תלויה ליניארית ב־ V, אזי
$$5v_1 - 4v_2, v_2, v_3, \dots, v_m$$
בלתי תלויה ליניארית.

10 Prove or give a counterexample: If $v_1, v_2, \dots, v_m$ is a linearly independent list of vectors in V and $\lambda \in \mathbf{F}$ with $\lambda \neq 0$, then $\lambda v_1, \lambda v_2, \dots, \lambda v_m$ is linearly independent.

11 Prove or give a counterexample: If $v_1, \dots, v_m$ and $w_1, \dots, w_m$ are linearly independent lists of vectors in V, then the list $v_1 + w_1, \dots, v_m + w_m$ is linearly independent.

12 Suppose $v_1, \dots, v_m$ is linearly independent in V and $w \in V$. Prove that if $v_1 + w, \dots, v_m + w$ is linearly dependent, then $w \in \operatorname{span}(v_1, \dots, v_m)$.

13 Suppose $v_1, \dots, v_m$ is linearly independent in V and $w \in V$. Show that

$$v_1, \dots, v_m, w \text{ is linearly independent} \iff w \notin \operatorname{span}(v_1, \dots, v_m).$$

14 Suppose $v_1, \dots, v_m$ is a list of vectors in V. For $k \in \{1, \dots, m\}$, let

$$w_k = v_1 + \cdots + v_k.$$

Show that the list $v_1, \dots, v_m$ is linearly independent if and only if the list $w_1, \dots, w_m$ is linearly independent.

15 Explain why there does not exist a list of six polynomials that is linearly independent in $\mathcal{P}_4(\mathbf{F})$.

16 Explain why no list of four polynomials spans $\mathcal{P}_4(\mathbf{F})$.

17 Prove that V is infinite-dimensional if and only if there is a sequence $v_1, v_2, \dots$ of vectors in V such that $v_1, \dots, v_m$ is linearly independent for every positive integer m.

18 Prove that $\mathbf{F}^\infty$ is infinite-dimensional.

19 Prove that the real vector space of all continuous real-valued functions on the interval $[0, 1]$ is infinite-dimensional.

20 Suppose $p_0, p_1, \dots, p_m$ are polynomials in $\mathcal{P}_m(\mathbf{F})$ such that $p_k(2) = 0$ for each $k \in \{0, \dots, m\}$. Prove that $p_0, p_1, \dots, p_m$ is not linearly independent in $\mathcal{P}_m(\mathbf{F})$.

10 הוכיחו או הביאו דוגמה נגדית: אם $v_1, v_2, \dots, v_m$ רשימת וקטורים בלתי תלויה ליניארית ב־ V וגם $\lambda \in \mathbf{F}$ כאשר $\lambda \neq 0$, אזי $\lambda v_1, \lambda v_2, \dots, \lambda v_m$ בלתי תלויה ליניארית.

11 הוכיחו או הביאו דוגמה נגדית: אם $v_1, \dots, v_m$ וגם $w_1, \dots, w_m$ רשימות וקטורים בלתי תלויות ליניארית ב־ V, אזי הרשימה $v_1 + w_1, \dots, v_m + w_m$ בלתי תלויה ליניארית.

12 נניח ש־ $v_1, \dots, v_m$ בלתי תלויה ליניארית ב־ V ו־ $w \in V$. הוכיחו כי אם $v_1 + w, \dots, v_m + w$ תלויה ליניארית, אזי $w \in \operatorname{span}(v_1, \dots, v_m)$.

13 נניח ש־ $v_1, \dots, v_m$ בלתי תלויה ליניארית ב־ V ו־ $w \in V$. הראו כי $v_1, \dots, v_m, w$ בלתי תלויה ליניארית אם־ורק־אם $w \notin \operatorname{span}(v_1, \dots, v_m)$.

14 נניח ש־ $v_1, \dots, v_m$ רשימת וקטורים ב־ V. עבור $k \in \{1, \dots, m\}$, יהי

$$w_k = v_1 + \cdots + v_k.$$

הראו שהרשימה $v_1, \dots, v_m$ בלתי תלויה ליניארית אם־ורק־אם $w_1, \dots, w_m$ בלתי תלויה ליניארית.

15 הסבירו מדוע לא קיימת רשימה בלתי תלויה ליניארית של שישה פולינומים ב־ $\mathcal{P}_4(\mathbf{F})$.

16 הסבירו מדוע אין רשימה בת ארבעה פולינומים הפורסת את $\mathcal{P}_4(\mathbf{F})$.

17 הוכיחו כי V הוא מממד אינסופי אם־ורק־אם קיימת סדרת וקטורים $v_1, v_2, \dots$ ב־ V כך ש־ $v_1, \dots, v_m$ בלתי תלויה ליניארית לכל m שלם חיובי.

18 הוכיחו כי $\mathbf{F}^\infty$ הוא מממד אינסופי.

19 הוכיחו כי המרחב הווקטורי הממשי של כל הפונקציות הרציפות הממשיות מעל הקטע $[0, 1]$ הוא מממד אינסופי.

20 נניח ש־ $p_0, p_1, \dots, p_m$ פולינומים ב־ $\mathcal{P}_m(\mathbf{F})$ כך ש־ $p_k(2) = 0$ עבור כל $k \in \{0, \dots, m\}$. הוכיחו כי $p_0, p_1, \dots, p_m$ אינה בלתי תלויה ליניארית ב־ $\mathcal{P}_m(\mathbf{F})$.

2B Bases

In the previous section, we discussed linearly independent lists and we also discussed spanning lists. Now we bring these concepts together by considering lists that have both properties.

2.26 definition: *basis*

A *basis* of V is a list of vectors in V that is linearly independent and spans V.

2.27 example: *bases*

(a) The list $(1, 0, \dots, 0), (0, 1, 0, \dots, 0), \dots, (0, \dots, 0, 1)$ is a basis of $\mathbf{F}^n$, called the *standard basis* of $\mathbf{F}^n$.

(b) The list $(1, 2), (3, 5)$ is a basis of $\mathbf{F}^2$. Note that this list has length two, which is the same as the length of the standard basis of $\mathbf{F}^2$. In the next section, we will see that this is not a coincidence.

(c) The list $(1, 2, -4), (7, -5, 6)$ is linearly independent in $\mathbf{F}^3$ but is not a basis of $\mathbf{F}^3$ because it does not span $\mathbf{F}^3$.

(d) The list $(1, 2), (3, 5), (4, 13)$ spans $\mathbf{F}^2$ but is not a basis of $\mathbf{F}^2$ because it is not linearly independent.

(e) The list $(1, 1, 0), (0, 0, 1)$ is a basis of $\{(x, x, y) \in \mathbf{F}^3 : x, y \in \mathbf{F}\}$.

(f) The list $(1, -1, 0), (1, 0, -1)$ is a basis of

$$\{(x, y, z) \in \mathbf{F}^3 : x + y + z = 0\}.$$

(g) The list $1, z, \dots, z^m$ is a basis of $\mathcal{P}_m(\mathbf{F})$, called the *standard basis* of $\mathcal{P}_m(\mathbf{F})$.

In addition to the standard basis, $\mathbf{F}^n$ has many other bases. For example,

$$(7, 5), (-4, 9) \quad \text{and} \quad (1, 2), (3, 5)$$

are both bases of $\mathbf{F}^2$.

The next result helps explain why bases are useful. Recall that "uniquely" means "in only one way".

2.28 *criterion for basis*

A list $v_1, \dots, v_n$ of vectors in V is a basis of V if and only if every $v \in V$ can be written uniquely in the form

2.29
$$v = a_1 v_1 + \cdots + a_n v_n,$$

where $a_1, \dots, a_n \in \mathbf{F}$.

בסיסים 2ב

ביחידה הקודמת עסקנו ברשימות בלתי תלויות ליניארית, וברשימות פורסות. כעת נאחד שני רעיונות אלו כשנדון ברשימות המקיימות את שתי התכונות יחד.

2.26 הגדרה: בסיס

בסיס של V הוא רשימת וקטורים בלתי תלויה ליניארית ב־ V הפורסת את V.

2.27 דוגמה: בסיסים

א. הרשימה $(1,0,\dots,0),(0,1,0,\dots,0),\dots,(0,\dots,0,1)$ היא בסיס של $\mathbf{F}^n$, ונקראת הבסיס הסטנדרטי של $\mathbf{F}^n$.

ב. הרשימה $(1,2),(3,5)$ היא בסיס של $\mathbf{F}^2$. שימו לב כי לרשימה שני איברים, וזהו גם אורכו של הבסיס הסטנדרטי של $\mathbf{F}^2$. ביחידה הבאה נראה שאין מדובר בצירוף מקרים.

ג. הרשימה $(1,2,-4),(7,-5,6)$ בלתי תלויה ליניארית ב־ $\mathbf{F}^3$ אך אינה בסיס של $\mathbf{F}^3$, מאחר שאינה פורסת את $\mathbf{F}^3$.

ד. הרשימה $(1,2),(3,5),(4,13)$ פורסת את $\mathbf{F}^2$ אך אינה בסיס של $\mathbf{F}^2$, מאחר שהיא תלויה ליניארית.

ה. הרשימה $(1,1,0),(0,0,1)$ היא בסיס של $\{(x,x,y)\in\mathbf{F}^3 : x,y\in\mathbf{F}\}$.

ו. הרשימה $(1,-1,0),(1,0,-1)$ היא בסיס של
$$\{(x,y,z)\in\mathbf{F}^3 : x+y+z=0\}.$$

ז. הרשימה $1,z,\dots,z^m$ היא בסיס של $\mathcal{P}_m(\mathbf{F})$, ונקראת הבסיס הסטנדרטי של $\mathcal{P}_m(\mathbf{F})$.

בנוסף לבסיס הסטנדרטי, יש ל־ $\mathbf{F}^n$ בסיסים רבים אחרים. לדוגמה,
$$(1,2),(3,5) \quad \text{וגם} \quad (7,5),(-4,9)$$
שניהם בסיסים של $\mathbf{F}^2$.

המסקנה הבאה תסייע להבין מדוע בסיסים כה שימושיים. נזכיר ש־"באופן יחיד" משמעותו "אפשרות אחת בלבד".

2.28 תנחין לבסיס

הרשימה $v_1,\dots,v_n$ של וקטורים ב־ V היא בסיס של V אם־ורק־אם כל $v\in V$ יכול להיכתב באופן יחיד בצורה הבאה

2.29
$$v=a_1v_1+\cdots+a_nv_n,$$

כאשר $a_1,\dots,a_n\in\mathbf{F}$.

This proof is essentially a repetition of the ideas that led us to the definition of linear independence.

Proof First suppose that $v_1, \dots, v_n$ is a basis of V. Let $v \in V$. Because $v_1, \dots, v_n$ spans V, there exist $a_1, \dots, a_n \in \mathbf{F}$ such that 2.29 holds. To show that the representation in 2.29 is unique, suppose $c_1, \dots, c_n$ are scalars such that we also have

$$v = c_1 v_1 + \cdots + c_n v_n.$$

Subtracting the last equation from 2.29, we get

$$0 = (a_1 - c_1)v_1 + \cdots + (a_n - c_n)v_n.$$

This implies that each $a_k - c_k$ equals 0 (because $v_1, \dots, v_n$ is linearly independent). Hence $a_1 = c_1, \dots, a_n = c_n$. We have the desired uniqueness, completing the proof in one direction.

For the other direction, suppose every $v \in V$ can be written uniquely in the form given by 2.29. This implies that the list $v_1, \dots, v_n$ spans V. To show that $v_1, \dots, v_n$ is linearly independent, suppose $a_1, \dots, a_n \in \mathbf{F}$ are such that

$$0 = a_1 v_1 + \cdots + a_n v_n.$$

The uniqueness of the representation 2.29 (taking $v = 0$) now implies that $a_1 = \cdots = a_n = 0$. Thus $v_1, \dots, v_n$ is linearly independent and hence is a basis of V. ∎

A spanning list in a vector space may not be a basis because it is not linearly independent. Our next result says that given any spanning list, some (possibly none) of the vectors in it can be discarded so that the remaining list is linearly independent and still spans the vector space.

As an example in the vector space $\mathbf{F}^2$, if the procedure in the proof below is applied to the list $(1, 2), (3, 6), (4, 7), (5, 9)$, then the second and fourth vectors will be removed. This leaves $(1, 2), (4, 7)$, which is a basis of $\mathbf{F}^2$.

2.30 *every spanning list contains a basis*

Every spanning list in a vector space can be reduced to a basis of the vector space.

Proof Suppose $v_1, \dots, v_n$ spans V. We want to remove some of the vectors from $v_1, \dots, v_n$ so that the remaining vectors form a basis of V. We do this through the multistep process described below.

Start with B equal to the list $v_1, \dots, v_n$.

Step 1

If $v_1 = 0$, then delete v_1 from B. If $v_1 \neq 0$, then leave B unchanged.

Step *k*

If v_k is in $\operatorname{span}(v_1, \dots, v_{k-1})$, then delete v_k from the list B. If v_k is not in $\operatorname{span}(v_1, \dots, v_{k-1})$, then leave B unchanged.

הוכחה ראשית נניח ש־ $v_1, \dots, v_n$ בסיס של V. יהי $v \in V$. מאחר ש־ $v_1, \dots, v_n$ פורסת את V, קיימים $a_1, \dots, a_n \in \mathbf{F}$ כך ש־ 2.29 מתקיימת. כדי להראות שהייצוג ב־ 2.29 יחיד, נניח ש־ $c_1, \dots, c_n$ סקלרים כך ש־

הוכחה זו היא בעצם חזרה על הרעיונות שהובילו אותנו להגדרת אִי־תְּלוּת ליניארית.

$$v = c_1 v_1 + \cdots + c_n v_n.$$

בחיסור המשוואה האחרונה מ־ 2.29, נקבל

$$0 = (a_1 - c_1)v_1 + \cdots + (a_n - c_n)v_n.$$

שוויון זה מוביל למסקנה כי כל $a_k - c_k$ שווה 0 (מאחר ש־ $v_1, \dots, v_n$ בלתי תלויה ליניארית). לכן $a_1 = c_1, \dots, a_n = c_n$. בכך קיבלנו את היחידות הרצויה, והשלמנו את ההוכחה בכיוון אחד.

להוכחת הכיוון השני, נניח שכל $v \in V$ יכול להיכתב באופן יחיד בצורה המוצגת ב־ 2.29. כלומר, הרשימה $v_1, \dots, v_n$ פורסת את V. כדי להראות ש־ $v_1, \dots, v_n$ בלתי תלויה ליניארית, נניח שהסקלרים $a_1, \dots, a_n \in \mathbf{F}$ מקיימים

$$0 = a_1 v_1 + \cdots + a_n v_n.$$

יחידות הייצוג ב־ 2.29 (בהצבת $v = 0$) מורה לנו כעת ש־ $a_1 = \cdots = a_n = 0$. לפיכך, $v_1, \dots, v_n$ היא בלתי תלויה ליניארית, ולכן היא בסיס של V. ■

רשימה פורסת במרחב וקטורי עשויה שלא להיות בסיס מכיוון שהיא תלויה ליניארית. במסקנה הבאה נטען כי עבור כל רשימה פורסת, נוכל להסיר חלק מאיבריה (ואולי אף לא אחד) כך שנקבל רשימה בלתי תלויה ליניארית שעדיין פורסת את המרחב הווקטורי.

לדוגמה, במרחב הווקטורי $\mathbf{F}^2$, אם התהליך בהוכחה להלן יופעל על הרשימה $(1,2), (3,6), (4,7), (5,9)$, הווקטורים השני והרביעי יוסרו. הווקטורים $(1,2), (4,7)$ נותרו ברשימה, וכעת היא בסיס של $\mathbf{F}^2$.

2.30 כל רשימה פורסת מכילה בסיס

כל רשימה פורסת במרחב וקטורי ניתנת לצמצום לכדי בסיס של המרחב.

הוכחה נניח ש־ $v_1, \dots, v_n$ פורסת את V. נרצה להסיר מספר וקטורים מ־ $v_1, \dots, v_n$ כך שהווקטורים הנותרים מהווים בסיס של V. נעשה זאת באמצעות התהליך הרב־שלבי המתואר להלן.

ראשית נגדיר את B כרשימה $v_1, \dots, v_n$.

שלב 1

אם $v_1 = 0$, נסיר את v_1 מ־ B. אם $v_1 \neq 0$, נשאיר את B ללא שינוי.

שלב k

אם v_k ב־ $\operatorname{span}(v_1, \dots, v_{k-1})$, נסיר אותו מ־ B. אם v_k אינו ב־ $\operatorname{span}(v_1, \dots, v_{k-1})$, נשאיר את B ללא שינוי.

Stop the process after step n, getting a list B. This list B spans V because our original list spanned V and we have discarded only vectors that were already in the span of the previous vectors. The process ensures that no vector in B is in the span of the previous ones. Thus B is linearly independent, by the linear dependence lemma (2.19). Hence B is a basis of V. ∎

We now come to an important corollary of the previous result.

2.31 *basis of finite-dimensional vector space*

Every finite-dimensional vector space has a basis.

Proof By definition, a finite-dimensional vector space has a spanning list. The previous result tells us that each spanning list can be reduced to a basis. ∎

Our next result is in some sense a dual of 2.30, which said that every spanning list can be reduced to a basis. Now we show that given any linearly independent list, we can adjoin some additional vectors (this includes the possibility of adjoining no additional vectors) so that the extended list is still linearly independent but also spans the space.

2.32 *every linearly independent list extends to a basis*

Every linearly independent list of vectors in a finite-dimensional vector space can be extended to a basis of the vector space.

Proof Suppose $u_1, \dots, u_m$ is linearly independent in a finite-dimensional vector space V. Let $w_1, \dots, w_n$ be a list of vectors in V that spans V. Thus the list

$$u_1, \dots, u_m, w_1, \dots, w_n$$

spans V. Applying the procedure of the proof of 2.30 to reduce this list to a basis of V produces a basis consisting of the vectors $u_1, \dots, u_m$ and some of the w's (none of the u's get deleted in this procedure because $u_1, \dots, u_m$ is linearly independent). ∎

As an example in $\mathbf{F}^3$, suppose we start with the linearly independent list $(2, 3, 4), (9, 6, 8)$. If we take w_1, w_2, w_3 to be the standard basis of $\mathbf{F}^3$, then applying the procedure in the proof above produces the list

$$(2, 3, 4), (9, 6, 8), (0, 1, 0),$$

which is a basis of $\mathbf{F}^3$.

As an application of the result above, we now show that every subspace of a finite-dimensional vector space can be paired with another subspace to form a direct sum of the whole space.

Using the same ideas but more advanced tools, the next result can be proved without the hypothesis that V is finite-dimensional.

את התהליך נעצור אחרי שלב n, כשבידינו הרשימה B. רשימה זו – B – פורסת את V, מאחר שהרשימה המקורית שלנו פרסה את V, ואנו הסרנו רק את הווקטורים שכבר היו בפריסה של הווקטורים שקדמו להם. התהליך הבטיח שאין ברשימה B וקטורים שהם בפריסה של קודמיהם, ולפיכך B בלתי תלויה ליניארית, לפי למת התלות הליניארית (2.19). לכן B בסיס של V. ■

כעת נגיע למסקנה חשובה הנגררת מהמסקנה הקודמת.

2.31 בסיס של מרחב וקטורי מממד סופי

לכל מרחב וקטורי מממד סופי יש בסיס.

הוכחה לפי הגדרה, למרחב וקטורי מממד סופי יש רשימה פורסת. המסקנה הקודמת מלמדת אותנו שניתן לצמצם כל רשימה פורסת לכדי בסיס. ■

המסקנה הבאה היא במובן מסוים תמונת ראי של 2.30, שבה נקבע שניתן לצמצם כל רשימה פורסת לכדי בסיס. כעת נראה שכל רשימה בלתי תלויה ליניארית ניתנת להשלמה באמצעות הוספת וקטורים (כולל האפשרות שלא יתווספו וקטורים כלל) כך שנקבל רשימה מורחבת הפורסת את המרחב, אך עדיין בלתי תלויה ליניארית.

2.32 כל רשימה בלתי תלויה ליניארית ניתנת להשלמה לכדי בסיס

כל רשימת וקטורים בלתי תלויה ליניארית במרחב וקטורי מממד סופי ניתנת להשלמה לבסיס של המרחב הווקטורי.

הוכחה נניח ש־ $u_1, \dots, u_m$ רשימה בלתי תלויה ליניארית במרחב וקטורי מממד סופי V. תהי $w_1, \dots, w_n$ רשימת וקטורים הפורסת את V. לפיכך, הרשימה

$$u_1, \dots, u_m, w_1, \dots, w_n$$

פורסת את V. הפעלת התהליך שפירטנו בהוכחה 2.30 לצמצום הרשימה לבסיס של V תניב בסיס המורכב מה־ u־ים בתוספת חלק מה־ w־ים (את וקטורי u־ים לא נמחק הפעם שכן הרשימה $u_1, \dots, u_m$ בלתי תלויה ליניארית). ■

כדוגמה ב־ $\mathbf{F}^3$, נניח שאנו מתחילים את התהליך עם הרשימה הבלתי תלויה ליניארית $(2,3,4), (9,6,8)$. אם נגדיר את w_1, w_2, w_3 כבסיס הסטנדרטי של $\mathbf{F}^3$, נוכל ליישם את התהליך שתואר בהוכחה ולקבל כתוצאה את הרשימה

$$(2,3,4), (9,6,8), (0,1,0),$$

שהיא בסיס של $\mathbf{F}^3$.

כיישום של המסקנה הקודמת, נראה כי כל תת־מרחב של מרחב וקטורי מממד סופי יכול להצטרף לתת־מרחב נוסף וליצור יחד איתו סכום ישר השווה למרחב כולו.

בעזרת אותם רעיונות וכלים מתקדמים יותר, ניתן להוכיח את המסקנה הבאה גם ללא ההנחה ש־ V מממד סופי.

2.33 *every subspace of V is part of a direct sum equal to V*

Suppose V is finite-dimensional and U is a subspace of V. Then there is a subspace W of V such that $V = U \oplus W$.

Proof Because V is finite-dimensional, so is U (see 2.25). Thus there is a basis $u_1, \dots, u_m$ of U (by 2.31). Of course $u_1, \dots, u_m$ is a linearly independent list of vectors in V. Hence this list can be extended to a basis $u_1, \dots, u_m, w_1, \dots, w_n$ of V (by 2.32). Let $W = \operatorname{span}(w_1, \dots, w_n)$.

To prove that $V = U \oplus W$, by 1.46 we only need to show that

$$V = U + W \quad \text{and} \quad U \cap W = \{0\}.$$

To prove the first equation above, suppose $v \in V$. Then, because the list $u_1, \dots, u_m, w_1, \dots, w_n$ spans V, there exist $a_1, \dots, a_m, b_1, \dots, b_n \in \mathbf{F}$ such that

$$v = \underbrace{a_1 u_1 + \cdots + a_m u_m}_{u} + \underbrace{b_1 w_1 + \cdots + b_n w_n}_{w}.$$

We have $v = u + w$, where $u \in U$ and $w \in W$ are defined as above. Thus $v \in U + W$, completing the proof that $V = U + W$.

To show that $U \cap W = \{0\}$, suppose $v \in U \cap W$. Then there exist scalars $a_1, \dots, a_m, b_1, \dots, b_n \in \mathbf{F}$ such that

$$v = a_1 u_1 + \cdots + a_m u_m = b_1 w_1 + \cdots + b_n w_n.$$

Thus

$$a_1 u_1 + \cdots + a_m u_m - b_1 w_1 - \cdots - b_n w_n = 0.$$

Because $u_1, \dots, u_m, w_1, \dots, w_n$ is linearly independent, this implies that

$$a_1 = \cdots = a_m = b_1 = \cdots = b_n = 0.$$

Thus $v = 0$, completing the proof that $U \cap W = \{0\}$. ∎

Exercises 2B

1 Find all vector spaces that have exactly one basis.

2 Verify all assertions in Example 2.27.

3 (a) Let U be the subspace of $\mathbf{R}^5$ defined by

$$U = \{(x_1, x_2, x_3, x_4, x_5) \in \mathbf{R}^5 : x_1 = 3x_2 \text{ and } x_3 = 7x_4\}.$$

Find a basis of U.

(b) Extend the basis in (a) to a basis of $\mathbf{R}^5$.

(c) Find a subspace W of $\mathbf{R}^5$ such that $\mathbf{R}^5 = U \oplus W$.

2.33 כל תת־מרחב של V הוא חלק מסכום ישר השווה ל־ V

נניח ש־ V הוא מרחב מממד סופי, ו־ U תת־מרחב של V. אזי קיים תת־מרחב W של V כך ש־ $V = U \oplus W$.

הוכחה מאחר ש־ V הוא מממד סופי, כך גם U (ראו 2.25). לפיכך, קיים בסיס $u_1, \dots, u_m$ של U (לפי 2.31). כמובן ש־ $u_1, \dots, u_m$ רשימה בלתי תלויה ליניארית של וקטורים ב V. לכן נוכל להרחיב רשימה זו לבסיס $u_1, \dots, u_m, w_1, \dots, w_n$ של V (לפי 2.32). נגדיר $W = \operatorname{span}(w_1, \dots, w_n)$.

כדי להוכיח כי $V = U \oplus W$, לפי 1.46 נצטרך רק להראות ש־

$$V = U + W \quad \text{וגם} \quad U \cap W = \{0\}.$$

להוכחת המשוואה הראשונה, נניח ש־ $v \in V$. מאחר שהרשימה $u_1, \dots, u_m, w_1, \dots, w_n$ פורסת את V, קיימים סקלרים $a_1, \dots, a_m, b_1, \dots, b_n \in \mathbf{F}$ כך ש־

$$v = \underbrace{a_1 u_1 + \cdots + a_m u_m}_{u} + \underbrace{b_1 w_1 + \cdots + b_n w_n}_{w}.$$

קיבלנו אפוא $v = u + w$, כאשר $u \in U$ ו־ $w \in W$ כפי שהוגדרו. לפיכך $v \in U + W$, ובכך השלמנו את ההוכחה ש־ $V = U + W$.

להוכחת המשוואה השנייה, $U \cap W = \{0\}$, נניח ש־ $v \in U \cap W$. אזי קיימים סקלרים $a_1, \dots, a_m, b_1, \dots, b_n \in \mathbf{F}$ כך ש־

$$v = a_1 u_1 + \cdots + a_m u_m = b_1 w_1 + \cdots + b_n w_n.$$

לפיכך נוכל לסדר מחדש את המשוואה כך

$$a_1 u_1 + \cdots + a_m u_m - b_1 w_1 - \cdots - b_n w_n = 0.$$

ומאחר ש־ $u_1, \dots, u_m, w_1, \dots, w_n$ בלתי תלויה ליניארית, נוכל לקבוע כי

$$a_1 = \cdots = a_m = b_1 = \cdots = b_n = 0.$$

לפיכך $v = 0$, ובכך השלמנו את ההוכחה ש־ $U \cap W = \{0\}$. ■

תרגילים ב2

1 מצאו את כל המרחבים הווקטוריים שיש להם בסיס אחד בלבד.

2 ודאו את כל הקביעות בדוגמה 2.27.

3 א. יהי U התת־מרחב של $\mathbf{R}^5$ המוגדר על ידי

$$U = \{(x_1, x_2, x_3, x_4, x_5) \in \mathbf{R}^5 : x_1 = 3x_2 \text{ and } x_3 = 7x_4\}.$$

מצאו בסיס של U.

ב. השלימו את הבסיס ב־ (א) לבסיס של $\mathbf{R}^5$.

ג. מצאו תת־מרחב W של $\mathbf{R}^5$ כך ש־ $\mathbf{R}^5 = U \oplus W$.

4 (a) Let U be the subspace of $\mathbf{C}^5$ defined by

$$U = \{(z_1, z_2, z_3, z_4, z_5) \in \mathbf{C}^5 : 6z_1 = z_2 \text{ and } z_3 + 2z_4 + 3z_5 = 0\}.$$

Find a basis of U.

(b) Extend the basis in (a) to a basis of $\mathbf{C}^5$.

(c) Find a subspace W of $\mathbf{C}^5$ such that $\mathbf{C}^5 = U \oplus W$.

5 Suppose V is finite-dimensional and U, W are subspaces of V such that $V = U + W$. Prove that there exists a basis of V consisting of vectors in $U \cup W$.

6 Prove or give a counterexample: If p_0, p_1, p_2, p_3 is a list in $\mathcal{P}_3(\mathbf{F})$ such that none of the polynomials p_0, p_1, p_2, p_3 has degree 2, then p_0, p_1, p_2, p_3 is not a basis of $\mathcal{P}_3(\mathbf{F})$.

7 Suppose v_1, v_2, v_3, v_4 is a basis of V. Prove that

$$v_1 + v_2, v_2 + v_3, v_3 + v_4, v_4$$

is also a basis of V.

8 Prove or give a counterexample: If v_1, v_2, v_3, v_4 is a basis of V and U is a subspace of V such that $v_1, v_2 \in U$ and $v_3 \notin U$ and $v_4 \notin U$, then v_1, v_2 is a basis of U.

9 Suppose $v_1, \dots, v_m$ is a list of vectors in V. For $k \in \{1, \dots, m\}$, let

$$w_k = v_1 + \cdots + v_k.$$

Show that $v_1, \dots, v_m$ is a basis of V if and only if $w_1, \dots, w_m$ is a basis of V.

10 Suppose U and W are subspaces of V such that $V = U \oplus W$. Suppose also that $u_1, \dots, u_m$ is a basis of U and $w_1, \dots, w_n$ is a basis of W. Prove that

$$u_1, \dots, u_m, w_1, \dots, w_n$$

is a basis of V.

11 Suppose V is a real vector space. Show that if $v_1, \dots, v_n$ is a basis of V (as a real vector space), then $v_1, \dots, v_n$ is also a basis of the complexification $V_{\mathbf{C}}$ (as a complex vector space).

See Exercise 8 in Section 1B for the definition of the complexification $V_{\mathbf{C}}$.

4 א. יהי U התת־מרחב של $\mathbf{C}^5$ המוגדר על ידי

$$U = \{(z_1, z_2, z_3, z_4, z_5) \in \mathbf{C}^5 : 6z_1 = z_2 \text{ and } z_3 + 2z_4 + 3z_5 = 0\}.$$

מצאו בסיס של U.

ב. השלימו את הבסיס ב־ (א) לבסיס של $\mathbf{C}^5$.

ג. מצאו תת־מרחב W של $\mathbf{C}^5$ כך ש־ $\mathbf{C}^5 = U \oplus W$.

5 נניח ש־ V הוא מרחב מממד סופי ו־ U, W תתי־מרחב של V כך ש־ $V = U + W$. הוכיחו כי קיים בסיס של V המורכב מווקטורים ב־ $U \cup W$.

6 הוכיחו או הביאו דוגמה נגדית: אם p_0, p_1, p_2, p_3 רשימה ב־ $\mathcal{P}_3(\mathbf{F})$ כך שאף אחד מהפולינומים p_0, p_1, p_2, p_3 אינו ממעלה 2, אזי p_0, p_1, p_2, p_3 אינה בסיס של $\mathcal{P}_3(\mathbf{F})$.

7 נניח ש־ v_1, v_2, v_3, v_4 בסיס של V. הוכיחו כי

$$v_1 + v_2, v_2 + v_3, v_3 + v_4, v_4$$

גם היא בסיס של V.

8 הוכיחו או הביאו דוגמה נגדית: אם v_1, v_2, v_3, v_4 בסיס של V ו־ U תת־מרחב של V כך ש־ $v_1, v_2 \in U$ ו־ $v_3 \notin U$ ו־ $v_4 \notin U$, אזי v_1, v_2 בסיס של U.

9 נניח ש־ $v_1, \dots, v_m$ רשימת וקטורים ב־ V. עבור $k \in \{1, \dots, m\}$, יהי

$$w_k = v_1 + \cdots + v_k.$$

הראו ש־ $v_1, \dots, v_m$ בסיס של V אם־ורק־אם $w_1, \dots, w_m$ בסיס של V.

10 נניח ש־ U ו־ W תתי־מרחב של V כך ש־ $V = U \oplus W$. נניח בנוסף ש־ $u_1, \dots, u_m$ בסיס של U ו־ $w_1, \dots, w_n$ בסיס של W. הוכיחו כי

$$u_1, \dots, u_m, w_1, \dots, w_n$$

בסיס של V.

11 נניח ש־ V מרחב וקטורי ממשי. הראו כי אם $v_1, \dots, v_n$ בסיס של V (כמרחב וקטורי ממשי), אזי $v_1, \dots, v_n$ היא גם בסיס של המירכוב $V_{\mathbf{C}}$ (כמרחב וקטורי מרוכב).

ראו תרגיל 8 ביחידה אז להגדרת המירכוב $V_{\mathbf{C}}$.

אלגברה ליניארית בדרך הישר, מהדורה רביעית, מחבר: שלדון אקסלר
Linear Algebra Done Right, fourth edition, by Sheldon Axler

2C Dimension

Although we have been discussing finite-dimensional vector spaces, we have not yet defined the dimension of such an object. How should dimension be defined? A reasonable definition should force the dimension of $\mathbf{F}^n$ to equal n. Notice that the standard basis

$$(1, 0, \dots, 0), (0, 1, 0, \dots, 0), \dots, (0, \dots, 0, 1)$$

of $\mathbf{F}^n$ has length n. Thus we are tempted to define the dimension as the length of a basis. However, a finite-dimensional vector space in general has many different bases, and our attempted definition makes sense only if all bases in a given vector space have the same length. Fortunately that turns out to be the case, as we now show.

2.34 *basis length does not depend on basis*

Any two bases of a finite-dimensional vector space have the same length.

Proof Suppose V is finite-dimensional. Let B_1 and B_2 be two bases of V. Then B_1 is linearly independent in V and B_2 spans V, so the length of B_1 is at most the length of B_2 (by 2.22). Interchanging the roles of B_1 and B_2, we also see that the length of B_2 is at most the length of B_1. Thus the length of B_1 equals the length of B_2, as desired. ∎

Now that we know that any two bases of a finite-dimensional vector space have the same length, we can formally define the dimension of such spaces.

2.35 definition: *dimension,* $\dim V$

- The *dimension* of a finite-dimensional vector space is the length of any basis of the vector space.
- The dimension of a finite-dimensional vector space V is denoted by $\dim V$.

2.36 example: *dimensions*

- $\dim \mathbf{F}^n = n$ because the standard basis of $\mathbf{F}^n$ has length n.
- $\dim \mathcal{P}_m(\mathbf{F}) = m + 1$ because the standard basis $1, z, \dots, z^m$ of $\mathcal{P}_m(\mathbf{F})$ has length $m + 1$.
- If $U = \{(x, x, y) \in \mathbf{F}^3 : x, y \in \mathbf{F}\}$, then $\dim U = 2$ because $(1, 1, 0), (0, 0, 1)$ is a basis of U.
- If $U = \{(x, y, z) \in \mathbf{F}^3 : x + y + z = 0\}$, then $\dim U = 2$ because the list $(1, -1, 0), (1, 0, -1)$ is a basis of U.

ממד

אף שעד כה עסקנו במרחבים וקטוריים מממד סופי, עדיין לא הגדרנו את הממד של אובייקט מסוג זה. באיזה אופן נגדיר את ממד המרחב? ההיגיון מורה שממד המרחב $\mathbf{F}^n$ יהיה n. נשים לב כי בבסיס הסטנדרטי

$$(1, 0, \dots, 0), (0, 1, 0, \dots, 0), \dots, (0, \dots, 0, 1)$$

של $\mathbf{F}^n$ יש n איברים. לפיכך הנטייה הטבעית היא להגדיר את הממד כאורך הבסיס. אולם, למרחב וקטורי מממד סופי יש בסיסים רבים ושונים, ולכן ניסיון הגדרת הממד בהתאם למספר האיברים בבסיס יהיה תקף רק אם לכל הבסיסים אותו מספר איברים. למרבה המזל זה אכן המצב, כפי שנראה כעת.

2.34 מספר האיברים בבסיס אינו תלוי בבסיס

לכל שני בסיסים של מרחב וקטורי מממד סופי אותו אורך.

הוכחה נניח ש־ V מממד סופי. יהיו B_1 ו־ B_2 שני בסיסים של V. אזי, B_1 הוא בלתי תלוי ליניארית ב־ V ו־ B_2 פורס את V, ולכן אורכו של B_1 הוא לכל היותר כאורכו של B_2 (לפי 2.22). בחילוף תפקידים בין B_1 ל־ B_2, נקבל שגם אורכו של B_2 הוא לכל היותר כאורכו של B_1. לפיכך אורכו של B_1 שווה לאורכו של B_2, כפי שרצינו. ■

כעת משאנו יודעים כי לכל שני בסיסים של מרחב וקטורי מממד סופי אותו אורך, נוכל להגדיר את ממדם של מרחבים אלו באופן רשמי.

2.35 הגדרה: ממד, $\dim V$

- הממד של מרחב וקטורי מממד סופי הוא מספר האיברים של כל אחד מבסיסי המרחב.
- ממד מרחב וקטורי מממד סופי V מסומן ב־ $\dim V$.

2.36 דוגמה: ממדים

- $\dim \mathbf{F}^n = n$ מאחר שלבסיס הסטנדרטי של $\mathbf{F}^n$ יש n איברים.
- $\dim \mathcal{P}_m(\mathbf{F}) = m + 1$ מאחר שלבסיס הסטנדרטי $1, z, \dots, z^m$ של $\mathcal{P}_m(\mathbf{F})$ יש $m + 1$ איברים.
- אם $U = \{(x, x, y) \in \mathbf{F}^3 : x, y \in \mathbf{F}\}$, אזי $\dim U = 2$ מאחר ש־ $(1, 1, 0), (0, 0, 1)$ בסיס של U.
- אם $U = \{(x, y, z) \in \mathbf{F}^3 : x + y + z = 0\}$, אזי $\dim U = 2$ מאחר שהרשימה $(1, -1, 0), (1, 0, -1)$ בסיס של U.

Every subspace of a finite-dimensional vector space is finite-dimensional (by 2.25) and so has a dimension. The next result gives the expected inequality about the dimension of a subspace.

2.37 *dimension of a subspace*

If V is finite-dimensional and U is a subspace of V, then $\dim U \le \dim V$.

Proof Suppose V is finite-dimensional and U is a subspace of V. Think of a basis of U as a linearly independent list in V, and think of a basis of V as a spanning list in V. Now use 2.22 to conclude that $\dim U \le \dim V$.

To check that a list of vectors in V is a basis of V, we must, according to the definition, show that the list in question satisfies two properties: it must be linearly independent and it must span V. The next two results show that if the list in question has the right length, then we only need to check that it satisfies one of the two required properties. First we prove that every linearly independent list of the right length is a basis.

The real vector space $\mathbf{R}^2$ has dimension two; the complex vector space $\mathbf{C}$ has dimension one. As sets, $\mathbf{R}^2$ can be identified with $\mathbf{C}$ (and addition is the same on both spaces, as is scalar multiplication by real numbers). Thus when we talk about the dimension of a vector space, the role played by the choice of $\mathbf{F}$ cannot be neglected.

2.38 *linearly independent list of the right length is a basis*

Suppose V is finite-dimensional. Then every linearly independent list of vectors in V of length $\dim V$ is a basis of V.

Proof Suppose $\dim V = n$ and $v_1, \dots, v_n$ is linearly independent in V. The list $v_1, \dots, v_n$ can be extended to a basis of V (by 2.32). However, every basis of V has length n, so in this case the extension is the trivial one, meaning that no elements are adjoined to $v_1, \dots, v_n$. Thus $v_1, \dots, v_n$ is a basis of V, as desired.

The next result is a useful consequence of the previous result.

2.39 *subspace of full dimension equals the whole space*

Suppose that V is finite-dimensional and U is a subspace of V such that $\dim U = \dim V$. Then $U = V$.

Proof Let $u_1, \dots, u_n$ be a basis of U. Thus $n = \dim U$, and by hypothesis we also have $n = \dim V$. Thus $u_1, \dots, u_n$ is a linearly independent list of vectors in V (because it is a basis of U) of length $\dim V$. From 2.38, we see that $u_1, \dots, u_n$ is a basis of V. In particular every vector in V is a linear combination of $u_1, \dots, u_n$. Thus $U = V$.

כל תת־מרחב של מרחב וקטורי מממד סופי הוא מממד סופי (לפי 2.25) ולפיכך יש לו ממד. המסקנה הבאה מספקת את האי־שוויון הצפוי בנוגע לממד תת־המרחב.

2.37 ממד תת־מרחב

אם V מממד סופי, ו־ U תת־מרחב של V, אזי $\dim U \leq \dim V$.

הוכחה נניח ש־ V מרחב מממד סופי ו־ U תת־מרחב של V. נתייחס לבסיס של U כרשימת וקטורים בלתי תלויה ליניארית ב־ V, ונתייחס לבסיס של V כרשימה פורסת ב־ V. כעת נשתמש ב־ 2.22 להסיק כי $\dim U \leq \dim V$. ■

כדי לוודא שרשימת וקטורים ב־ V היא בסיס של V, עלינו, לפי ההגדרה, להראות שהרשימה מקיימת שני תנאים: היא בלתי תלויה ליניארית והיא פורסת את V. שתי המסקנות הבאות יראו לנו כי אם אורכה של רשימה תואם את ממד המרחב, מספיק שנוודא תנאי אחד מהשניים. נוכיח תחילה שכל רשימה בלתי תלויה ליניארית עם מספר האיברים הנדרש היא בסיס.

המרחב הווקטורי המפשי $\mathbf{R}^2$ הוא דו־ממדי; והמרחב הווקטורי המרוכב $\mathbf{C}$ הוא חד־ממדי. כקבוצות, ניתן לזהות את $\mathbf{R}^2$ עם $\mathbf{C}$ (ופעולת החיבור זהה, כמו גם פעולת הכפל בסקלרים ממשיים). לפיכך כאשר אנו עוסקים בממד מרחב וקטורי, לבחירת ה־ $\mathbf{F}$ תפקיד לא מבוטל.

2.38 רשימה בלתי תלויה ליניארית באורך מתאים היא בסיס

נניח ש־ V מרחב מממד סופי. אזי כל רשימת וקטורים בלתי תלויה ליניארית ב־ V שאורכה שווה ל־ $\dim V$ היא בסיס של V.

הוכחה נניח ש־ $\dim V = n$ ו־ $v_1, \ldots, v_n$ בלתי תלויה ליניארית ב־ V. את הרשימה $v_1, \ldots, v_n$ נוכל להרחיב לבסיס של V (לפי 2.32). אולם, לכל בסיס של V יש n איברים, ולכן במקרה זה פעולת ההרחבה טריוויאלית, כלומר, לא נוסיף כלל איברים ל־ $v_1, \ldots, v_n$. לפיכך $v_1, \ldots, v_n$ בסיס של V כפי שרצינו להוכיח. ■

המסקנה הבאה היא תוצאה שימושית של המסקנה הקודמת.

2.39 תת־מרחב בממד מרבי שווה למרחב כולו

נניח ש־ V מרחב מממד סופי ו־ U הוא תת־מרחב של V כך ש־ $\dim U = \dim V$. אזי $U = V$.

הוכחה יהי $u_1, \ldots, u_n$ בסיס של U. לפיכך $n = \dim U$, ולפי ההנחה נתון גם ש־ $n = \dim V$. לפיכך, $u_1, \ldots, u_n$ רשימה בלתי תלויה ליניארית של וקטורים ב־ V (שכן היא בסיס של U) ואורכה $\dim V$. לפי 2.38 נקבל ש־ $u_1, \ldots, u_n$ היא בסיס של V. בפרט, כל וקטור ב־ V הוא צירוף ליניארי של $u_1, \ldots, u_n$. לפיכך $U = V$. ■

2.40 example: *a basis of* $\mathbf{F}^2$

Consider the list $(5,7), (4,3)$ of vectors in $\mathbf{F}^2$. This list of length two is linearly independent in $\mathbf{F}^2$ (because neither vector is a scalar multiple of the other). Note that $\mathbf{F}^2$ has dimension two. Thus 2.38 implies that the linearly independent list $(5,7), (4,3)$ of length two is a basis of $\mathbf{F}^2$ (we do not need to bother checking that it spans $\mathbf{F}^2$).

2.41 example: *a basis of a subspace of* $\mathcal{P}_3(\mathbf{R})$

Let U be the subspace of $\mathcal{P}_3(\mathbf{R})$ defined by

$$U = \{p \in \mathcal{P}_3(\mathbf{R}) : p'(5) = 0\}.$$

To find a basis of U, first note that each of the polynomials 1, $(x-5)^2$, and $(x-5)^3$ is in U.

Suppose $a, b, c \in \mathbf{R}$ and

$$a + b(x-5)^2 + c(x-5)^3 = 0$$

for every $x \in \mathbf{R}$. Without explicitly expanding the left side of the equation above, we can see that the left side has a cx^3 term. Because the right side has no x^3 term, this implies that $c = 0$. Because $c = 0$, we see that the left side has a bx^2 term, which implies that $b = 0$. Because $b = c = 0$, we can also conclude that $a = 0$. Thus the equation above implies that $a = b = c = 0$. Hence the list $1, (x-5)^2, (x-5)^3$ is linearly independent in U. Thus $3 \le \dim U$. Hence

$$3 \le \dim U \le \dim \mathcal{P}_3(\mathbf{R}) = 4,$$

where we have used 2.37.

The polynomial x is not in U because its derivative is the constant function 1. Thus $U \ne \mathcal{P}_3(\mathbf{R})$. Hence $\dim U \ne 4$ (by 2.39). The inequality above now implies that $\dim U = 3$. Thus the linearly independent list $1, (x-5)^2, (x-5)^3$ in U has length $\dim U$ and hence is a basis of U (by 2.38).

Now we prove that a spanning list of the right length is a basis.

2.42 *spanning list of the right length is a basis*

Suppose V is finite-dimensional. Then every list of vectors in V that spans V and has length $\dim V$ is a basis of V.

Proof Suppose $\dim V = n$ and $v_1, \dots, v_n$ spans V. The list $v_1, \dots, v_n$ can be reduced to a basis of V (by 2.30). However, every basis of V has length n, so in this case the reduction is the trivial one, meaning that no elements are deleted from $v_1, \dots, v_n$. Thus $v_1, \dots, v_n$ is a basis of V, as desired. ∎

2.40 דוגמה: בסיס של $\mathbf{F}^2$

ניקח כדוגמה את רשימת הווקטורים $(5,7),(4,3)$ ב־ $\mathbf{F}^2$. רשימה זו בת שני וקטורים היא בלתי תלויה ליניארית ב־ $\mathbf{F}^2$ (שכן אף אחד מהווקטורים אינו כפולת סקלר של האחר). מאחר שממד $\mathbf{F}^2$ הוא 2, לפיכך מ־ 2.38 נסיק שהרשימה הבלתי תלויה ליניארית, בת שני האיברים $(5,7),(4,3)$, היא בסיס של $\mathbf{F}^2$ (ואין צורך לבדוק בנוסף שהיא גם פורסת את $\mathbf{F}^2$).

2.41 דוגמה: בסיס של תת־מרחב של $\mathcal{P}_3(\mathbf{R})$

יהי U תת־מרחב של $\mathcal{P}_3(\mathbf{R})$ המוגדר על ידי

$$U = \{p \in \mathcal{P}_3(\mathbf{R}) : p'(5) = 0\}.$$

למציאת בסיס של U, נשים לב תחילה, שכל אחד מהפולינומים 1, $(x-5)^2$, ו־ $(x-5)^3$ ב־ U.

נניח ש־ $a, b, c \in \mathbf{R}$ וגם

$$a + b(x-5)^2 + c(x-5)^3 = 0$$

לכל $x \in \mathbf{R}$. מבלי שנפתח סוגריים באגף השמאלי של המשוואה, אנו רואים כי באגף השמאלי יש איבר של cx^3. מאחר שבאגף הימני אין איבר עם x^3, מכאן נובע ש־ $c = 0$. מאחר ש־ $c = 0$, האיבר היחיד עם x^2 באגף השמאלי הוא bx^2, ומכאן ש־ $b = 0$. מאחר ש־ $b = c = 0$, נוכל להסיק שגם $a = 0$. לפיכך מהמשוואה אנו מסיקים ש־ $a = b = c = 0$. לכן הרשימה $1, (x-5)^2, (x-5)^3$ בלתי תלויה ליניארית ב־ U. לפיכך $\dim U \geq 3$. לכן

$$3 \leq \dim U \leq \dim \mathcal{P}_3(\mathbf{R}) = 4,$$

כאשר השתמשנו ב־ 2.37.

הפולינום x אינו ב־ U מאחר שהנגזרת שלו היא הפונקציה הקבועה 1. לפיכך $U \neq \mathcal{P}_3(\mathbf{R})$. לכן $\dim U \neq 4$ (לפי 2.39). מהאי־שוויון לעיל אנו מסיקים ש־ $\dim U = 3$. לפיכך, מאחר ואורך הרשימה הבלתי תלויה ליניארית $1, (x-5)^2, (x-5)^3$ ב־ U הוא $\dim U$, היא בסיס של U (לפי 2.38).

כעת נוכיח שרשימה פורסת באורך מתאים היא בסיס.

2.42 רשימה פורסת באורך מתאים היא בסיס

נניח ש־ V מרחב וקטורי מממד סופי. אזי, כל רשימת וקטורים ב־ V הפורסת את V ואורכה $\dim V$ היא בסיס של V.

הוכחה נניח ש־ $\dim V = n$ וגם $v_1, \ldots, v_n$ פורסת את V. נוכל לצמצם את הרשימה $v_1, \ldots, v_n$ לכדי בסיס של V (לפי 2.30). אולם בכל בסיס של V יש n איברים, כך שבמקרה זה הצמצום הוא טריוויאלי, כלומר לא נסיר אף וקטור מ־ $v_1, \ldots, v_n$. לפיכך, $v_1, \ldots, v_n$ בסיס של V, כפי שרצינו להוכיח. ■

The next result gives a formula for the dimension of the sum of two subspaces of a finite-dimensional vector space. This formula is analogous to a familiar counting formula: the number of elements in the union of two finite sets equals the number of elements in the first set, plus the number of elements in the second set, minus the number of elements in the intersection of the two sets.

2.43 *dimension of a sum*

If V_1 and V_2 are subspaces of a finite-dimensional vector space, then

$$\dim(V_1 + V_2) = \dim V_1 + \dim V_2 - \dim(V_1 \cap V_2).$$

Proof Let $v_1, \dots, v_m$ be a basis of $V_1 \cap V_2$; thus $\dim(V_1 \cap V_2) = m$. Because $v_1, \dots, v_m$ is a basis of $V_1 \cap V_2$, it is linearly independent in V_1. Hence this list can be extended to a basis $v_1, \dots, v_m, u_1, \dots, u_j$ of V_1 (by 2.32). Thus $\dim V_1 = m + j$. Also extend $v_1, \dots, v_m$ to a basis $v_1, \dots, v_m, w_1, \dots, w_k$ of V_2; thus $\dim V_2 = m + k$.

We will show that

2.44 $$v_1, \dots, v_m, u_1, \dots, u_j, w_1, \dots, w_k$$

is a basis of $V_1 + V_2$. This will complete the proof, because then we will have

$$\begin{aligned}\dim(V_1 + V_2) &= m + j + k \\ &= (m + j) + (m + k) - m \\ &= \dim V_1 + \dim V_2 - \dim(V_1 \cap V_2).\end{aligned}$$

The list 2.44 is contained in $V_1 \cup V_2$ and thus is contained in $V_1 + V_2$. The span of this list contains V_1 and contains V_2 and hence is equal to $V_1 + V_2$. Thus to show that 2.44 is a basis of $V_1 + V_2$ we only need to show that it is linearly independent.

To prove that 2.44 is linearly independent, suppose

$$a_1 v_1 + \cdots + a_m v_m + b_1 u_1 + \cdots + b_j u_j + c_1 w_1 + \cdots + c_k w_k = 0,$$

where all the a's, b's, and c's are scalars. We need to prove that all the a's, b's, and c's equal 0. The equation above can be rewritten as

2.45 $$c_1 w_1 + \cdots + c_k w_k = -a_1 v_1 - \cdots - a_m v_m - b_1 u_1 - \cdots - b_j u_j,$$

which shows that $c_1 w_1 + \cdots + c_k w_k \in V_1$. All the w's are in V_2, so this implies that $c_1 w_1 + \cdots + c_k w_k \in V_1 \cap V_2$. Because $v_1, \dots, v_m$ is a basis of $V_1 \cap V_2$, we have

$$c_1 w_1 + \cdots + c_k w_k = d_1 v_1 + \cdots + d_m v_m$$

for some scalars $d_1, \dots, d_m$. But $v_1, \dots, v_m, w_1, \dots, w_k$ is linearly independent, so the last equation implies that all the c's (and d's) equal 0. Thus 2.45 becomes the equation

$$a_1 v_1 + \cdots + a_m v_m + b_1 u_1 + \cdots + b_j u_j = 0.$$

Because the list $v_1, \dots, v_m, u_1, \dots, u_j$ is linearly independent, this equation implies that all the a's and b's are 0, completing the proof. ∎

המסקנה הבאה מספקת נוסחה לחישוב ממד הסכום של שני תתי־מרחב של מרחב וקטורי מממד סופי. נוסחה זו אנלוגית לנוסחת ספירה מוכרת: מספר האיברים באיחוד של שתי קבוצות סופיות שווה למספר האיברים בקבוצה הראשונה, בתוספת מספר האיברים בקבוצה השנייה, ובחיסור מספר האיברים שבחיתוך שתי הקבוצות.

2.43 ממד של סכום

אם V_1 ו־ V_2 תתי־מרחב של מרחב וקטורי מממד סופי, אזי
$$\dim(V_1 + V_2) = \dim V_1 + \dim V_2 - \dim(V_1 \cap V_2).$$

הוכחה יהי $v_1, \dots, v_m$ בסיס של $V_1 \cap V_2$; לפיכך, $\dim(V_1 \cap V_2) = m$. מאחר ש־ $v_1, \dots, v_m$ בסיס של $V_1 \cap V_2$, הרשימה בלתי תלויה ליניארית ב־ V_1. לכן נרחיב את הרשימה לבסיס של V_1 (לפי 2.32) באופן הבא: $v_1, \dots, v_m, u_1, \dots, u_j$. לפיכך $\dim V_1 = m + j$. בנוסף, נרחיב את $v_1, \dots, v_m$ לבסיס $v_1, \dots, v_m, w_1, \dots, w_k$ של V_2; לפיכך $\dim V_2 = m + k$.

אנו נראה ש־

2.44
$$v_1, \dots, v_m, u_1, \dots, u_j, w_1, \dots, w_k$$

בסיס של $V_1 + V_2$. בכך תושלם ההוכחה, שכן אז נקבל
$$\begin{aligned}\dim(V_1 + V_2) &= m + j + k \\ &= (m + j) + (m + k) - m \\ &= \dim V_1 + \dim V_2 - \dim(V_1 \cap V_2).\end{aligned}$$

הרשימה 2.44 מוכלת ב־ $V_1 \cup V_2$ ולפיכך מוכלת גם ב־ $V_1 + V_2$. הפריסה של הרשימה מכילה את V_1 ומכילה את V_2 ולכן שווה ל־ $V_1 + V_2$. לפיכך כדי להראות ש־ 2.44 בסיס של $V_1 + V_2$ כל שעלינו להראות הוא שהיא בלתי תלויה ליניארית.

כדי להוכיח ש־ 2.44 בלתי תלויה ליניארית, נניח ש־
$$a_1 v_1 + \cdots + a_m v_m + b_1 u_1 + \cdots + b_j u_j + c_1 w_1 + \cdots + c_k w_k = 0,$$
כאשר כל ה־ a־ים, b־ים, ו־ c־ים סקלרים. עלינו להראות שכל ה־ a־ים, b־ים, ו־ c־ים שווים 0. נשכתב את המשוואה באופן הבא:

2.45
$$c_1 w_1 + \cdots + c_k w_k = -a_1 v_1 - \cdots - a_m v_m - b_1 u_1 - \cdots - b_j u_j,$$

מה שמראה כי $c_1 w_1 + \cdots + c_k w_k \in V_1$. מאחר שכל ה־ w־ים ב־ V_2, נסיק ש־ $c_1 w_1 + \cdots + c_k w_k \in V_1 \cap V_2$. מאחר ש־ $v_1, \dots, v_m$ בסיס של $V_1 \cap V_2$, נקבל
$$c_1 w_1 + \cdots + c_k w_k = d_1 v_1 + \cdots + d_m v_m$$
עבור $d_1, \dots, d_m$ סקלרים כלשהם. אולם מאחר ש־ $w_1, \dots, w_k$ בלתי תלויה ליניארית, מהמשוואה האחרונה נסיק כי כל ה־ c־ים (וה־ d־ים) שווים 0. לפיכך הופכת 2.45 למשוואה
$$a_1 v_1 + \cdots + a_m v_m + b_1 u_1 + \cdots + b_j u_j = 0.$$
מאחר שהרשימה $v_1, \dots, v_m, u_1, \dots, u_j$ בלתי תלויה ליניארית, נסיק כי כל ה־ a־ים וה־ b־ים שווים גם הם 0 ובזאת השלמנו את ההוכחה. ∎

For S a finite set, let $\#S$ denote the number of elements of S. The table below compares finite sets with finite-dimensional vector spaces, showing the analogy between $\#S$ (for sets) and $\dim V$ (for vector spaces), as well as the analogy between unions of subsets (in the context of sets) and sums of subspaces (in the context of vector spaces).

sets	vector spaces
S is a finite set	V is a finite-dimensional vector space
$\#S$	$\dim V$
for subsets S_1, S_2 of S, the union $S_1 \cup S_2$ is the smallest subset of S containing S_1 and S_2	for subspaces V_1, V_2 of V, the sum V_1+V_2 is the smallest subspace of V containing V_1 and V_2
$\#(S_1 \cup S_2)$ $= \#S_1 + \#S_2 - \#(S_1 \cap S_2)$	$\dim(V_1 + V_2)$ $= \dim V_1 + \dim V_2 - \dim(V_1 \cap V_2)$
$\#(S_1 \cup S_2) = \#S_1 + \#S_2$ $\iff S_1 \cap S_2 = \emptyset$	$\dim(V_1 + V_2) = \dim V_1 + \dim V_2$ $\iff V_1 \cap V_2 = \{0\}$
$S_1 \cup \cdots \cup S_m$ is a disjoint union $\iff$ $\#(S_1 \cup \cdots \cup S_m) = \#S_1 + \cdots + \#S_m$	$V_1 + \cdots + V_m$ is a direct sum $\iff$ $\dim(V_1 + \cdots + V_m)$ $= \dim V_1 + \cdots + \dim V_m$

The last row above focuses on the analogy between disjoint unions (for sets) and direct sums (for vector spaces). The proof of the result in the last box above will be given in 3.94.

You should be able to find results about sets that correspond, via analogy, to the results about vector spaces in Exercises 12 through 18.

Exercises 2C

1 Show that the subspaces of $\mathbf{R}^2$ are precisely $\{0\}$, all lines in $\mathbf{R}^2$ containing the origin, and $\mathbf{R}^2$.

2 Show that the subspaces of $\mathbf{R}^3$ are precisely $\{0\}$, all lines in $\mathbf{R}^3$ containing the origin, all planes in $\mathbf{R}^3$ containing the origin, and $\mathbf{R}^3$.

3 (a) Let $U = \{p \in \mathcal{P}_4(\mathbf{F}) : p(6) = 0\}$. Find a basis of U.
(b) Extend the basis in (a) to a basis of $\mathcal{P}_4(\mathbf{F})$.
(c) Find a subspace W of $\mathcal{P}_4(\mathbf{F})$ such that $\mathcal{P}_4(\mathbf{F}) = U \oplus W$.

4 (a) Let $U = \{p \in \mathcal{P}_4(\mathbf{R}) : p''(6) = 0\}$. Find a basis of U.
(b) Extend the basis in (a) to a basis of $\mathcal{P}_4(\mathbf{R})$.
(c) Find a subspace W of $\mathcal{P}_4(\mathbf{R})$ such that $\mathcal{P}_4(\mathbf{R}) = U \oplus W$.

5 (a) Let $U = \{p \in \mathcal{P}_4(\mathbf{F}) : p(2) = p(5)\}$. Find a basis of U.
(b) Extend the basis in (a) to a basis of $\mathcal{P}_4(\mathbf{F})$.
(c) Find a subspace W of $\mathcal{P}_4(\mathbf{F})$ such that $\mathcal{P}_4(\mathbf{F}) = U \oplus W$.

עבור S קבוצה סופית נסמן ב־ $\#S$ את מספר איבריה של S. הטבלה להלן משווה קבוצות סופיות ומרחבים מממד סופי, מציגה את האנלוגיה בין $\#S$ (עבור קבוצות) ו־ $\dim V$ (עבור מרחבים וקטוריים), וכן את האנלוגיה בין איחוד תתי־קבוצה (בהקשר של קבוצות) וסכום תתי־מרחב (בהקשר של מרחבים וקטוריים).

מרחבים וקטוריים	קבוצות
V מרחב וקטורי מממד סופי	S קבוצה סופית
$\dim V$	$\#S$
עבור תתי־המרחב V_1, V_2 של V, הסכום $V_1 + V_2$ הוא התת־מרחב הקטן ביותר של V המכיל את V_1 ו־ V_2	עבור תתי־קבוצה S_1, S_2 של S, האיחוד $S_1 \cup S_2$ הוא התת־קבוצה הקטנה ביותר של S המכילה את S_1 ו־ S_2
$\dim(V_1 + V_2)$ $= \dim V_1 + \dim V_2 - \dim(V_1 \cap V_2)$	$\#(S_1 \cup S_2)$ $= \#S_1 + \#S_2 - \#(S_1 \cap S_2)$
$\dim(V_1 + V_2) = \dim V_1 + \dim V_2$ $\iff V_1 \cap V_2 = \{0\}$	$\#(S_1 \cup S_2) = \#S_1 + \#S_2$ $\iff S_1 \cap S_2 = \emptyset$
$V_1 + \cdots + V_m$ הוא סכום ישר $\iff$ $\dim(V_1 + \cdots + V_m)$ $= \dim V_1 + \cdots + \dim V_m$	$S_1 \cup \cdots \cup S_m$ הוא איחוד זר $\iff$ $\#(S_1 \cup \cdots \cup S_m) = \#S_1 + \cdots + \#S_m$

השורה האחרונה בטבלה מתארת את האנלוגיה בין איחוד זר עבור קבוצות וסכום ישר עבור מרחבים וקטוריים. ההוכחה למסקנה שבשורה האחרונה בטבלה תובא ב־ 3.94.

תוכלו למצוא מסקנות על קבוצות, המותאמות באנלוגיה למסקנות על מרחבים וקטוריים המופיעות בתרגילים 12 עד 18.

תרגילים בג

1 הראו שתתי־המרחב של $\mathbf{R}^2$ הם אך ורק $\{0\}$, כל הישרים ב־ $\mathbf{R}^2$ העוברים דרך הראשית, ו־ $\mathbf{R}^2$.

2 הראו כי תתי־המרחב של $\mathbf{R}^3$ הם אך ורק $\{0\}$, כל הישרים ב־ $\mathbf{R}^3$ העוברים דרך הראשית, כל המישורים ב־ $\mathbf{R}^3$ העוברים דרך הראשית, ו־ $\mathbf{R}^3$.

3 א. יהי $U = \{p \in \mathcal{P}_4(\mathbf{F}) : p(6) = 0\}$. מצאו בסיס של U.
ב. השלימו את הבסיס ב־ (א) לבסיס של $\mathcal{P}_4(\mathbf{F})$.
ג. מצאו תת־מרחב W של $\mathcal{P}_4(\mathbf{F})$ כך ש־ $\mathcal{P}_4(\mathbf{F}) = U \oplus W$.

4 א. יהי $U = \{p \in \mathcal{P}_4(\mathbf{R}) : p''(6) = 0\}$. מצאו בסיס של U.
ב. השלימו את הבסיס ב־ (א) לבסיס של $\mathcal{P}_4(\mathbf{R})$.
ג. מצאו תת־מרחב W של $\mathcal{P}_4(\mathbf{R})$ כך ש־ $\mathcal{P}_4(\mathbf{R}) = U \oplus W$.

5 א. יהי $U = \{p \in \mathcal{P}_4(\mathbf{F}) : p(2) = p(5)\}$. מצאו בסיס של U.
ב. השלימו את הבסיס ב־ (א) לבסיס של $\mathcal{P}_4(\mathbf{F})$.
ג. מצאו תת־מרחב W של $\mathcal{P}_4(\mathbf{F})$ כך ש־ $\mathcal{P}_4(\mathbf{F}) = U \oplus W$.

6 (a) Let $U = \{p \in \mathcal{P}_4(\mathbf{F}) : p(2) = p(5) = p(6)\}$. Find a basis of U.
(b) Extend the basis in (a) to a basis of $\mathcal{P}_4(\mathbf{F})$.
(c) Find a subspace W of $\mathcal{P}_4(\mathbf{F})$ such that $\mathcal{P}_4(\mathbf{F}) = U \oplus W$.

7 (a) Let $U = \left\{p \in \mathcal{P}_4(\mathbf{R}) : \int_{-1}^{1} p = 0\right\}$. Find a basis of U.
(b) Extend the basis in (a) to a basis of $\mathcal{P}_4(\mathbf{R})$.
(c) Find a subspace W of $\mathcal{P}_4(\mathbf{R})$ such that $\mathcal{P}_4(\mathbf{R}) = U \oplus W$.

8 Suppose $v_1, \dots, v_m$ is linearly independent in V and $w \in V$. Prove that

$$\dim \operatorname{span}(v_1 + w, \dots, v_m + w) \geq m - 1.$$

9 Suppose m is a positive integer and $p_0, p_1, \dots, p_m \in \mathcal{P}(\mathbf{F})$ are such that each p_k has degree k. Prove that $p_0, p_1, \dots, p_m$ is a basis of $\mathcal{P}_m(\mathbf{F})$.

10 Suppose m is a positive integer. For $0 \leq k \leq m$, let

$$p_k(x) = x^k(1 - x)^{m-k}.$$

Show that $p_0, \dots, p_m$ is a basis of $\mathcal{P}_m(\mathbf{F})$.

The basis in this exercise leads to what are called ***Bernstein polynomials****. You can do a web search to learn how Bernstein polynomials are used to approximate continuous functions on* $[0, 1]$.

11 Suppose U and W are both four-dimensional subspaces of $\mathbf{C}^6$. Prove that there exist two vectors in $U \cap W$ such that neither of these vectors is a scalar multiple of the other.

12 Suppose that U and W are subspaces of $\mathbf{R}^8$ such that $\dim U = 3$, $\dim W = 5$, and $U + W = \mathbf{R}^8$. Prove that $\mathbf{R}^8 = U \oplus W$.

13 Suppose U and W are both five-dimensional subspaces of $\mathbf{R}^9$. Prove that $U \cap W \neq \{0\}$.

14 Suppose V is a ten-dimensional vector space and V_1, V_2, V_3 are subspaces of V with $\dim V_1 = \dim V_2 = \dim V_3 = 7$. Prove that $V_1 \cap V_2 \cap V_3 \neq \{0\}$.

15 Suppose V is finite-dimensional and V_1, V_2, V_3 are subspaces of V with $\dim V_1 + \dim V_2 + \dim V_3 > 2 \dim V$. Prove that $V_1 \cap V_2 \cap V_3 \neq \{0\}$.

16 Suppose V is finite-dimensional and U is a subspace of V with $U \neq V$. Let $n = \dim V$ and $m = \dim U$. Prove that there exist $n - m$ subspaces of V, each of dimension $n - 1$, whose intersection equals U.

17 Suppose that $V_1, \dots, V_m$ are finite-dimensional subspaces of V. Prove that $V_1 + \cdots + V_m$ is finite-dimensional and

$$\dim(V_1 + \cdots + V_m) \leq \dim V_1 + \cdots + \dim V_m.$$

The inequality above is an equality if and only if $V_1 + \cdots + V_m$ *is a direct sum, as will be shown in 3.94.*

6 א. יהי $U = \{p \in \mathcal{P}_4(\mathbf{F}) : p(2) = p(5) = p(6)\}$. מצאו בסיס של U.
ב. השלימו את הבסיס ב־ (א) לבסיס של $\mathcal{P}_4(\mathbf{F})$.
ג. מצאו תת־מרחב W של $\mathcal{P}_4(\mathbf{F})$ כך ש־ $\mathcal{P}_4(\mathbf{F}) = U \oplus W$.

7 א. יהי $U = \left\{p \in \mathcal{P}_4(\mathbf{R}) : \int_{-1}^{1} p = 0\right\}$. מצאו בסיס של U.
ב. השלימו את הבסיס ב־ (א) לבסיס של $\mathcal{P}_4(\mathbf{R})$.
ג. מצאו תת־מרחב W של $\mathcal{P}_4(\mathbf{R})$ כך ש־ $\mathcal{P}_4(\mathbf{R}) = U \oplus W$.

8 נניח ש־ $v_1, \dots, v_m$ בלתי תלויה ליניארית ב־ V ו־ $w \in V$. הוכיחו כי
$$\dim \operatorname{span}(v_1 + w, \dots, v_m + w) \geq m - 1.$$

9 נניח ש־ m מספר שלם חיובי, ו־ $p_0, p_1, \dots, p_m \in \mathcal{P}(\mathbf{F})$ כך שכל p_k הוא ממעלה k. הוכיחו כי $p_0, p_1, \dots, p_m$ בסיס של $\mathcal{P}_m(\mathbf{F})$.

10 נניח ש־ m מספר שלם חיובי. עבור $0 \leq k \leq m$, יהי
$$p_k(x) = x^k(1-x)^{m-k}.$$
הראו ש־ $p_0, \dots, p_m$ בסיס של $\mathcal{P}_m(\mathbf{F})$.

הבסיס בתרגיל זה מוביל למה שקרוי **פולינומי ברנשטיין**. תוכלו לחפש מידע על האופן שבו פולינומי ברנשטיין משמשים לקירוב פונקציות רציפות מעל $[0,1]$.

11 נניח ש־ U ו־ W שניהם תתי־מרחב מממד 4 של $\mathbf{C}^6$. הוכיחו כי קיימים שני וקטורים ב־ $U \cap W$ כך שאף אחד מהווקטורים הללו אינו כפולת סקלר של האחר.

12 נניח ש־ U ו־ W תתי־מרחב של $\mathbf{R}^8$ כך ש־ $\dim W = 5$, $\dim U = 3$, ו־ $U + W = \mathbf{R}^8$. הוכיחו כי $\mathbf{R}^8 = U \oplus W$.

13 נניח ש־ U ו־ W שניהם תתי־מרחב מממד 5 של $\mathbf{R}^9$. הוכיחו כי $U \cap W \neq \{0\}$.

14 נניח ש־ V הוא מרחב וקטורי מממד 10 ו־ V_1, V_2, V_3 תתי־מרחב של V, כאשר $\dim V_1 = \dim V_2 = \dim V_3 = 7$. הוכיחו כי $V_1 \cap V_2 \cap V_3 \neq \{0\}$.

15 נניח ש־ V מרחב וקטורי מממד סופי ו־ V_1, V_2, V_3 תתי־מרחב של V כאשר $\dim V_1 + \dim V_2 + \dim V_3 > 2 \dim V$. הוכיחו כי $V_1 \cap V_2 \cap V_3 \neq \{0\}$.

16 נניח ש־ V מרחב וקטורי מממד סופי ו־ U תת־מרחב של V כאשר $U \neq V$. נסמן $n = \dim V$ ו־ $m = \dim U$. הוכיחו כי קיימים $n - m$ תתי־מרחב של V, כל אחד מממד $n - 1$, והחיתוך שלהם שווה ל־ U.

17 נניח ש־ $V_1, \dots, V_m$ תתי־מרחב מממד סופי של V. הוכיחו כי $V_1 + \cdots + V_m$ הוא מממד סופי, וגם
$$\dim(V_1 + \cdots + V_m) \leq \dim V_1 + \cdots + \dim V_m.$$

אי־שוויון זה הוא שוויון אס־ורק־אס $V_1 + \cdots + V_m$ הוא סכום ישר, כפי שנראה ב־ 3.94.

18 Suppose V is finite-dimensional, with $\dim V = n \geq 1$. Prove that there exist one-dimensional subspaces $V_1, \ldots, V_n$ of V such that

$$V = V_1 \oplus \cdots \oplus V_n.$$

19 Explain why you might guess, motivated by analogy with the formula for the number of elements in the union of three finite sets, that if V_1, V_2, V_3 are subspaces of a finite-dimensional vector space, then

$$\begin{aligned}\dim(V_1 + V_2 + V_3)&\\ &= \dim V_1 + \dim V_2 + \dim V_3\\ &\quad - \dim(V_1 \cap V_2) - \dim(V_1 \cap V_3) - \dim(V_2 \cap V_3)\\ &\quad + \dim(V_1 \cap V_2 \cap V_3).\end{aligned}$$

Then either prove the formula above or give a counterexample.

20 Prove that if V_1, V_2, and V_3 are subspaces of a finite-dimensional vector space, then

$$\begin{aligned}&\dim(V_1 + V_2 + V_3)\\ &= \dim V_1 + \dim V_2 + \dim V_3\\ &\quad - \frac{\dim(V_1 \cap V_2) + \dim(V_1 \cap V_3) + \dim(V_2 \cap V_3)}{3}\\ &\quad - \frac{\dim((V_1+V_2)\cap V_3) + \dim((V_1+V_3)\cap V_2) + \dim((V_2+V_3)\cap V_1)}{3}.\end{aligned}$$

The formula above may seem strange because the right side does not look like an integer.

I at once gave up my former occupations, set down natural history and all its progeny as a deformed and abortive creation, and entertained the greatest disdain for a would-be science which could never even step within the threshold of real knowledge. In this mood I betook myself to the mathematics and the branches of study appertaining to that science as being built upon secure foundations, and so worthy of my consideration.

—*Frankenstein*, Mary Wollstonecraft Shelley

18 נניח ש־ V מרחב וקטורי מממד סופי, עם $\dim V = n \geq 1$. הוכיחו כי קיימים תתי־מרחב חד־ממדיים $V_1, \ldots, V_n$ של V כך ש־

$$V = V_1 \oplus \cdots \oplus V_n.$$

19 הסבירו מדוע ניתן לשער, בהשראת האנלוגיה לנוסחה למספר האיברים באיחוד של שלוש קבוצות סופיות, שאם V_1, V_2, V_3 הם תתי־מרחב של מרחב וקטורי מממד סופי, אזי

$$\begin{aligned}\dim(V_1 + V_2 + V_3) &= \dim V_1 + \dim V_2 + \dim V_3 \\ &\quad - \dim(V_1 \cap V_2) - \dim(V_1 \cap V_3) - \dim(V_2 \cap V_3) \\ &\quad + \dim(V_1 \cap V_2 \cap V_3).\end{aligned}$$

ואז הוכיחו את הנוסחה או הביאו דוגמה נגדית.

20 הוכיחו כי אם V_1, V_2, ו־ V_3 תתי־מרחב של מרחב וקטורי מממד סופי, אזי מתקיים

$$\begin{aligned}&\dim(V_1 + V_2 + V_3) \\ &= \dim V_1 + \dim V_2 + \dim V_3 \\ &\quad - \frac{\dim(V_1 \cap V_2) + \dim(V_1 \cap V_3) + \dim(V_2 \cap V_3)}{3} \\ &\quad - \frac{\dim((V_1+V_2)\cap V_3) + \dim((V_1+V_3)\cap V_2) + \dim((V_2+V_3)! \cap V_1)}{3}.\end{aligned}$$

ייתכן שהנוסחה לעיל נראית מוזרה מאחר שהאגף הימני אינו נראה כמספר שלם.

בְּאִבְחָה אַחַת נָטַשְׁתִּי אֶת עִסּוּקַי הַקּוֹדְמִים,
אֶל תְּהוֹם הִשְׁלַכְתִּי תּוֹלְדֹת הַטֶּבַע וְצֶאֱצָאָיו כִּבְאוּשֵׁי נְפֹלֶת עִוּוּיִים,
וּבְקִרְבִּי, בּוּז עָמוֹק אֶל הַמִּתְחַזֶּה לְמַדָּע,
אֲשֶׁר רַגְלוֹ לֹא דָּרְכָה מֵעוֹלָם עַל מִפְתַּן שַׁעֲרֵי הַחָכְמָה.
חָדוּר בְּתוֹדָעָה זוֹ,
הֶעְתַּקְתִּי עַצְמִי אֶל מַעֲמַקֵּי הַמָּתֵמָטִיקָה וְשָׂרִיגֶיהָ,
כִּי בְּסֶלַע אֵיתָנִים חֲצוּבִים שָׁרָשֶׁיהָ,
וְעַד אֲבַק כּוֹכָבִים יַגִּיעוּ חִשּׁוּבֶיהָ.

—פְּרַנְקֶנְשְׁטַיְן, מַרִי וּולסֵטוֹנקרַאפט שֶׁלִי

אלגברה ליניארית בדרך הישר, מהדורה רביעית, מחבר: שלדון אקסלר
Linear Algebra Done Right, fourth edition, by Sheldon Axler

Chapter 3

Linear Maps

So far our attention has focused on vector spaces. No one gets excited about vector spaces. The interesting part of linear algebra is the subject to which we now turn—linear maps.

We will frequently use the powerful fundamental theorem of linear maps, which states that the dimension of the domain of a linear map equals the dimension of the subspace that gets sent to 0 plus the dimension of the range. This will imply the striking result that a linear map from a finite-dimensional vector space to itself is one-to-one if and only if its range is the whole space.

A major concept that we will introduce in this chapter is the matrix associated with a linear map and with a basis of the domain space and a basis of the target space. This correspondence between linear maps and matrices provides much insight into key aspects of linear algebra.

This chapter concludes by introducing product, quotient, and dual spaces.

In this chapter we will need additional vector spaces, which we call U and W, in addition to V. Thus our standing assumptions are now as follows.

standing assumptions for this chapter

- $\mathbf{F}$ denotes $\mathbf{R}$ or $\mathbf{C}$.
- U, V, and W denote vector spaces over $\mathbf{F}$.

Stefan Schäfer CC BY-SA

The twelfth-century Dankwarderode Castle in Brunswick (Braunschweig), where Carl Friedrich Gauss (1777–1855) was born and grew up. In 1809 Gauss published a method for solving systems of linear equations. This method, now called Gaussian elimination, was used in a Chinese book written over 1600 years earlier.

פרק ג

העתקות ליניאריות

עד כה תשומת הלב שלנו התמקדה במרחבים וקטוריים. איש אינו מתרגש ממרחבים וקטוריים. החלק המרתק באלגברה ליניארית הוא הנושא שאליו נפנה כעת—העתקות ליניאריות.

לעיתים קרובות נשתמש במשפט היסודי רב־העוצמה של ההעתקות הליניאריות, הקובע כי ממד התחום של העתקה ליניארית שווה לממד תת־המרחב הנשלח ל־ 0 בתוספת ממד הטווח. קביעה זו גוררת את המסקנה המפתיעה שהעתקה ליניארית ממרחב וקטורי מממד סופי לעצמו היא חד־חד־ערכית אם־ורק־אם טווח ההעתקה הוא המרחב כולו.

אחד הרעיונות המרכזיים שנכיר בפרק זה הוא המטריצה המשויכת להעתקה ליניארית, לבסיס מרחב התחום ולבסיס מרחב היעד. התאמה זו בין העתקות ליניאריות ומטריצות מספקת תובנות רבות על היבטים מרכזיים באלגברה ליניארית.

את פרק זה נחתום בהיכרות עם מכפלות ומנות של מרחבים וקטוריים, ומרחבים דואליים.

בפרק זה נזדקק למרחבים וקטוריים נוספים, שלהם נקרא U ו־ W בנוסף ל־ V. לפיכך הנחות המוצא שלנו הן כדלהלן.

הנחות המוצא לפרק זה

- $\mathbf{F}$ מסמן $\mathbf{R}$ או $\mathbf{C}$.
- U, V ו־ W מסמנים מרחבים וקטוריים מעל $\mathbf{F}$.

Stefan Schäfer CC BY-SA

טירת דנקוורדרודה (Dankwarderode) מהמאה השתים־עשרה בבראונשוויג (Braunschweig), היכן שנולד וגדל קרל פרידריך גָאוּס (1855–1777). בשנת 1809 פרסם גאוס שיטה לפתרון מערכות משוואות. שיטה זו, הנקראת כיום שיטת החילוץ של גאוס, הייתה בשימוש בספרים סיניים יותר מ־1600 שנה קודם לכן.

3A Vector Space of Linear Maps

Definition and Examples of Linear Maps

Now we are ready for one of the key definitions in linear algebra.

3.1 definition: *linear map*

A *linear map* from V to W is a function $T : V \to W$ with the following properties.

additivity
$T(u + v) = Tu + Tv$ for all $u, v \in V$.

homogeneity
$T(\lambda v) = \lambda(Tv)$ for all $\lambda \in \mathbf{F}$ and all $v \in V$.

Note that for linear maps we often use the notation Tv as well as the usual function notation $T(v)$.

*Some mathematicians use the phrase **linear transformation**, which means the same as linear map.*

3.2 notation: $\mathcal{L}(V, W)$, $\mathcal{L}(V)$

- The set of linear maps from V to W is denoted by $\mathcal{L}(V, W)$.
- The set of linear maps from V to V is denoted by $\mathcal{L}(V)$. In other words, $\mathcal{L}(V) = \mathcal{L}(V, V)$.

Let's look at some examples of linear maps. Make sure you verify that each of the functions defined in the next example is indeed a linear map:

3.3 example: *linear maps*

zero
In addition to its other uses, we let the symbol 0 denote the linear map that takes every element of some vector space to the additive identity of another (or possibly the same) vector space. To be specific, $0 \in \mathcal{L}(V, W)$ is defined by

$$0v = 0.$$

The 0 on the left side of the equation above is a function from V to W, whereas the 0 on the right side is the additive identity in W. As usual, the context should allow you to distinguish between the many uses of the symbol 0.

identity operator
The *identity operator*, denoted by I, is the linear map on some vector space that takes each element to itself. To be specific, $I \in \mathcal{L}(V)$ is defined by

$$Iv = v.$$

ג1 מרחב וקטורי של העתקות ליניאריות

הגדרה ודוגמאות של העתקות ליניאריות

ניגש כעת אל אחת מהגדרות המפתח באלגברה ליניארית.

3.1 הגדרה: העתקה ליניארית

העתקה ליניארית מ־ V ל־ W היא פונקציה $T: V \to W$ המקיימת את התכונות הבאות.

אדיטיביות
$T(u+v) = Tu + Tv$ לכל $u, v \in V$.

הומוגניות
$T(\lambda v) = \lambda(Tv)$ לכל $\lambda \in \mathbf{F}$ ולכל $v \in V$.

נשים לב כי עבור העתקות ליניאריות, לעיתים קרובות נשתמש בסימון Tv במקביל לשימוש ב־ $T(v)$ המקובל לסימון פונקציות.

יש מתמטיקאים המשתמשים במונח **טרנספורמציה ליניארית**, שמשמעותו זהה להעתקה ליניארית.

3.2 סימון: $\mathcal{L}(V, W)$, $\mathcal{L}(V)$

- קבוצת ההעתקות הליניאריות מ־ V ל־ W מסומנת על ידי $\mathcal{L}(V, W)$.
- קבוצת ההעתקות הליניאריות מ־ V ל־ V מסומנת על ידי $\mathcal{L}(V)$. כלומר, $\mathcal{L}(V) = \mathcal{L}(V, V)$.

כעת נראה מספר דוגמאות של העתקות ליניאריות. עליכם לוודא כי כל אחת מהפונקציות המוגדרות בדוגמה הבאה היא אכן העתקה ליניארית.

3.3 דוגמה: העתקות ליניאריות

העתקת האפס
בנוסף לשימושים האחרים, אנו מגדירים את 0 כסימון להעתקה הליניארית המעתיקה כל איבר במרחב וקטורי כלשהו אל איבר הזהות לחיבור של מרחב וקטורי אחר (או לאותו מרחב). ליתר דיוק, $0 \in \mathcal{L}(V, W)$ מוגדר על ידי

$$0v = 0.$$

ה־ 0 באגף השמאלי של המשוואה הוא פונקציה מ־ V ל־ W, בעוד ה־ 0 באגף הימני הוא איבר הזהות לחיבור ב־ W. כהרגלנו, הֶקְשֵׁר אמור לאפשר לכם להבחין בין השימושים הרבים של הסימן 0.

אופרטור הזהות
אופרטור הזהות, המסומן ב־ I, הוא ההעתקה הליניארית על מרחב וקטורי כלשהו המעתיקה כל איבר לעצמו. ליתר דיוק $I \in \mathcal{L}(V)$ מוגדר על ידי

$$Iv = v.$$

differentiation
Define $D \in \mathcal{L}(\mathcal{P}(\mathbf{R}))$ by

$$Dp = p'.$$

The assertion that this function is a linear map is another way of stating a basic result about differentiation: $(f + g)' = f' + g'$ and $(\lambda f)' = \lambda f'$ whenever f, g are differentiable and λ is a constant.

integration
Define $T \in \mathcal{L}(\mathcal{P}(\mathbf{R}), \mathbf{R})$ by

$$Tp = \int_0^1 p.$$

The assertion that this function is linear is another way of stating a basic result about integration: the integral of the sum of two functions equals the sum of the integrals, and the integral of a constant times a function equals the constant times the integral of the function.

multiplication by x^2
Define a linear map $T \in \mathcal{L}(\mathcal{P}(\mathbf{R}))$ by

$$(Tp)(x) = x^2 p(x)$$

for each $x \in \mathbf{R}$.

backward shift
Recall that $\mathbf{F}^\infty$ denotes the vector space of all sequences of elements of $\mathbf{F}$. Define a linear map $T \in \mathcal{L}(\mathbf{F}^\infty)$ by

$$T(x_1, x_2, x_3, \dots) = (x_2, x_3, \dots).$$

from $\mathbf{R}^3$ to $\mathbf{R}^2$
Define a linear map $T \in \mathcal{L}(\mathbf{R}^3, \mathbf{R}^2)$ by

$$T(x, y, z) = (2x - y + 3z, 7x + 5y - 6z).$$

from $\mathbf{F}^n$ to $\mathbf{F}^m$
To generalize the previous example, let m and n be positive integers, let $A_{j,k} \in \mathbf{F}$ for each $j = 1, \dots, m$ and each $k = 1, \dots, n$, and define a linear map $T \in \mathcal{L}(\mathbf{F}^n, \mathbf{F}^m)$ by

$$T(x_1, \dots, x_n) = (A_{1,1}x_1 + \cdots + A_{1,n}x_n, \dots, A_{m,1}x_1 + \cdots + A_{m,n}x_n).$$

Actually every linear map from $\mathbf{F}^n$ to $\mathbf{F}^m$ is of this form.

composition
Fix a polynomial $q \in \mathcal{P}(\mathbf{R})$. Define a linear map $T \in \mathcal{L}(\mathcal{P}(\mathbf{R}))$ by

$$(Tp)(x) = p(q(x)).$$

The existence part of the next result means that we can find a linear map that takes on whatever values we wish on the vectors in a basis. The uniqueness part of the next result means that a linear map is completely determined by its values on a basis.

גזירה

נגדיר העתקה ליניארית $D \in \mathcal{L}(\mathcal{P}(\mathbf{R}))$ על ידי

$$Dp = p'.$$

הקביעה כי פונקציה זו היא ליניארית היא דרך נוספת לנסח את אחת המסקנות הבסיסיות בנושא נגזרות: $(f+g)' = f' + g'$ ו־ $(\lambda f)' = \lambda f'$ כאשר f, g גְּזִירוֹת ו־ λ מספר קבוע.

אינטגרציה

נגדיר העתקה ליניארית $T \in \mathcal{L}(\mathcal{P}(\mathbf{R}), \mathbf{R})$ על ידי

$$Tp = \int_0^1 p.$$

הקביעה כי פונקציה זו היא ליניארית היא דרך נוספת לנסח את אחת המסקנות הבסיסיות בנושא אינטגרלים: האינטגרל של סכום שתי פונקציות שווה לסכום האינטגרלים, והאינטגרל של מכפלת קבוע בפונקציה שווה למכפלת הקבוע באינטגרל של הפונקציה.

הכפלה ב־ x^2

נגדיר העתקה ליניארית $T \in \mathcal{L}(\mathcal{P}(\mathbf{R}))$ על ידי

$$(Tp)(x) = x^2 p(x)$$

לכל $x \in \mathbf{R}$.

הסטה לאחור

נזכיר ש־ $\mathbf{F}^\infty$ מסמן את המרחב הווקטורי של כל הסדרות של איברי $\mathbf{F}$. נגדיר העתקה ליניארית $T \in \mathcal{L}(\mathbf{F}^\infty)$ על ידי

$$T(x_1, x_2, x_3, \dots) = (x_2, x_3, \dots).$$

מ־ $\mathbf{R}^3$ ל־ $\mathbf{R}^2$

נגדיר העתקה ליניארית $T \in \mathcal{L}(\mathbf{R}^3, \mathbf{R}^2)$ על ידי

$$T(x, y, z) = (2x - y + 3z, 7x + 5y - 6z).$$

מ־ $\mathbf{F}^n$ ל־ $\mathbf{F}^m$

כהכללה של המקרה הקודם, יהיו m ו־ n מספרים שלמים חיוביים, נגדיר את $A_{j,k} \in \mathbf{F}$ עבור כל $j = 1, \dots, m$ וכל $k = 1, \dots, n$, ואת ההעתקה הליניארית $T \in \mathcal{L}(\mathbf{F}^n, \mathbf{F}^m)$ על ידי

$$T(x_1, \dots, x_n) = (A_{1,1}x_1 + \cdots + A_{1,n}x_n, \dots, A_{m,1}x_1 + \cdots + A_{m,n}x_n).$$

למעשה כל העתקה ליניארית מ־ $\mathbf{F}^n$ ל־ $\mathbf{F}^m$ היא בתבנית זו.

הרכבה

נְקַבֵּעַ פולינום $q \in \mathcal{P}(\mathbf{R})$. נגדיר העתקה ליניארית $T \in \mathcal{L}(\mathcal{P}(\mathbf{R}))$ על ידי

$$(Tp)(x) = p(q(x)).$$

במסקנה להלן, החלק העוסק בקיומה של ההעתקה מורה לנו כי לכל קבוצת ערכים שנבחר להתאים לווקטורי הבסיס, נמצא העתקה ליניארית מתאימה. החלק העוסק ביחידותה של ההעתקה מורה לנו כי העתקה ליניארית נקבעת לחלוטין על ידי הערכים שלה מעל הבסיס.

3.4 *linear map lemma*

Suppose $v_1, \dots, v_n$ is a basis of V and $w_1, \dots, w_n \in W$. Then there exists a unique linear map $T\colon V \to W$ such that

$$Tv_k = w_k$$

for each $k = 1, \dots, n$.

Proof First we show the existence of a linear map T with the desired property. Define $T\colon V \to W$ by

$$T(c_1v_1 + \cdots + c_nv_n) = c_1w_1 + \cdots + c_nw_n,$$

where $c_1, \dots, c_n$ are arbitrary elements of $\mathbf{F}$. The list $v_1, \dots, v_n$ is a basis of V. Thus the equation above does indeed define a function T from V to W (because each element of V can be uniquely written in the form $c_1v_1 + \cdots + c_nv_n$).

For each k, taking $c_k = 1$ and the other c's equal to 0 in the equation above shows that $Tv_k = w_k$.

If $u, v \in V$ with $u = a_1v_1 + \cdots + a_nv_n$ and $v = c_1v_1 + \cdots + c_nv_n$, then

$$\begin{aligned} T(u + v) &= T((a_1 + c_1)v_1 + \cdots + (a_n + c_n)v_n) \\ &= (a_1 + c_1)w_1 + \cdots + (a_n + c_n)w_n \\ &= (a_1w_1 + \cdots + a_nw_n) + (c_1w_1 + \cdots + c_nw_n) \\ &= Tu + Tv. \end{aligned}$$

Similarly, if $\lambda \in \mathbf{F}$ and $v = c_1v_1 + \cdots + c_nv_n$, then

$$\begin{aligned} T(\lambda v) &= T(\lambda c_1v_1 + \cdots + \lambda c_nv_n) \\ &= \lambda c_1w_1 + \cdots + \lambda c_nw_n \\ &= \lambda(c_1w_1 + \cdots + c_nw_n) \\ &= \lambda Tv. \end{aligned}$$

Thus T is a linear map from V to W.

To prove uniqueness, now suppose that $T \in \mathcal{L}(V, W)$ and that $Tv_k = w_k$ for each $k = 1, \dots, n$. Let $c_1, \dots, c_n \in \mathbf{F}$. Then the homogeneity of T implies that $T(c_kv_k) = c_kw_k$ for each $k = 1, \dots, n$. The additivity of T now implies that

$$T(c_1v_1 + \cdots + c_nv_n) = c_1w_1 + \cdots + c_nw_n.$$

Thus T is uniquely determined on $\operatorname{span}(v_1, \dots, v_n)$ by the equation above. Because $v_1, \dots, v_n$ is a basis of V, this implies that T is uniquely determined on V, as desired. ∎

3.4 לֶמַת ההעתקה הליניארית

נניח ש־ $v_1, \dots, v_n$ בסיס של V ו־ $w_1, \dots, w_n \in W$. אזי קיימת העתקה ליניארית יחידה $T: V \to W$ כך ש־

$$Tv_k = w_k$$

לכל $k = 1, \dots, n$.

הוכחה ראשית נוכיח את קיומה של ההעתקה T עם התכונה הרצויה. נגדיר $T: V \to W$ על ידי

$$T(c_1v_1 + \cdots + c_nv_n) = c_1w_1 + \cdots + c_nw_n,$$

כאשר $c_1, \dots, c_n$ איברים כלשהם ב־ $\mathbf{F}$. הרשימה $v_1, \dots, v_n$ היא בסיס של V, ולפיכך המשוואה לעיל אכן מגדירה פונקציה T מ־ V ל־ W (מאחר שכל איבר של V יכול להיכתב באופן יחיד בצורה $c_1v_1 + \cdots + c_nv_n$).

עבור כל k, נבחר $c_k = 1$ ו־ 0 עבור שאר ה־ c־ים במשוואה לעיל ונקבל ש־ $Tv_k = w_k$.

אם $u, v \in V$ כאשר $u = a_1v_1 + \cdots + a_nv_n$ ו־ $v = c_1v_1 + \cdots + c_nv_n$, אזי

$$\begin{aligned} T(u + v) &= T((a_1 + c_1)v_1 + \cdots + (a_n + c_n)v_n) \\ &= (a_1 + c_1)w_1 + \cdots + (a_n + c_n)w_n \\ &= (a_1w_1 + \cdots + a_nw_n) + (c_1w_1 + \cdots + c_nw_n) \\ &= Tu + Tv. \end{aligned}$$

ובאותו האופן, עבור $\lambda \in \mathbf{F}$ ו־ $v = c_1v_1 + \cdots + c_nv_n$, נקבל

$$\begin{aligned} T(\lambda v) &= T(\lambda c_1v_1 + \cdots + \lambda c_nv_n) \\ &= \lambda c_1w_1 + \cdots + \lambda c_nw_n \\ &= \lambda(c_1w_1 + \cdots + c_nw_n) \\ &= \lambda Tv. \end{aligned}$$

לפיכך T העתקה ליניארית מ־ V ל־ W.

להוכחת היחידות, נניח כעת ש־ $T \in \mathcal{L}(V, W)$ ו־ $Tv_k = w_k$ לכל $k = 1, \dots, n$. יהיו $c_1, \dots, c_n \in \mathbf{F}$. אזי ההומוגניות של T גוררת ש־ $T(c_kv_k) = c_kw_k$ לכל $k = 1, \dots, n$. האדיטיביות של T גוררת ש־

$$T(c_1v_1 + \cdots + c_nv_n) = c_1w_1 + \cdots + c_nw_n.$$

לפיכך T נקבעת באופן יחיד מעל $\operatorname{span}(v_1, \dots, v_n)$ על ידי המשוואה לעיל. ומאחר ש־ $v_1, \dots, v_n$ בסיס של V, נסיק כי T נקבעת באופן יחיד מעל V, כפי שרצינו. ■

Algebraic Operations on $\mathcal{L}(V, W)$

We begin by defining addition and scalar multiplication on $\mathcal{L}(V, W)$.

3.5 definition: *addition and scalar multiplication on $\mathcal{L}(V, W)$*

Suppose $S, T \in \mathcal{L}(V, W)$ and $\lambda \in \mathbf{F}$. The *sum* $S + T$ and the *product* λT are the linear maps from V to W defined by

$$(S + T)(v) = Sv + Tv \quad \text{and} \quad (\lambda T)(v) = \lambda(Tv)$$

for all $v \in V$.

You should verify that $S + T$ and λT as defined above are indeed linear maps. In other words, if $S, T \in \mathcal{L}(V, W)$ and $\lambda \in \mathbf{F}$, then $S + T \in \mathcal{L}(V, W)$ and $\lambda T \in \mathcal{L}(V, W)$.

Because we took the trouble to define addition and scalar multiplication on $\mathcal{L}(V, W)$, the next result should not be a surprise.

Linear maps are pervasive throughout mathematics. However, they are not as ubiquitous as imagined by people who seem to think cos *is a linear map from* $\mathbf{R}$ *to* $\mathbf{R}$ *when they incorrectly write that* $\cos(x+y)$ *equals* $\cos x + \cos y$ *and that* $\cos 2x$ *equals* $2\cos x$.

3.6 $\mathcal{L}(V, W)$ *is a vector space*

With the operations of addition and scalar multiplication as defined above, $\mathcal{L}(V, W)$ is a vector space.

The routine proof of the result above is left to the reader. Note that the additive identity of $\mathcal{L}(V, W)$ is the zero linear map defined in Example 3.3.

Usually it makes no sense to multiply together two elements of a vector space, but for some pairs of linear maps a useful product exists, as in the next definition.

3.7 definition: *product of linear maps*

If $T \in \mathcal{L}(U, V)$ and $S \in \mathcal{L}(V, W)$, then the *product* $ST \in \mathcal{L}(U, W)$ is defined by

$$(ST)(u) = S(Tu)$$

for all $u \in U$.

Thus ST is just the usual composition $S \circ T$ of two functions, but when both functions are linear, we usually write ST instead of $S \circ T$. The product notation ST helps make the distributive properties (see next result) seem natural.

Note that ST is defined only when T maps into the domain of S. You should verify that ST is indeed a linear map from U to W whenever $T \in \mathcal{L}(U, V)$ and $S \in \mathcal{L}(V, W)$.

פעולות אלגבריות מעל $\mathcal{L}(V,W)$

אנו מתחילים בהגדרת חיבור וכפל בסקלר מעל $\mathcal{L}(V,W)$.

3.5 הגדרה: חיבור וכפל בסקלר מעל $\mathcal{L}(V,W)$

נניח ש־ $S,T \in \mathcal{L}(V,W)$ ו־ $\lambda \in \mathbf{F}$. הסכום $S+T$ והמכפלה λT הם ההעתקות הליניאריות מ־ V ל־ W המוגדרות על ידי

$$(S+T)(v) = Sv + Tv \quad \text{וגם} \quad (\lambda T)(v) = \lambda(Tv)$$

לכל $v \in V$.

עליכם לוודא כי $S+T$ ו־ λT כפי שהוגדרו לעיל הן אכן העתקות ליניאריות. כלומר, אם $S,T \in \mathcal{L}(V,W)$ ו־ $\lambda \in \mathbf{F}$, אזי $S+T \in \mathcal{L}(V,W)$ ו־ $\lambda T \in \mathcal{L}(V,W)$.

לאור העובדה כי טרחנו להגדיר חיבור וכפל בסקלר מעל $\mathcal{L}(V,W)$, המסקנה הבאה אינה אמורה להפתיע.

העתקות ליניאריות נפוצות בענפי מתמטיקה רבים. עם זאת, הן פחות נפוצות משנוטים לחשוב: יש הסבורים ש־ cos היא העתקה ליניארית מ־ $\mathbf{R}$ ל־ $\mathbf{R}$ לאחר שכתבו בטעות ש־ $\cos(x+y) = \cos x + \cos y$ וגם ש־ $\cos 2x = 2\cos x$.

3.6 $\mathcal{L}(V,W)$ הוא מרחב וקטורי

עם פעולות החיבור והכפל בסקלר כפי שהוגדרו לעיל, $\mathcal{L}(V,W)$ מרחב וקטורי.

את ההוכחה השגרתית למסקנה זו נשאיר עבורכם, הקוראים. שימו לב כי איבר הזהות לחיבור של $\mathcal{L}(V,W)$ הוא העתקת האפס המוגדרת בדוגמה 3.3.

בדרך כלל פעולת ההכפלה של שני איברים זה בזה במרחב וקטורי חסרת משמעות. אך ישנם זוגות של העתקות ליניאריות שלמכפלת אחת באחרת יש שימוש רב, כפי שנראה בהגדרה הבאה.

3.7 הגדרה: מכפלת העתקות ליניאריות

אם $T \in \mathcal{L}(U,V)$ ו־ $S \in \mathcal{L}(V,W)$, אזי המכפלה $ST \in \mathcal{L}(U,W)$ מוגדרת על ידי

$$(ST)(u) = S(Tu)$$

לכל $u \in U$.

לפיכך ST היא פעולת ההרכבה הרגילה $S \circ T$ של שתי פונקציות, אך כאשר שתי הפונקציות ליניאריות, נכתוב ST בדרך כלל במקום $S \circ T$. סימון המכפלה ST מסייע לכך שתכונות הדיסטריבוטיביות (ראו במסקנה הבאה) תיראינה טבעיות.

שימו לב כי ST מוגדרת רק כאשר T מעתיקה אל התחום של S. עליכם לוודא ש־ ST אכן העתקה ליניארית מ־ U ל־ W כאשר $T \in \mathcal{L}(U,V)$ ו־ $S \in \mathcal{L}(V,W)$.

3.8 *algebraic properties of products of linear maps*

associativity
$(T_1T_2)T_3 = T_1(T_2T_3)$ whenever T_1, T_2, and T_3 are linear maps such that the products make sense (meaning T_3 maps into the domain of T_2, and T_2 maps into the domain of T_1).

identity
$TI = IT = T$ whenever $T \in \mathcal{L}(V, W)$; here the first I is the identity operator on V, and the second I is the identity operator on W.

distributive properties
$(S_1 + S_2)T = S_1T + S_2T$ and $S(T_1 + T_2) = ST_1 + ST_2$ whenever $T, T_1, T_2 \in \mathcal{L}(U, V)$ and $S, S_1, S_2 \in \mathcal{L}(V, W)$.

The routine proof of the result above is left to the reader.

Multiplication of linear maps is not commutative. In other words, it is not necessarily true that $ST = TS$, even if both sides of the equation make sense.

3.9 example: *two noncommuting linear maps from $\mathcal{P}(\mathbf{R})$ to $\mathcal{P}(\mathbf{R})$*

Suppose $D \in \mathcal{L}(\mathcal{P}(\mathbf{R}))$ is the differentiation map defined in Example 3.3 and $T \in \mathcal{L}(\mathcal{P}(\mathbf{R}))$ is the multiplication by x^2 map defined earlier in this section. Then

$$((TD)p)(x) = x^2p'(x) \quad \text{but} \quad ((DT)p)(x) = x^2p'(x) + 2xp(x).$$

Thus $TD \neq DT$—differentiating and then multiplying by x^2 is not the same as multiplying by x^2 and then differentiating.

3.10 *linear maps take 0 to 0*

Suppose T is a linear map from V to W. Then $T(0) = 0$.

Proof By additivity, we have

$$T(0) = T(0 + 0) = T(0) + T(0).$$

Add the additive inverse of $T(0)$ to each side of the equation above to conclude that $T(0) = 0$.

Suppose $m, b \in \mathbf{R}$. The function $f : \mathbf{R} \to \mathbf{R}$ defined by

$$f(x) = mx + b$$

is a linear map if and only if $b = 0$ (use 3.10). Thus the linear functions of high school algebra are not the same as linear maps in the context of linear algebra.

3.8 תכונות אלגבריות של מכפלות העתקות ליניאריות

אסוציאטיביות
$(T_1T_2)T_3 = T_1(T_2T_3)$ כאשר T_1, T_2, ו־ T_3 העתקות ליניאריות במקרים שבהם למכפלות יש משמעות (כלומר T_3 מעתיקה אל התחום של T_2, ו־ T_2 מעתיקה אל התחום של T_1).

זהות
$TI = IT = T$ כאשר $T \in \mathcal{L}(V, W)$; במקרה זה ה־ I הראשון הוא אופרטור הזהות ב־ V, ואילו ה־ I השני הוא אופרטור הזהות ב־ W.

תכונות הדיסטריבוטיביות
$S(T_1 + T_2) = ST_1 + ST_2$ וגם $(S_1 + S_2)T = S_1T + S_2T$ כאשר $T, T_1, T_2 \in \mathcal{L}(U, V)$ ו־ $S, S_1, S_2 \in \mathcal{L}(V, W)$.

נשאיר לקוראים את ההוכחה השגרתית של המסקנה שלעיל.

פעולת הכפל של העתקות ליניאריות אינה קומוטטיבית. כלומר, $ST = TS$ אינו מתקיים בהכרח, גם אם לכל מכפלה יש משמעות בפני עצמה.

3.9 דוגמה: שתי העתקות ליניאריות מ־ $\mathcal{P}(\mathbf{R})$ ל־ $\mathcal{P}(\mathbf{R})$ שאינן קומוטטיביות

נניח ש־ $D \in \mathcal{L}(\mathcal{P}(\mathbf{R}))$ היא העתקת הגזירה שהוגדרה בדוגמה 3.3 ו־ $T \in \mathcal{L}(\mathcal{P}(\mathbf{R}))$ היא העתקת הכפלה ב־ x^2 שהוגדרה גם היא לעיל ביחידה זו. אזי
$$((DT)p)(x) = x^2p'(x) + 2xp(x) \quad \text{אולם} \quad ((TD)p)(x) = x^2p'(x).$$
לפיכך $DT \neq TD$—גזירה ואז הכפלה ב־ x^2 אינה זהה להכפלה ב־ x^2 ואז גזירה.

3.10 העתקות ליניאריות מעתיקות 0 ל־ 0

נניח ש־ T העתקה ליניארית מ־ V ל־ W. אזי $T(0) = 0$.

הוכחה לפי האדיטיביות מתקיים
$$T(0) = T(0 + 0) = T(0) + T(0).$$
נוסיף את האיבר הנגדי לחיבור של $T(0)$ לכל אגף במשוואה לעיל ונסיק ש־ $T(0) = 0$. ■

נניח ש־ $m, b \in \mathbf{R}$. הפונקציה $f: \mathbf{R} \to \mathbf{R}$ המוגדרת על ידי
$$f(x) = mx + b$$
היא העתקה ליניארית אם־ורק־אם $b = 0$ (השתמשו ב־ 3.10). לפיכך, הפונקציות הליניאריות משיעורי האלגברה בתיכון שונות מהעתקות ליניאריות בהקשר של אלגברה ליניארית.

Exercises 3A

1 Suppose $b, c \in \mathbf{R}$. Define $T\colon \mathbf{R}^3 \to \mathbf{R}^2$ by

$$T(x, y, z) = (2x - 4y + 3z + b, 6x + cxyz).$$

Show that T is linear if and only if $b = c = 0$.

2 Suppose $b, c \in \mathbf{R}$. Define $T\colon \mathcal{P}(\mathbf{R}) \to \mathbf{R}^2$ by

$$Tp = \Big(3p(4) + 5p'(6) + bp(1)p(2), \int_{-1}^{2} x^3 p(x)\,dx + c \sin p(0)\Big).$$

Show that T is linear if and only if $b = c = 0$.

3 Suppose that $T \in \mathcal{L}(\mathbf{F}^n, \mathbf{F}^m)$. Show that there exist scalars $A_{j,k} \in \mathbf{F}$ for $j = 1, \dots, m$ and $k = 1, \dots, n$ such that

$$T(x_1, \dots, x_n) = (A_{1,1}x_1 + \cdots + A_{1,n}x_n, \dots, A_{m,1}x_1 + \cdots + A_{m,n}x_n)$$

for every $(x_1, \dots, x_n) \in \mathbf{F}^n$.

This exercise shows that the linear map T has the form promised in the second to last item of Example 3.3.

4 Suppose $T \in \mathcal{L}(V, W)$ and $v_1, \dots, v_m$ is a list of vectors in V such that $Tv_1, \dots, Tv_m$ is a linearly independent list in W. Prove that $v_1, \dots, v_m$ is linearly independent.

5 Prove that $\mathcal{L}(V, W)$ is a vector space, as was asserted in 3.6.

6 Prove that multiplication of linear maps has the associative, identity, and distributive properties asserted in 3.8.

7 Show that every linear map from a one-dimensional vector space to itself is multiplication by some scalar. More precisely, prove that if $\dim V = 1$ and $T \in \mathcal{L}(V)$, then there exists $\lambda \in \mathbf{F}$ such that $Tv = \lambda v$ for all $v \in V$.

8 Give an example of a function $\varphi\colon \mathbf{R}^2 \to \mathbf{R}$ such that

$$\varphi(av) = a\varphi(v)$$

for all $a \in \mathbf{R}$ and all $v \in \mathbf{R}^2$ but φ is not linear.

This exercise and the next exercise show that neither homogeneity nor additivity alone is enough to imply that a function is a linear map.

9 Give an example of a function $\varphi\colon \mathbf{C} \to \mathbf{C}$ such that

$$\varphi(w + z) = \varphi(w) + \varphi(z)$$

for all $w, z \in \mathbf{C}$ but φ is not linear. (Here $\mathbf{C}$ is thought of as a complex vector space.)

There also exists a function $\varphi\colon \mathbf{R} \to \mathbf{R}$ such that φ satisfies the additivity condition above but φ is not linear. However, showing the existence of such a function involves considerably more advanced tools.

תרגילים ג1

1 נניח ש־ $b, c \in \mathbf{R}$. הגדירו $T: \mathbf{R}^3 \to \mathbf{R}^2$ על ידי

$$T(x, y, z) = (2x - 4y + 3z + b, 6x + cxyz).$$

הראו כי T העתקה ליניארית אם־ורק־אם $b = c = 0$.

2 נניח ש־ $b, c \in \mathbf{R}$. הגדירו $T: \mathcal{P}(\mathbf{R}) \to \mathbf{R}^2$ על ידי

$$Tp = \Big(3p(4) + 5p'(6) + bp(1)p(2), \int_{-1}^{2} x^3 p(x)\, dx + c \sin p(0)\Big).$$

הראו כי T העתקה ליניארית אם־ורק־אם $b = c = 0$.

3 נניח ש־ $T \in \mathcal{L}(\mathbf{F}^n, \mathbf{F}^m)$. הראו כי קיימים סקלרים $A_{j,k} \in \mathbf{F}$ עבור $j = 1, \dots, m$ ו־ $k = 1, \dots, n$ כך ש־

$$T(x_1, \dots, x_n) = (A_{1,1}x_1 + \cdots + A_{1,n}x_n, \dots, A_{m,1}x_1 + \cdots + A_{m,n}x_n)$$

לכל $(x_1, \dots, x_n) \in \mathbf{F}^n$.

תרגיל זה מראה שיש להעתקה הליניארית T את הצורה שהובטחה בסעיף שלפני האחרון בדוגמה 3.3.

4 נניח ש־ $T \in \mathcal{L}(V, W)$ ו־ $v_1, \dots, v_m$ רשימת וקטורים ב־ V כך ש־ $Tv_1, \dots, Tv_m$ רשימה בלתי תלויה ליניארית ב־ W. הוכיחו כי $v_1, \dots, v_m$ בלתי תלויה ליניארית.

5 הוכיחו כי $\mathcal{L}(V, W)$ מרחב וקטורי, כפי שנטען ב־ 3.6.

6 הוכיחו כי מכפלת העתקות ליניאריות מקיימת את תכונות האסוציאטיביות, הזהות, והדיסטריבוטיביות שנטענו ב־ 3.8.

7 הראו כי כל העתקה ליניארית ממרחב וקטורי חד־ממדי לעצמו היא כפל בסקלר. ביתר דיוק, הוכיחו כי אם $\dim V = 1$ ו־ $T \in \mathcal{L}(V)$, אזי קיים $\lambda \in \mathbf{F}$ כך ש־ $Tv = \lambda v$ לכל $v \in V$.

8 הביאו דוגמה של פונקציה $\varphi: \mathbf{R}^2 \to \mathbf{R}$ כך ש־

$$\varphi(av) = a\varphi(v)$$

לכל $a \in \mathbf{R}$ ולכל $v \in \mathbf{R}^2$ אבל φ אינה ליניארית.

תרגיל זה והבא אחריו מיועדים שההוכחות או האזיטיביות לבדן אינן מספיקות כדי לקבוע כי פונקציה היא העתקה ליניארית.

9 הביאו דוגמה לפונקציה $\varphi: \mathbf{C} \to \mathbf{C}$ כך ש־

$$\varphi(w + z) = \varphi(w) + \varphi(z)$$

לכל $w, z \in \mathbf{C}$ אך φ אינה ליניארית. (בשאלה זו $\mathbf{C}$ הוא מרחב וקטורי מרוכב.)

קיימת גם פונקציה $\varphi: \mathbf{R} \to \mathbf{R}$ הפקיימת את האדיטיביות אך אינה ליניארית. אולם, הוכחת קיומה של פונקציה כזו מצריכה כלים מתקדמים בהרבה.

10 Prove or give a counterexample: If $q \in \mathcal{P}(\mathbf{R})$ and $T\colon \mathcal{P}(\mathbf{R}) \to \mathcal{P}(\mathbf{R})$ is defined by $Tp = q \circ p$, then T is a linear map.

The function T defined here differs from the function T defined in the last bullet point of 3.3 by the order of the functions in the compositions.

11 Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Prove that T is a scalar multiple of the identity if and only if $ST = TS$ for every $S \in \mathcal{L}(V)$.

12 Suppose U is a subspace of V with $U \neq V$. Suppose $S \in \mathcal{L}(U, W)$ and $S \neq 0$ (which means that $Su \neq 0$ for some $u \in U$). Define $T\colon V \to W$ by

$$Tv = \begin{cases} Sv & \text{if } v \in U, \\ 0 & \text{if } v \in V \text{ and } v \notin U. \end{cases}$$

Prove that T is not a linear map on V.

13 Suppose V is finite-dimensional. Prove that every linear map on a subspace of V can be extended to a linear map on V. In other words, show that if U is a subspace of V and $S \in \mathcal{L}(U, W)$, then there exists $T \in \mathcal{L}(V, W)$ such that $Tu = Su$ for all $u \in U$.

The result in this exercise is used in the proof of 3.125.

14 Suppose V is finite-dimensional with $\dim V > 0$, and suppose W is infinite-dimensional. Prove that $\mathcal{L}(V, W)$ is infinite-dimensional.

15 Suppose $v_1, \dots, v_m$ is a linearly dependent list of vectors in V. Suppose also that $W \neq \{0\}$. Prove that there exist $w_1, \dots, w_m \in W$ such that no $T \in \mathcal{L}(V, W)$ satisfies $Tv_k = w_k$ for each $k = 1, \dots, m$.

16 Suppose V is finite-dimensional with $\dim V > 1$. Prove that there exist $S, T \in \mathcal{L}(V)$ such that $ST \neq TS$.

17 Suppose V is finite-dimensional. Show that the only two-sided ideals of $\mathcal{L}(V)$ are $\{0\}$ and $\mathcal{L}(V)$.

*A subspace $\mathcal{E}$ of $\mathcal{L}(V)$ is called a **two-sided ideal** of $\mathcal{L}(V)$ if $TE \in \mathcal{E}$ and $ET \in \mathcal{E}$ for all $E \in \mathcal{E}$ and all $T \in \mathcal{L}(V)$.*

10 הוכיחו או הביאו דוגמה נגדית: אם $q \in \mathcal{P}(\mathbf{R})$ ו־ $T\colon \mathcal{P}(\mathbf{R}) \to \mathcal{P}(\mathbf{R})$ מוגדרת על ידי $Tp = q \circ p$, אזי T העתקה ליניארית.

הפונקציה T המוגדרת כאן שונה מהפונקציה T בסעיף האחרון של 3.3 בסדר ההרכבה של הפונקציות.

11 נניח ש־ V מרחב מממד סופי ו־ $T \in \mathcal{L}(V)$. הוכיחו כי T היא כפולת סקלר של אופרטור הזהות אם־ורק־אם $ST = TS$ לכל $S \in \mathcal{L}(V)$.

12 נניח ש־ U תת־מרחב של V כאשר $U \neq V$. נניח ש־ $S \in \mathcal{L}(U, W)$ וגם $S \neq 0$ (כלומר $Su \neq 0$ עבור $u \in U$ כלשהו). הגדירו $T\colon V \to W$ על ידי

$$Tv = \begin{cases} Sv & \text{if } v \in U, \\ 0 & \text{if } v \in V,\ v \notin U. \end{cases}$$

והוכיחו כי T אינה העתקה ליניארית מעל V.

13 נניח ש־ V מרחב מממד סופי. הוכיחו כי ניתן להרחיב כל העתקה ליניארית מעל תת־מרחב של V כך שתפעל מעל V. כלומר, הראו כי אם U תת־מרחב של V ו־ $S \in \mathcal{L}(U, W)$, אזי קיימת $T \in \mathcal{L}(V, W)$ כזו המקיימת $Tu = Su$ לכל $u \in U$.

התוצאה בתרגיל זה משמשת בהוכחת 3.125.

14 נניח כי V מרחב מממד סופי המקיים $\dim V > 0$, ונניח בנוסף כי W הוא מרחב מממד אינסופי. הוכיחו כי $\mathcal{L}(V, W)$ מרחב מממד אינסופי.

15 נניח כי $v_1, \dots, v_m$ רשימת וקטורים תלויה ליניארית ב־ V. נניח בנוסף כי $W \neq \{0\}$. הוכיחו שקיימים $w_1, \dots, w_m \in W$ כך שאין $T \in \mathcal{L}(V, W)$ המקיימת $Tv_k = w_k$ לכל $k = 1, \dots, m$.

16 נניח ש־ V מרחב מממד סופי עם $\dim V > 1$. הוכיחו כי קיימות $S, T \in \mathcal{L}(V)$ כך ש־ $ST \neq TS$.

17 נניח ש־ V מרחב מממד סופי. הראו כי עבור $\mathcal{L}(V)$, רק $\{0\}$ ו־ $\mathcal{L}(V)$ הם אידיאלים דו־צדדיים.

תת־מרחב $\mathcal{E}$ של $\mathcal{L}(V)$ נקרא **אידיאל דו־צדדי** של $\mathcal{L}(V)$ אם $TE \in \mathcal{E}$ וגם $ET \in \mathcal{E}$ לכל $E \in \mathcal{E}$ וכל $T \in \mathcal{L}(V)$.

3B Null Spaces and Ranges

Null Space and Injectivity

In this section we will learn about two subspaces that are intimately connected with each linear map. We begin with the set of vectors that get mapped to 0.

3.11 definition: *null space,* null T

For $T \in \mathcal{L}(V, W)$, the *null space* of T, denoted by $\operatorname{null} T$, is the subset of V consisting of those vectors that T maps to 0:

$$\operatorname{null} T = \{v \in V : Tv = 0\}.$$

3.12 example: *null space*

- If T is the zero map from V to W, meaning that $Tv = 0$ for every $v \in V$, then $\operatorname{null} T = V$.
- Suppose $\varphi \in \mathcal{L}(\mathbf{C}^3, \mathbf{C})$ is defined by $\varphi(z_1, z_2, z_3) = z_1 + 2z_2 + 3z_3$. Then $\operatorname{null} \varphi$ equals $\{(z_1, z_2, z_3) \in \mathbf{C}^3 : z_1 + 2z_2 + 3z_3 = 0\}$, which is a subspace of the domain of φ. We will soon see that the null space of each linear map is a subspace of its domain.
- Suppose $D \in \mathcal{L}(\mathcal{P}(\mathbf{R}))$ is the differentiation map defined by $Dp = p'$. The only functions whose derivative equals the zero function are the constant functions. Thus the null space of D equals the set of constant functions.

*The word "null" means zero. Thus the term "null space" should remind you of the connection to 0. Some mathematicians use the term **kernel** instead of null space.*

- Suppose that $T \in \mathcal{L}(\mathcal{P}(\mathbf{R}))$ is the multiplication by x^2 map defined by $(Tp)(x) = x^2 p(x)$. The only polynomial p such that $x^2 p(x) = 0$ for all $x \in \mathbf{R}$ is the 0 polynomial. Thus $\operatorname{null} T = \{0\}$.
- Suppose $T \in \mathcal{L}(\mathbf{F}^\infty)$ is the backward shift defined by

$$T(x_1, x_2, x_3, \dots) = (x_2, x_3, \dots).$$

Then $T(x_1, x_2, x_3, \dots)$ equals 0 if and only if the numbers $x_2, x_3, \dots$ are all 0. Thus $\operatorname{null} T = \{(a, 0, 0, \dots) : a \in \mathbf{F}\}$.

The next result shows that the null space of each linear map is a subspace of the domain. In particular, 0 is in the null space of every linear map.

3.13 *the null space is a subspace*

Suppose $T \in \mathcal{L}(V, W)$. Then $\operatorname{null} T$ is a subspace of V.

ג מרחבי אפס וטווחים

מרחב אפס וחד־חד־ערכיות

ביחידה זו נלמד על שני תתי־מרחב הקשורים בקשר הדוק עם כל העתקה ליניארית. נתחיל עם קבוצת הווקטורים המועתקת אל 0.

3.11 הגדרה: מרחב האפס, $\operatorname{null} T$

עבור $T \in \mathcal{L}(V, W)$, מרחב האפס של T, המסומן ב־ $\operatorname{null} T$, הוא תת־הקבוצה של V המורכבת מכל אותם וקטורים ש־ T מעתיקה אל איבר האפס:

$$\operatorname{null} T = \{v \in V : Tv = 0\}.$$

3.12 דוגמה: מרחב האפס

- אם T היא העתקת האפס מ־ V ל־ W, כלומר $Tv = 0$ לכל $v \in V$, אזי $\operatorname{null} T = V$.
- נניח ש־ $\varphi \in \mathcal{L}(\mathbf{C}^3, \mathbf{C})$ מוגדרת באופן הבא $\varphi(z_1, z_2, z_3) = z_1 + 2z_2 + 3z_3$. אזי $\operatorname{null} \varphi$ שווה ל־ $\{(z_1, z_2, z_3) \in \mathbf{C}^3 : z_1 + 2z_2 + 3z_3 = 0\}$, שהוא תת־מרחב של התחום של φ. בהמשך נראה כי מרחב האפס של כל העתקה ליניארית הוא תת־מרחב של התחום שלה.
- נניח ש־ $D \in \mathcal{L}(\mathcal{P}(\mathbf{R}))$ העתקת הגזירה המוגדרת על ידי $Dp = p'$. הפונקציות היחידות שהנגזרת שלהן שווה לפונקציית האפס הן הפונקציות הקבועות. לפיכך, מרחב האפס של D שווה לקבוצת הפונקציות הקבועות.

המילה null משמעותה אפס. לפיכך המונח "מרחב האפס" מייצג באופן ישיר את הקשר ל־ 0. ישנם מתמטיקאים המשתמשים במונח **גרעין** במקום מרחב האפס.

- נניח ש־ $T \in \mathcal{L}(\mathcal{P}(\mathbf{R}))$ היא העתקת המכפילה ב־ x^2 ומוגדרת באופן הבא: $(Tp)(x) = x^2 p(x)$. הפולינום היחיד p כך ש־ $x^2 p(x) = 0$ לכל $x \in \mathbf{R}$ הוא פולינום האפס. לפיכך $\operatorname{null} T = \{0\}$.
- נניח ש־ $T \in \mathcal{L}(\mathbf{F}^\infty)$ היא העתקת הסטה לאחור המוגדרת על ידי
$$T(x_1, x_2, x_3, \dots) = (x_2, x_3, \dots).$$
אזי $T(x_1, x_2, x_3, \dots)$ שווה 0 אם־ורק־אם המספרים $x_2, x_3, \dots$ כולם 0. לפיכך $\operatorname{null} T = \{(a, 0, 0, \dots) : a \in \mathbf{F}\}$.

המסקנה הבאה מראה שמרחב האפס של כל העתקה ליניארית הוא תת־מרחב של התחום. בפרט, 0 הוא במרחב האפס של כל העתקה ליניארית.

3.13 מרחב האפס הוא תת־מרחב

נניח ש־ $T \in \mathcal{L}(V, W)$. אזי $\operatorname{null} T$ הוא תת־מרחב של V.

Proof Because T is a linear map, $T(0) = 0$ (by 3.10). Thus $0 \in \operatorname{null} T$.

Suppose $u, v \in \operatorname{null} T$. Then

$$T(u + v) = Tu + Tv = 0 + 0 = 0.$$

Hence $u + v \in \operatorname{null} T$. Thus $\operatorname{null} T$ is closed under addition.

Suppose $u \in \operatorname{null} T$ and $\lambda \in \mathbf{F}$. Then

$$T(\lambda u) = \lambda Tu = \lambda 0 = 0.$$

Hence $\lambda u \in \operatorname{null} T$. Thus $\operatorname{null} T$ is closed under scalar multiplication.

We have shown that $\operatorname{null} T$ contains 0 and is closed under addition and scalar multiplication. Thus $\operatorname{null} T$ is a subspace of V (by 1.34). ∎

As we will soon see, for a linear map the next definition is closely connected to the null space.

3.14 definition: *injective*

A function $T : V \to W$ is called *injective* if $Tu = Tv$ implies $u = v$.

*The term **one-to-one** means the same as injective.*

We could rephrase the definition above to say that T is injective if $u \neq v$ implies that $Tu \neq Tv$. Thus T is injective if and only if it maps distinct inputs to distinct outputs.

The next result says that we can check whether a linear map is injective by checking whether 0 is the only vector that gets mapped to 0. As a simple application of this result, we see that of the linear maps whose null spaces we computed in 3.12, only multiplication by x^2 is injective (except that the zero map is injective in the special case $V = \{0\}$).

3.15 *injectivity* ⟺ *null space equals* $\{0\}$

Let $T \in \mathcal{L}(V, W)$. Then T is injective if and only if $\operatorname{null} T = \{0\}$.

Proof First suppose T is injective. We want to prove that $\operatorname{null} T = \{0\}$. We already know that $\{0\} \subseteq \operatorname{null} T$ (by 3.10). To prove the inclusion in the other direction, suppose $v \in \operatorname{null} T$. Then

$$T(v) = 0 = T(0).$$

Because T is injective, the equation above implies that $v = 0$. Thus we can conclude that $\operatorname{null} T = \{0\}$, as desired.

To prove the implication in the other direction, now suppose $\operatorname{null} T = \{0\}$. We want to prove that T is injective. To do this, suppose $u, v \in V$ and $Tu = Tv$. Then

$$0 = Tu - Tv = T(u - v).$$

Thus $u - v$ is in $\operatorname{null} T$, which equals $\{0\}$. Hence $u - v = 0$, which implies that $u = v$. Hence T is injective, as desired. ∎

הוכחה מאחר ש־ T העתקה ליניארית, $T(0) = 0$ (לפי 3.10). לפיכך $0 \in \operatorname{null} T$.
נניח ש־ $u, v \in \operatorname{null} T$. אזי

$$T(u + v) = Tu + Tv = 0 + 0 = 0.$$

לכן $u + v \in \operatorname{null} T$. לפיכך $\operatorname{null} T$ סגור תחת פעולת החיבור.
נניח ש־ $u \in \operatorname{null} T$ ו־ $\lambda \in \mathbf{F}$. אזי

$$T(\lambda u) = \lambda Tu = \lambda 0 = 0.$$

לכן $\lambda u \in \operatorname{null} T$. לפיכך $\operatorname{null} T$ סגור תחת כפל בסקלר.
הראינו אפוא ש־ $\operatorname{null} T$ מכיל 0 וסגור תחת פעולות חיבור וכפל בסקלר. לפיכך $\operatorname{null} T$ תת־מרחב של V (לפי 1.34). ■

כפי שנראה מייד, להגדרה הבאה קשר הדוק למרחב האפס של העתקה ליניארית.

3.14 הגדרה: חד־חד־ערכית

פונקציה $T : V \to W$ נקראת חד־חד־ערכית אם $Tu = Tv$ גורר ש־ $u = v$.

משמעות המונח ***one-to-one*** זהה לחד־חד־ערכי.

נוכל לנסח מחדש את ההגדרה לעיל ולומר ש־ T חד־חד־ערכית אם $u \neq v$ גורר ש־ $Tu \neq Tv$. לפיכך T חד־חד־ערכית אם־ורק־אם היא מעתיקה קלטים שונים אל פלטים שונים.

המסקנה הבאה קובעת כי נוכל לבדוק אם העתקה ליניארית היא חד־חד־ערכית באמצעות הבדיקה האם 0 הוא האיבר היחיד המועתק ל־ 0. ביישום פשוט של מסקנה זו, אנו רואים כי מִבֵּין ההעתקות הליניאריות שחישבנו את מרחב האפס שלהן ב־ 3.12, רק ההכפלה ב־ x^2 היא חד־חד־ערכית (למעט העובדה שהעתקת האפס היא חד־חד־ערכית במקרה המיוחד שבו $V = \{0\}$).

3.15 חד־חד־ערכיות $\iff$ מרחב האפס שווה ל־ $\{0\}$

תהי $T \in \mathcal{L}(V, W)$. אזי T חד־חד־ערכית אם־ורק־אם $\operatorname{null} T = \{0\}$.

הוכחה תחילה נניח ש־ T חד־חד־ערכית. נרצה להוכיח ש־ $\operatorname{null} T = \{0\}$. כבר ביססנו את העובדה ש־ $\{0\} \subseteq \operatorname{null} T$ (לפי 3.10). להוכחת ההכלה בכיוון ההפוך, נניח ש־ $v \in \operatorname{null} T$. נקבל

$$T(v) = 0 = T(0).$$

מאחר ש־ T חד־חד־ערכית, המשוואה לעיל גוררת ש־ $v = 0$. לפיכך אנו יכולים להסיק כי $\operatorname{null} T = \{0\}$, כפי שרצינו.

להוכחת הגרירה בכיוון השני, נניח כעת ש־ $\operatorname{null} T = \{0\}$. ברצוננו להוכיח כי T חד־חד־ערכית. לצורך כך נניח ש־ $u, v \in V$ ושו־ $Tu = Tv$. אזי

$$0 = Tu - Tv = T(u - v).$$

לפיכך $u - v$ נמצא ב־ $\operatorname{null} T$, השווה ל־ $\{0\}$. לכן $u - v = 0$, ומכאן ש־ $u = v$. לפיכך T חד־חד־ערכית כפי שרצינו. ■

Range and Surjectivity

Now we give a name to the set of outputs of a linear map.

3.16 definition: *range*

For $T \in \mathcal{L}(V, W)$, the *range* of T is the subset of W consisting of those vectors that are equal to Tv for some $v \in V$:

$$\text{range } T = \{Tv : v \in V\}.$$

3.17 example: *range*

- If T is the zero map from V to W, meaning that $Tv = 0$ for every $v \in V$, then $\text{range } T = \{0\}$.
- Suppose $T \in \mathcal{L}(\mathbf{R}^2, \mathbf{R}^3)$ is defined by $T(x, y) = (2x, 5y, x + y)$. Then

  $$\text{range } T = \{(2x, 5y, x + y) : x, y \in \mathbf{R}\}.$$

  Note that $\text{range } T$ is a subspace of $\mathbf{R}^3$. We will soon see that the range of each element of $\mathcal{L}(V, W)$ is a subspace of W.
- Suppose $D \in \mathcal{L}(\mathcal{P}(\mathbf{R}))$ is the differentiation map defined by $Dp = p'$. Because for every polynomial $q \in \mathcal{P}(\mathbf{R})$ there exists a polynomial $p \in \mathcal{P}(\mathbf{R})$ such that $p' = q$, the range of D is $\mathcal{P}(\mathbf{R})$.

The next result shows that the range of each linear map is a subspace of the vector space into which it is being mapped.

3.18 *the range is a subspace*

If $T \in \mathcal{L}(V, W)$, then $\text{range } T$ is a subspace of W.

Proof Suppose $T \in \mathcal{L}(V, W)$. Then $T(0) = 0$ (by 3.10), which implies that $0 \in \text{range } T$.

If $w_1, w_2 \in \text{range } T$, then there exist $v_1, v_2 \in V$ such that $Tv_1 = w_1$ and $Tv_2 = w_2$. Thus

$$T(v_1 + v_2) = Tv_1 + Tv_2 = w_1 + w_2.$$

Hence $w_1 + w_2 \in \text{range } T$. Thus $\text{range } T$ is closed under addition.

If $w \in \text{range } T$ and $\lambda \in \mathbf{F}$, then there exists $v \in V$ such that $Tv = w$. Thus

$$T(\lambda v) = \lambda Tv = \lambda w.$$

Hence $\lambda w \in \text{range } T$. Thus $\text{range } T$ is closed under scalar multiplication.

We have shown that $\text{range } T$ contains 0 and is closed under addition and scalar multiplication. Thus $\text{range } T$ is a subspace of W (by 1.34). ∎

טווח ותכונת העל

כעת נעניק שם לקבוצת התוצאות של העתקה ליניארית.

3.16 הגדרה: טווח, range

עבור $T \in \mathcal{L}(V, W)$, הטווח של T הוא תת־הקבוצה של W המורכבת מכל אותם וקטורים השווים ל־ Tv עבור $v \in V$ כלשהו:

$$\operatorname{range} T = \{Tv : v \in V\}.$$

3.17 דוגמה: טווח, range

- אם T היא העתקת האפס מ־ V ל־ W, כלומר $Tv = 0$ לכל $v \in V$, אזי $\operatorname{range} T = \{0\}$.
- נניח ש־ $T \in \mathcal{L}(\mathbf{R}^2, \mathbf{R}^3)$ מוגדרת על ידי $T(x, y) = (2x, 5y, x + y)$. אזי
$$\operatorname{range} T = \{(2x, 5y, x + y) : x, y \in \mathbf{R}\}.$$
שימו לב ש־ $\operatorname{range} T$ הוא תת־מרחב של $\mathbf{R}^3$. בהמשך נראה שהטווח של כל איבר ב־ $\mathcal{L}(V, W)$ הוא תת־מרחב של W.
- נניח ש־ $D \in \mathcal{L}(\mathcal{P}(\mathbf{R}))$ היא העתקת הגזירה המוגדרת על ידי $Dp = p'$. מאחר שלכל פולינום $q \in \mathcal{P}(\mathbf{R})$ קיים פולינום $p \in \mathcal{P}(\mathbf{R})$ כך ש־ $p' = q$, הטווח של D הוא $\mathcal{P}(\mathbf{R})$.

המסקנה הבאה קובעת שהטווח של כל העתקה ליניארית הוא תת־מרחב של המרחב הווקטורי שאליו היא מעתיקה.

3.18 הטווח הוא תת־מרחב

אם $T \in \mathcal{L}(V, W)$, אזי $\operatorname{range} T$ תת־מרחב של W.

הוכחה נניח ש־ $T \in \mathcal{L}(V, W)$. אזי $T(0) = 0$ (לפי 3.10), מה שמוביל למסקנה ש־ $0 \in \operatorname{range} T$.

אם $w_1, w_2 \in \operatorname{range} T$, אזי קיימים $v_1, v_2 \in V$ כך ש־ $Tv_1 = w_1$ ו־ $Tv_2 = w_2$. לפיכך

$$T(v_1 + v_2) = Tv_1 + Tv_2 = w_1 + w_2.$$

לכן $w_1 + w_2 \in \operatorname{range} T$. לפיכך $\operatorname{range} T$ סגור תחת פעולת החיבור.

אם $w \in \operatorname{range} T$ ו־ $\lambda \in \mathbf{F}$, אזי קיים $v \in V$ כך ש־ $Tv = w$. לפיכך

$$T(\lambda v) = \lambda Tv = \lambda w.$$

לכן $\lambda w \in \operatorname{range} T$. לפיכך $\operatorname{range} T$ סגור תחת כפל בסקלר.

הראינו ש־ $\operatorname{range} T$ מכיל את איבר האפס והוא סגור תחת פעולות החיבור והכפל בסקלר. לפיכך $\operatorname{range} T$ הוא תת־מרחב של W (לפי 1.34). ■

3.19 definition: *surjective*

A function $T : V \to W$ is called *surjective* if its range equals W.

To illustrate the definition above, note that of the ranges we computed in 3.17, only the differentiation map is surjective (except that the zero map is surjective in the special case $W = \{0\}$).

Whether a linear map is surjective depends on what we are thinking of as the vector space into which it maps.

*Some people use the term **onto**, which means the same as surjective.*

3.20 example: *surjectivity depends on the target space*

The differentiation map $D \in \mathcal{L}(\mathcal{P}_5(\mathbf{R}))$ defined by $Dp = p'$ is not surjective, because the polynomial x^5 is not in the range of D. However, the differentiation map $S \in \mathcal{L}(\mathcal{P}_5(\mathbf{R}), \mathcal{P}_4(\mathbf{R}))$ defined by $Sp = p'$ is surjective, because its range equals $\mathcal{P}_4(\mathbf{R})$, which is the vector space into which S maps.

Fundamental Theorem of Linear Maps

The next result is so important that it gets a dramatic name.

3.21 *fundamental theorem of linear maps*

Suppose V is finite-dimensional and $T \in \mathcal{L}(V, W)$. Then range T is finite-dimensional and

$$\dim V = \dim \operatorname{null} T + \dim \operatorname{range} T.$$

Proof Let $u_1, \dots, u_m$ be a basis of null T; thus $\dim \operatorname{null} T = m$. The linearly independent list $u_1, \dots, u_m$ can be extended to a basis

$$u_1, \dots, u_m, v_1, \dots, v_n$$

of V (by 2.32). Thus $\dim V = m + n$. To complete the proof, we need to show that range T is finite-dimensional and $\dim \operatorname{range} T = n$. We will do this by proving that $Tv_1, \dots, Tv_n$ is a basis of range T.

Let $v \in V$. Because $u_1, \dots, u_m, v_1, \dots, v_n$ spans V, we can write

$$v = a_1 u_1 + \cdots + a_m u_m + b_1 v_1 + \cdots + b_n v_n,$$

where the a's and b's are in $\mathbf{F}$. Applying T to both sides of this equation, we get

$$Tv = b_1 Tv_1 + \cdots + b_n Tv_n,$$

where the terms of the form Tu_k disappeared because each u_k is in null T. The last equation implies that the list $Tv_1, \dots, Tv_n$ spans range T. In particular, range T is finite-dimensional.

3.19 הגדרה: עַל

פונקציה $T: V \to W$ נקראת עַל אם הטווח שלה שווה ל־ W.

להמחשת ההגדרה לעיל, שימו לב כי מכל הטווחים שחישבנו ב־ 3.17, רק העתקת הגזירה היא עַל (למעט העובדה שהעתקת האפס היא עַל במקרה המיוחד שבו $W = \{0\}$).

ההכרעה האם העתקה ליניארית היא על תלויה במה שאנו רואים כמרחב היעד של ההעתקה.

יש המשתמשים במונח **סורג'קטיבית**, שמשמעותו זהה לזו של עַל.

3.20 דוגמה: תכונת העל תלויה במרחב היעד

העתקת הגזירה $D \in \mathcal{L}(\mathcal{P}_5(\mathbf{R}))$ כשהיא מוגדרת על ידי $Dp = p'$ אינה על מהסיבה שהפולינום x^5 אינו בטווח של D. לעומת זאת, העתקת הגזירה $S \in \mathcal{L}(\mathcal{P}_5(\mathbf{R}), \mathcal{P}_4(\mathbf{R}))$ המוגדרת על ידי $Sp = p'$ היא על, מאחר שהטווח שלה שווה ל־ $\mathcal{P}_4(\mathbf{R})$, שהוא המרחב הווקטורי אליו S מעתיקה.

המשפט היסודי של ההעתקות הליניאריות

המסקנה הבאה היא כֹּה חשובה שהוענק לה שם דרמטי.

3.21 המשפט היסודי של ההעתקות הליניאריות

נניח ש־ V מממד סופי ו־ $T \in \mathcal{L}(V, W)$. אזי $\operatorname{range} T$ מממד סופי, וגם
$$\dim V = \dim \operatorname{null} T + \dim \operatorname{range} T.$$

הוכחה יהי $u_1, \dots, u_m$ בסיס של $\operatorname{null} T$; לפיכך $\dim \operatorname{null} T = m$. את הרשימה הבלתי תלויה ליניארית $u_1, \dots, u_m$ נשלים לבסיס

$$u_1, \dots, u_m, v_1, \dots, v_n$$

של V (לפי 2.32). לפיכך $\dim V = m + n$. להשלמת ההוכחה, עלינו להראות ש־ $\operatorname{range} T$ מממד סופי ו־ $\dim \operatorname{range} T = n$. נעשה זאת בכך שנוכיח ש־ $Tv_1, \dots, Tv_n$ בסיס של $\operatorname{range} T$.

יהי $v \in V$. מאחר ש־ $u_1, \dots, u_m, v_1, \dots, v_n$ פורסת את V, נוכל לכתוב

$$v = a_1 u_1 + \cdots + a_m u_m + b_1 v_1 + \cdots + b_n v_n,$$

כאשר ה־ a־ים וה־ b־ים ב־ $\mathbf{F}$. בהפעלת T על שני אגפי המשוואה, נקבל

$$Tv = b_1 Tv_1 + \cdots + b_n Tv_n,$$

כאשר הביטויים מהצורה Tu_k נעלמים שכן כל u_k ב־ $\operatorname{null} T$. מהמשוואה האחרונה נסיק שהרשימה $Tv_1, \dots, Tv_n$ פורסת את $\operatorname{range} T$. בפרט, $\operatorname{range} T$ מממד סופי.

To show $Tv_1, \dots, Tv_n$ is linearly independent, suppose $c_1, \dots, c_n \in \mathbf{F}$ and

$$c_1 Tv_1 + \cdots + c_n Tv_n = 0.$$

Then

$$T(c_1 v_1 + \cdots + c_n v_n) = 0.$$

Hence

$$c_1 v_1 + \cdots + c_n v_n \in \operatorname{null} T.$$

Because $u_1, \dots, u_m$ spans $\operatorname{null} T$, we can write

$$c_1 v_1 + \cdots + c_n v_n = d_1 u_1 + \cdots + d_m u_m,$$

where the d's are in $\mathbf{F}$. This equation implies that all the c's (and d's) are 0 (because $u_1, \dots, u_m, v_1, \dots, v_n$ is linearly independent). Thus $Tv_1, \dots, Tv_n$ is linearly independent and hence is a basis of range T, as desired. ∎

Now we can show that no linear map from a finite-dimensional vector space to a "smaller" vector space can be injective, where "smaller" is measured by dimension.

3.22 *linear map to a lower-dimensional space is not injective*

Suppose V and W are finite-dimensional vector spaces such that $\dim V > \dim W$. Then no linear map from V to W is injective.

Proof Let $T \in \mathcal{L}(V, W)$. Then

$$\begin{aligned} \dim \operatorname{null} T &= \dim V - \dim \operatorname{range} T \\ &\geq \dim V - \dim W \\ &> 0, \end{aligned}$$

where the first line above comes from the fundamental theorem of linear maps (3.21) and the second line follows from 2.37. The inequality above states that $\dim \operatorname{null} T > 0$. This means that $\operatorname{null} T$ contains vectors other than 0. Thus T is not injective (by 3.15). ∎

3.23 example: *linear map from* $\mathbf{F}^4$ *to* $\mathbf{F}^3$ *is not injective*

Define a linear map $T : \mathbf{F}^4 \to \mathbf{F}^3$ by

$$T(z_1, z_2, z_3, z_4) = \left(\sqrt{7} z_1 + \pi z_2 + z_4, 97 z_1 + 3 z_2 + 2 z_3, z_2 + 6 z_3 + 7 z_4\right).$$

Because $\dim \mathbf{F}^4 > \dim \mathbf{F}^3$, we can use 3.22 to assert that T is not injective, without doing any calculations.

כדי להראות ש־ $Tv_1, \dots, Tv_n$ בלתי תלויה ליניארית, נניח ש־ $c_1, \dots, c_n \in \mathbf{F}$ ומתקיים

$$c_1 Tv_1 + \cdots + c_n Tv_n = 0.$$

כלומר

$$T(c_1 v_1 + \cdots + c_n v_n) = 0.$$

לכן

$$c_1 v_1 + \cdots + c_n v_n \in \operatorname{null} T.$$

מאחר שהרשימה $u_1, \dots, u_m$ פורסת את $\operatorname{null} T$, נוכל להציב

$$c_1 v_1 + \cdots + c_n v_n = d_1 u_1 + \cdots + d_m u_m,$$

כאשר ה־ d־ים ב־ $\mathbf{F}$. ממשוואה זו עולה כי כל ה־ c־ים (וה־ d־ים) שווים ל־ 0 (שכן $u_1, \dots, u_m, v_1, \dots, v_n$ בלתי תלויה ליניארית). לפיכך $Tv_1, \dots, Tv_n$ בלתי תלויה ליניארית, לכן היא בסיס של $\operatorname{range} T$, כפי שרצינו להוכיח. ■

כעת אנו יכולים להראות שהעתקה ליניארית ממרחב וקטורי מממד סופי אל מרחב וקטורי "קטן יותר" אינה יכולה להיות חד־חד־ערכית, כאשר ממדי המרחבים הם המדד ליחס הגודל.

3.22 העתקה ליניארית למרחב מממד נמוך יותר אינה חד־חד־ערכית

נניח ש־ V ו־ W מרחבים וקטוריים מממד סופי כך שמתקיים $\dim V > \dim W$.
אזי לא קיימת העתקה ליניארית חד־חד־ערכית מ־ V אל W.

הוכחה תהי $T \in \mathcal{L}(V, W)$. אזי

$$\begin{aligned} \dim \operatorname{null} T &= \dim V - \dim \operatorname{range} T \\ &\geq \dim V - \dim W \\ &> 0, \end{aligned}$$

כאשר השורה הראשונה נובעת מהמשפט היסודי של ההעתקות הליניאריות (3.21) והשורה השנייה מ־ 2.37. האי־שוויון לעיל קובע כי $\dim \operatorname{null} T > 0$. כלומר, המרחב $\operatorname{null} T$ מכיל וקטורים שאינם 0. לפיכך, T אינה חד־חד־ערכית (לפי 3.15). ■

3.23 דוגמה: העתקה ליניארית מ־ $\mathbf{F}^4$ ל־ $\mathbf{F}^3$ אינה חד־חד־ערכית

נגדיר $T: \mathbf{F}^4 \to \mathbf{F}^3$ על ידי

$$T(z_1, z_2, z_3, z_4) = \left(\sqrt{7} z_1 + \pi z_2 + z_4, 97 z_1 + 3 z_2 + 2 z_3, z_2 + 6 z_3 + 7 z_4\right).$$

מאחר ש־ $\dim \mathbf{F}^4 > \dim \mathbf{F}^3$, נוכל להשתמש ב־ 3.22 ולקבוע ש־ T אינה חד־חד־ערכית, מבלי שנבצע חישובים כלשהם.

The next result shows that no linear map from a finite-dimensional vector space to a "bigger" vector space can be surjective, where "bigger" is measured by dimension.

3.24 *linear map to a higher-dimensional space is not surjective*

Suppose V and W are finite-dimensional vector spaces such that $\dim V < \dim W$. Then no linear map from V to W is surjective.

Proof Let $T \in \mathcal{L}(V, W)$. Then

$$\begin{aligned}\dim \operatorname{range} T &= \dim V - \dim \operatorname{null} T\\ &\le \dim V\\ &< \dim W,\end{aligned}$$

where the equality above comes from the fundamental theorem of linear maps (3.21). The inequality above states that $\dim \operatorname{range} T < \dim W$. This means that $\operatorname{range} T$ cannot equal W. Thus T is not surjective. ∎

As we will soon see, 3.22 and 3.24 have important consequences in the theory of linear equations. The idea is to express questions about systems of linear equations in terms of linear maps. Let's begin by rephrasing in terms of linear maps the question of whether a homogeneous system of linear equations has a nonzero solution.

Fix positive integers m and n, and let $A_{j,k} \in \mathbf{F}$ for $j = 1, \dots, m$ and $k = 1, \dots, n$. Consider the homogeneous system of linear equations

***Homogeneous**, in this context, means that the constant term on the right side of each equation below is 0.*

$$\begin{aligned}\sum_{k=1}^{n} A_{1,k} x_k &= 0\\ &\vdots\\ \sum_{k=1}^{n} A_{m,k} x_k &= 0.\end{aligned}$$

Clearly $x_1 = \cdots = x_n = 0$ is a solution of the system of equations above; the question here is whether any other solutions exist.

Define $T : \mathbf{F}^n \to \mathbf{F}^m$ by

3.25 $$T(x_1, \dots, x_n) = \Big(\sum_{k=1}^{n} A_{1,k} x_k, \dots, \sum_{k=1}^{n} A_{m,k} x_k\Big).$$

The equation $T(x_1, \dots, x_n) = 0$ (the 0 here is the additive identity in $\mathbf{F}^m$, namely, the list of length m of all 0's) is the same as the homogeneous system of linear equations above.

Thus we want to know if $\operatorname{null} T$ is strictly bigger than $\{0\}$, which is equivalent to T not being injective (by 3.15). The next result gives an important condition for ensuring that T is not injective.

המסקנה הבאה קובעת שהעתקה ליניארית ממרחב וקטורי מממד סופי אל מרחב וקטורי "גדול יותר" אינה יכולה להיות עַל, כשממדי המרחבים הם המדד ליחסי הגודל.

3.24 העתקה ליניארית למרחב מממד גבוה יותר אינה עַל

נניח ש־ V ו־ W מרחבים וקטוריים מממד סופי כך שמתקיים $\dim V < \dim W$. אזי לא קיימת העתקה ליניארית עַל מ־ V ל־ W.

הוכחה תהי $T \in \mathcal{L}(V, W)$. אזי

$$\begin{aligned}\dim \operatorname{range} T &= \dim V - \dim \operatorname{null} T \\ &\le \dim V \\ &< \dim W,\end{aligned}$$

כאשר השוויון לעיל מתקבל מהמשפט היסודי של ההעתקות הליניאריות (3.21). האי־שוויון לעיל קובע כי $\dim \operatorname{range} T < \dim W$. משמעות הדבר היא ש־ $\operatorname{range} T$ אינה יכולה להיות שווה ל־ W. לפיכך T אינה עַל. ■

כפי שנראה בקרוב, ל־ 3.22 ו־ 3.24 השלכות חשובות בתורת המשוואות הליניאריות. הרעיון הוא לנסח שאלות על מערכות של משוואות ליניאריות במונחים של העתקות ליניאריות. כצעד ראשון, ננסח מחדש, במונחים של העתקות ליניאריות, את השאלה האם למערכת הומוגנית של משוואות ליניאריות יש פתרון השונה מאפס.

הומוגניות, בהקשר זה, משמעותה שערך האיבר הקבוע באגף הימני של כל משוואה להלן הוא 0.

נְקַבֵּע מספרים שלמים חיוביים m ו־ n, ויהי $A_{j,k} \in \mathbf{F}$ עבור $j = 1, \dots, m$ ו־ $k = 1, \dots, n$. נתבונן במערכת ההומוגנית של המשוואות הליניאריות

$$\begin{aligned}\sum_{k=1}^{n} A_{1,k} x_k &= 0 \\ &\vdots \\ \sum_{k=1}^{n} A_{m,k} x_k &= 0.\end{aligned}$$

ברור ש־ $x_1 = \cdots = x_n = 0$ הוא פתרון למערכת המשוואות לעיל; השאלה היא האם קיימים פתרונות נוספים.

נגדיר $T \colon \mathbf{F}^n \to \mathbf{F}^m$ על ידי

$$T(x_1, \dots, x_n) = \Big(\sum_{k=1}^{n} A_{1,k} x_k, \dots, \sum_{k=1}^{n} A_{m,k} x_k \Big). \tag{3.25}$$

המשוואה $T(x_1, \dots, x_n) = 0$ (ה־ 0 כאן הוא איבר הזהות לחיבור ב־ $\mathbf{F}^m$, כלומר, רשימה בת m איברים שכולם אפסים) היא זהה למערכת ההומוגנית של המשוואות הליניאריות לעיל.

לפיכך, אנו רוצים לדעת האם $\operatorname{null} T$ גדול ממש מ־ $\{0\}$, תנאי השקול לכך ש־ T אינה חד־חד־ערכית (לפי 3.15). המסקנה הבאה מספקת תנאי חשוב המבטיח ש־ T אינה חד־חד־ערכית.

3.26 *homogeneous system of linear equations*

A homogeneous system of linear equations with more variables than equations has nonzero solutions.

Proof Use the notation and result from the discussion above. Thus T is a linear map from $\mathbf{F}^n$ to $\mathbf{F}^m$, and we have a homogeneous system of m linear equations with n variables $x_1, \dots, x_n$. From 3.22 we see that T is not injective if $n > m$. ∎

Example of the result above: a homogeneous system of four linear equations with five variables has nonzero solutions.

Now we consider the question of whether a system of linear equations has no solutions for some choice of the constant terms. To rephrase this question in terms of a linear map, fix positive integers m and n, and let $A_{j,k} \in \mathbf{F}$ for all $j = 1, \dots, m$ and all $k = 1, \dots, n$. For $c_1, \dots, c_m \in \mathbf{F}$, consider the system of linear equations

$$\begin{aligned} \sum_{k=1}^{n} A_{1,k} x_k &= c_1 \\ &\vdots \\ \sum_{k=1}^{n} A_{m,k} x_k &= c_m. \end{aligned} \tag{3.27}$$

With this notation, the question here is whether there is some choice of the constant terms $c_1, \dots, c_m \in \mathbf{F}$ such that no solution exists to the system above.

Define $T : \mathbf{F}^n \to \mathbf{F}^m$ as in 3.25. The equation $T(x_1, \dots, x_n) = (c_1, \dots, c_m)$ is the same as the system of equations 3.27. Thus we want to know if range $T \neq \mathbf{F}^m$. Hence we can rephrase our question about not having a solution for some choice of $c_1, \dots, c_m \in \mathbf{F}$ as follows: What condition ensures that T is not surjective? The next result gives one such condition.

The results 3.26 and 3.28, which compare the number of variables and the number of equations, can also be proved using Gaussian elimination. The abstract approach taken here seems to provide cleaner proofs.

3.28 *system of linear equations with more equations than variables*

A system of linear equations with more equations than variables has no solution for some choice of the constant terms.

Proof Use the notation from the discussion above. Thus T is a linear map from $\mathbf{F}^n$ to $\mathbf{F}^m$, and we have a system of m equations with n variables $x_1, \dots, x_n$; see 3.27. If $n < m$, then 3.24 implies that T is not surjective. As discussed above, this shows that if we have more equations than variables in a system of linear equations, then there is no solution for some choice of the constant terms. ∎

Example of the result above: a system of five linear equations with four variables has no solution for some choice of the constant terms.

3.26 מערכת הומוגנית של משוואות ליניאריות

למערכת הומוגנית של משוואות ליניאריות עם מספר משתנים גדול ממספר המשוואות יש פתרונות השונים מאפס.

הוכחה נשתמש בסימון ובמסקנה מהדיון לעיל. לפיכך T העתקה ליניארית מ־ $\mathbf{F}^n$ אל $\mathbf{F}^m$, וכעת לפנינו מערכת הומוגנית של m משוואות ליניאריות עם n משתנים $x_1, \dots, x_n$. מ־ 3.22 נסיק כי T אינה חד־חד־ערכית אם $n > m$. ■

דוגמה למסקנה לעיל: למערכת הומוגנית של ארבע משוואות ליניאריות עם חמישה משתנים יש פתרונות השונים מאפס.

כעת נתבונן בשאלה האם למערכת משוואות ליניאריות אין פתרונות עבור בחירה מסוימת של האיברים הקבועים. לניסוח מחדש של שאלה זו במונחי העתקה ליניארית, נְקַבֵּעַ מספרים שלמים חיוביים m ו־ n, ויהי $A_{j,k} \in \mathbf{F}$ לכל $j = 1, \dots, m$ ולכל $k = 1, \dots, n$. עבור $c_1, \dots, c_m \in \mathbf{F}$, נתבונן במערכת המשוואות הליניאריות

$$\begin{aligned} \sum_{k=1}^{n} A_{1,k} x_k &= c_1 \\ &\vdots \\ \sum_{k=1}^{n} A_{m,k} x_k &= c_m. \end{aligned} \tag{3.27}$$

עם סימון זה, השאלה העולה היא האם קיימת בחירה כלשהי של האיברים הקבועים $c_1, \dots, c_m \in \mathbf{F}$ כך שלמערכת לעיל לא יהיה פתרון.

נגדיר $T : \mathbf{F}^n \to \mathbf{F}^m$ כמו ב־ 3.25. המשוואה $T(x_1, \dots, x_n) = (c_1, \dots, c_m)$ זהה למערכת המשוואות 3.27. לפיכך אנו מבקשים לדעת האם $\operatorname{range} T \neq \mathbf{F}^m$. לכן נוכל לנסח מחדש את השאלה לגבי היעדר פתרון לקבועים כלשהם $c_1, \dots, c_m \in \mathbf{F}$ באופן הבא: איזה תנאי יבטיח ש־ T אינה עַל? המסקנה הבאה מספקת לנו תנאי שכזה.

את מסקנות 3.26 ו־ 3.28, המשוות בין מספר המשתנים למספר המשוואות, ניתן להוכיח גם בשיטת החילוץ של גאוס. נראה כי הגישה המופשטת שנקטנו כאן מספקת הוכחות נקיות יותר.

3.28 מערכת משוואות ליניאריות עם יותר משוואות ממשתנים

למערכת משוואות ליניאריות עם יותר משוואות ממשתנים אין פתרון עבור בחירה מסוימת של איברים קבועים.

הוכחה נשתמש בסימון מהסעיף לעיל. לפיכך T העתקה ליניארית מ־ $\mathbf{F}^n$ אל $\mathbf{F}^m$, וכעת לפנינו מערכת של m משוואות עם n משתנים $x_1, \dots, x_n$; ראו 3.27. אם $n < m$, אזי 3.24 קובע ש־ T אינה עַל. כפי שנאמר לעיל, משמעות קביעה זו היא שאם יש לנו יותר משוואות ממשתנים במערכת של משוואות ליניאריות, אזי למערכת אין פתרון עבור בחירה מסוימת של איברים קבועים. ■

דוגמה למסקנה לעיל: למערכת של חמש משוואות ליניאריות עם ארבעה משתנים אין פתרון עבור בחירה מסוימת של האיברים הקבועים.

Exercises 3B

1 Give an example of a linear map T with $\dim \operatorname{null} T = 3$ and $\dim \operatorname{range} T = 2$.

2 Suppose $S, T \in \mathcal{L}(V)$ are such that $\operatorname{range} S \subseteq \operatorname{null} T$. Prove that $(ST)^2 = 0$.

3 Suppose $v_1, \dots, v_m$ is a list of vectors in V. Define $T \in \mathcal{L}(\mathbf{F}^m, V)$ by

$$T(z_1, \dots, z_m) = z_1 v_1 + \cdots + z_m v_m.$$

(a) What property of T corresponds to $v_1, \dots, v_m$ spanning V?
(b) What property of T corresponds to the list $v_1, \dots, v_m$ being linearly independent?

4 Show that $\{T \in \mathcal{L}(\mathbf{R}^5, \mathbf{R}^4) : \dim \operatorname{null} T > 2\}$ is not a subspace of $\mathcal{L}(\mathbf{R}^5, \mathbf{R}^4)$.

5 Give an example of $T \in \mathcal{L}(\mathbf{R}^4)$ such that $\operatorname{range} T = \operatorname{null} T$.

6 Prove that there does not exist $T \in \mathcal{L}(\mathbf{R}^5)$ such that $\operatorname{range} T = \operatorname{null} T$.

7 Suppose V and W are finite-dimensional with $2 \le \dim V \le \dim W$. Show that $\{T \in \mathcal{L}(V, W) : T \text{ is not injective}\}$ is not a subspace of $\mathcal{L}(V, W)$.

8 Suppose V and W are finite-dimensional with $\dim V \ge \dim W \ge 2$. Show that $\{T \in \mathcal{L}(V, W) : T \text{ is not surjective}\}$ is not a subspace of $\mathcal{L}(V, W)$.

9 Suppose $T \in \mathcal{L}(V, W)$ is injective and $v_1, \dots, v_n$ is linearly independent in V. Prove that $Tv_1, \dots, Tv_n$ is linearly independent in W.

10 Suppose $v_1, \dots, v_n$ spans V and $T \in \mathcal{L}(V, W)$. Show that $Tv_1, \dots, Tv_n$ spans $\operatorname{range} T$.

11 Suppose that V is finite-dimensional and that $T \in \mathcal{L}(V, W)$. Prove that there exists a subspace U of V such that

$$U \cap \operatorname{null} T = \{0\} \quad \text{and} \quad \operatorname{range} T = \{Tu : u \in U\}.$$

12 Suppose T is a linear map from $\mathbf{F}^4$ to $\mathbf{F}^2$ such that

$$\operatorname{null} T = \{(x_1, x_2, x_3, x_4) \in \mathbf{F}^4 : x_1 = 5x_2 \text{ and } x_3 = 7x_4\}.$$

Prove that T is surjective.

13 Suppose U is a three-dimensional subspace of $\mathbf{R}^8$ and that T is a linear map from $\mathbf{R}^8$ to $\mathbf{R}^5$ such that $\operatorname{null} T = U$. Prove that T is surjective.

14 Prove that there does not exist a linear map from $\mathbf{F}^5$ to $\mathbf{F}^2$ whose null space equals $\{(x_1, x_2, x_3, x_4, x_5) \in \mathbf{F}^5 : x_1 = 3x_2 \text{ and } x_3 = x_4 = x_5\}$.

15 Suppose there exists a linear map on V whose null space and range are both finite-dimensional. Prove that V is finite-dimensional.

תרגילים ג

1 הביאו דוגמה להעתקה ליניארית T עם $\dim \operatorname{null} T = 3$ ו־ $\dim \operatorname{range} T = 2$.

2 נניח ש־ $S, T \in \mathcal{L}(V)$ כך שמתקיים $\operatorname{range} S \subseteq \operatorname{null} T$. הוכיחו ש־ $(ST)^2 = 0$.

3 נניח ש־ $v_1, \dots, v_m$ היא רשימת וקטורים ב־ V. הגדירו $T \in \mathcal{L}(\mathbf{F}^m, V)$ על ידי
$$T(z_1, \dots, z_m) = z_1 v_1 + \cdots + z_m v_m.$$
א. איזו תכונה של T מתאימה לכך ש־ $v_1, \dots, v_m$ פורסת את V?
ב. איזו תכונה של T מתאימה לכך שהרשימה $v_1, \dots, v_m$ בלתי תלויה ליניארית?

4 הראו ש־ $\{T \in \mathcal{L}(\mathbf{R}^5, \mathbf{R}^4) : \dim \operatorname{null} T > 2\}$ אינו תת־מרחב של $\mathcal{L}(\mathbf{R}^5, \mathbf{R}^4)$.

5 הביאו דוגמה ל־ $T \in \mathcal{L}(\mathbf{R}^4)$ כך ש־ $\operatorname{range} T = \operatorname{null} T$.

6 הוכיחו שלא קיימת $T \in \mathcal{L}(\mathbf{R}^5)$ כך ש־ $\operatorname{range} T = \operatorname{null} T$.

7 נניח ש־ V ו־ W מרחבים מממד סופי המקיימים $2 \le \dim V \le \dim W$. הראו ש־ $\{T \in \mathcal{L}(V, W) : T \text{ אינה חד־חד־ערכית}\}$ אינו תת־מרחב של $\mathcal{L}(V, W)$.

8 נניח ש־ V ו־ W מרחבים מממד סופי ומתקיים $\dim V \ge \dim W \ge 2$. הראו כי $\{T \in \mathcal{L}(V, W) : T \text{ אינה על}\}$ אינו תת־מרחב של $\mathcal{L}(V, W)$.

9 נניח ש־ $T \in \mathcal{L}(V, W)$ חד־חד־ערכית ו־ $v_1, \dots, v_n$ בלתי תלויה ליניארית ב־ V. הוכיחו כי $Tv_1, \dots, Tv_n$ בלתי תלויה ליניארית ב־ W.

10 נניח ש־ $v_1, \dots, v_n$ פורסת את V ו־ $T \in \mathcal{L}(V, W)$. הראו ש־ $Tv_1, \dots, Tv_n$ פורסת את $\operatorname{range} T$.

11 נניח ש־ V מרחב מממד סופי ו־ $T \in \mathcal{L}(V, W)$. הוכיחו כי קיים תת־מרחב U של V כך שמתקיים
$$\operatorname{range} T = \{Tu : u \in U\} \quad \text{וגם} \quad U \cap \operatorname{null} T = \{0\}.$$

12 נניח ש־ T העתקה ליניארית מ־ $\mathbf{F}^4$ אל $\mathbf{F}^2$ כך ש־
$$\operatorname{null} T = \{(x_1, x_2, x_3, x_4) \in \mathbf{F}^4 : x_1 = 5x_2 \text{ and } x_3 = 7x_4\}.$$
הוכיחו כי T על.

13 נניח ש־ U תת־מרחב תלת־ממדי של $\mathbf{R}^8$ ו־ T העתקה ליניארית מ־ $\mathbf{R}^8$ אל $\mathbf{R}^5$ כך ש־ $\operatorname{null} T = U$. הוכיחו כי T על.

14 הוכיחו כי לא קיימת העתקה ליניארית מ־ $\mathbf{F}^5$ אל $\mathbf{F}^2$ כזו שמרחב האפס שלה שווה ל־ $\{(x_1, x_2, x_3, x_4, x_5) \in \mathbf{F}^5 : x_1 = 3x_2 \quad \text{וגם} \quad x_3 = x_4 = x_5\}$.

15 נניח שקיימת העתקה ליניארית מעל V שמרחב האפס שלה והטווח שניהם מרחבים מממד סופי. הוכיחו כי V מממד סופי.

16 Suppose V and W are both finite-dimensional. Prove that there exists an injective linear map from V to W if and only if $\dim V \le \dim W$.

17 Suppose V and W are both finite-dimensional. Prove that there exists a surjective linear map from V onto W if and only if $\dim V \ge \dim W$.

18 Suppose V and W are finite-dimensional and that U is a subspace of V. Prove that there exists $T \in \mathcal{L}(V, W)$ such that $\operatorname{null} T = U$ if and only if $\dim U \ge \dim V - \dim W$.

19 Suppose W is finite-dimensional and $T \in \mathcal{L}(V, W)$. Prove that T is injective if and only if there exists $S \in \mathcal{L}(W, V)$ such that ST is the identity operator on V.

20 Suppose W is finite-dimensional and $T \in \mathcal{L}(V, W)$. Prove that T is surjective if and only if there exists $S \in \mathcal{L}(W, V)$ such that TS is the identity operator on W.

21 Suppose V is finite-dimensional, $T \in \mathcal{L}(V, W)$, and U is a subspace of W. Prove that $\{v \in V : Tv \in U\}$ is a subspace of V and

$$\dim\{v \in V : Tv \in U\} = \dim \operatorname{null} T + \dim(U \cap \operatorname{range} T).$$

22 Suppose U and V are finite-dimensional vector spaces and $S \in \mathcal{L}(V, W)$ and $T \in \mathcal{L}(U, V)$. Prove that

$$\dim \operatorname{null} ST \le \dim \operatorname{null} S + \dim \operatorname{null} T.$$

23 Suppose U and V are finite-dimensional vector spaces and $S \in \mathcal{L}(V, W)$ and $T \in \mathcal{L}(U, V)$. Prove that

$$\dim \operatorname{range} ST \le \min\{\dim \operatorname{range} S, \dim \operatorname{range} T\}.$$

24 (a) Suppose $\dim V = 5$ and $S, T \in \mathcal{L}(V)$ are such that $ST = 0$. Prove that $\dim \operatorname{range} TS \le 2$.

(b) Give an example of $S, T \in \mathcal{L}(\mathbf{F}^5)$ with $ST = 0$ and $\dim \operatorname{range} TS = 2$.

25 Suppose that W is finite-dimensional and $S, T \in \mathcal{L}(V, W)$. Prove that $\operatorname{null} S \subseteq \operatorname{null} T$ if and only if there exists $E \in \mathcal{L}(W)$ such that $T = ES$.

26 Suppose that V is finite-dimensional and $S, T \in \mathcal{L}(V, W)$. Prove that $\operatorname{range} S \subseteq \operatorname{range} T$ if and only if there exists $E \in \mathcal{L}(V)$ such that $S = TE$.

27 Suppose $P \in \mathcal{L}(V)$ and $P^2 = P$. Prove that $V = \operatorname{null} P \oplus \operatorname{range} P$.

28 Suppose $D \in \mathcal{L}(\mathcal{P}(\mathbf{R}))$ is such that $\deg Dp = (\deg p) - 1$ for every non-constant polynomial $p \in \mathcal{P}(\mathbf{R})$. Prove that D is surjective.

The notation D is used above to remind you of the differentiation map that sends a polynomial p to p'.

16 נניח ש־ V ו־ W שניהם מממד סופי. הוכיחו כי קיימת העתקה ליניארית חד־חד־ערכית מ־ V אל W אם־ורק־אם $\dim V \le \dim W$.

17 נניח ש־ V ו־ W שניהם מממד סופי. הוכיחו כי קיימת העתקה ליניארית עַל מ־ V ל־ W אם־ורק־אם $\dim V \ge \dim W$.

18 נניח ש־ V ו־ W מרחבים מממד סופי, ו־ U הוא תת־מרחב של V. הוכיחו כי קיימת העתקה $T \in \mathcal{L}(V,W)$ כך ש־ $\operatorname{null} T = U$ אם־ורק־אם $\dim U \ge \dim V - \dim W$.

19 נניח ש־ W מממד סופי ו־ $T \in \mathcal{L}(V,W)$. הוכיחו כי T היא חד־חד־ערכית אם־ורק־אם קיימת $S \in \mathcal{L}(W,V)$ כך ש־ ST הוא אופרטור הזהות מעל V.

20 נניח ש־ W מממד סופי ו־ $T \in \mathcal{L}(V,W)$. הוכיחו כי T עַל אם־ורק־אם קיימת $S \in \mathcal{L}(W,V)$ כך ש־ TS הוא אופרטור הזהות מעל W.

21 נניח ש־ V מממד סופי, $T \in \mathcal{L}(V,W)$, ו־ U תת־מרחב של W. הוכיחו כי $\{v \in V : Tv \in U\}$ תת־מרחב של V וגם

$$\dim\{v \in V : Tv \in U\} = \dim \operatorname{null} T + \dim(U \cap \operatorname{range} T).$$

22 נניח ש־ U ו־ V מרחבים וקטוריים מממד סופי ו־ $S \in \mathcal{L}(V,W)$ ו־ $T \in \mathcal{L}(U,V)$. הוכיחו כי

$$\dim \operatorname{null} ST \le \dim \operatorname{null} S + \dim \operatorname{null} T.$$

23 נניח ש־ U ו־ V מרחבים וקטוריים מממד סופי ו־ $S \in \mathcal{L}(V,W)$ ו־ $T \in \mathcal{L}(U,V)$. הוכיחו כי

$$\dim \operatorname{range} ST \le \min\{\dim \operatorname{range} S, \dim \operatorname{range} T\}.$$

24 א. נניח ש־ $\dim V = 5$ ו־ $S,T \in \mathcal{L}(V)$ כך ש־ $ST = 0$. הוכיחו כי $\dim \operatorname{range} TS \le 2$.

ב. הביאו דוגמה של $S,T \in \mathcal{L}(\mathbf{F}^5)$ כאשר $ST = 0$ ו־ $\dim \operatorname{range} TS = 2$.

25 נניח ש־ W מממד סופי ו־ $S,T \in \mathcal{L}(V,W)$. הוכיחו כי $\operatorname{null} S \subseteq \operatorname{null} T$ אם־ורק־אם קיימת $E \in \mathcal{L}(W)$ כך ש־ $T = ES$.

26 נניח ש־ V מממד סופי ו־ $S,T \in \mathcal{L}(V,W)$. הוכיחו כי $\operatorname{range} S \subseteq \operatorname{range} T$ אם־ורק־אם קיימת העתקה $E \in \mathcal{L}(V)$ כך ש־ $S = TE$.

27 נניח ש־ $P \in \mathcal{L}(V)$ ו־ $P^2 = P$. הוכיחו כי $V = \operatorname{null} P \oplus \operatorname{range} P$.

28 נניח ש־ $D \in \mathcal{L}(\mathcal{P}(\mathbf{R}))$ העתקה כך ש־ $\deg Dp = (\deg p) - 1$ לכל פולינום שאינו קבוע $p \in \mathcal{P}(\mathbf{R})$. הוכיחו ש־ D על.

השימוש בסימון D נועד להזכיר לכם את העתקת הגזירה השולחת את הפולינום p אל p'.

29 Suppose $p \in \mathcal{P}(\mathbf{R})$. Prove that there exists a polynomial $q \in \mathcal{P}(\mathbf{R})$ such that $5q'' + 3q' = p$.

This exercise can be done without linear algebra, but it's more fun to do it using linear algebra.

30 Suppose $\varphi \in \mathcal{L}(V, \mathbf{F})$ and $\varphi \neq 0$. Suppose $u \in V$ is not in $\operatorname{null} \varphi$. Prove that

$$V = \operatorname{null} \varphi \oplus \{au : a \in \mathbf{F}\}.$$

31 Suppose V is finite-dimensional, X is a subspace of V, and Y is a finite-dimensional subspace of W. Prove that there exists $T \in \mathcal{L}(V, W)$ such that $\operatorname{null} T = X$ and $\operatorname{range} T = Y$ if and only if $\dim X + \dim Y = \dim V$.

32 Suppose V is finite-dimensional with $\dim V > 1$. Show that if $\varphi : \mathcal{L}(V) \to \mathbf{F}$ is a linear map such that $\varphi(ST) = \varphi(S)\varphi(T)$ for all $S, T \in \mathcal{L}(V)$, then $\varphi = 0$.

Hint: The description of the two-sided ideals of $\mathcal{L}(V)$ given by Exercise 17 in Section 3A might be useful.

33 Suppose that V and W are real vector spaces and $T \in \mathcal{L}(V, W)$. Define $T_{\mathbf{C}} : V_{\mathbf{C}} \to W_{\mathbf{C}}$ by

$$T_{\mathbf{C}}(u + iv) = Tu + iTv$$

for all $u, v \in V$.

(a) Show that $T_{\mathbf{C}}$ is a (complex) linear map from $V_{\mathbf{C}}$ to $W_{\mathbf{C}}$.
(b) Show that $T_{\mathbf{C}}$ is injective if and only if T is injective.
(c) Show that $\operatorname{range} T_{\mathbf{C}} = W_{\mathbf{C}}$ if and only if $\operatorname{range} T = W$.

*See Exercise 8 in Section 1B for the definition of the complexification $V_{\mathbf{C}}$. The linear map $T_{\mathbf{C}}$ is called the **complexification** of the linear map T.*

29 נניח ש־ $p \in \mathcal{P}(\mathbf{R})$. הוכיחו כי קיים פולינום $q \in \mathcal{P}(\mathbf{R})$ כך ש־ $5q'' + 3q' = p$.

אף שניתן לפתור תרגיל זה ללא אלגברה ליניארית, שימוש באלגברה ליניארית יוביל לפתרון מהנה יותר.

30 נניח ש־ $\varphi \in \mathcal{L}(V, \mathbf{F})$ וגם $\varphi \neq 0$. בנוסף, נניח ש־ $u \in V$ וגם $u \notin \operatorname{null} \varphi$. הוכיחו כי

$$V = \operatorname{null} \varphi \oplus \{au : a \in \mathbf{F}\}.$$

31 נניח ש־ V מממד סופי, X תת־מרחב של V, ו־ Y תת־מרחב מממד סופי של W. הוכיחו כי קיימת $T \in \mathcal{L}(V, W)$ כך ש־ $\operatorname{null} T = X$ ו־ $\operatorname{range} T = Y$ אם־ורק־אם $\dim X + \dim Y = \dim V$.

32 נניח ש־ V מממד סופי שעבורו מתקיים $\dim V > 1$. הראו שאם $\varphi : \mathcal{L}(V) \to \mathbf{F}$ היא העתקה ליניארית כך ש־ $\varphi(ST) = \varphi(S)\varphi(T)$ לכל $S, T \in \mathcal{L}(V)$, אזי $\varphi = 0$.

רמז: התיאור של אידיאלים דו־צדדיים של $\mathcal{L}(V)$ שניתן בתרגיל 17 ביחידה ג1 יכול לסייע בפתרון.

33 נניח ש־ V ו־ W מרחבים וקטוריים ממשיים וגם $T \in \mathcal{L}(V, W)$. הגדירו $T_{\mathbf{C}} : V_{\mathbf{C}} \to W_{\mathbf{C}}$ על ידי

$$T_{\mathbf{C}}(u + iv) = Tu + iTv$$

לכל $u, v \in V$.

א. הראו כי $T_{\mathbf{C}}$ העתקה ליניארית מרוכבת מ־ $V_{\mathbf{C}}$ אל $W_{\mathbf{C}}$.

ב. הראו כי $T_{\mathbf{C}}$ חד־חד־ערכית אם־ורק־אם T חד־חד־ערכית.

ג. הראו כי $\operatorname{range} T_{\mathbf{C}} = W_{\mathbf{C}}$ אם־ורק־אם $\operatorname{range} T = W$.

ראו תרגיל 8 ביחידה א2 להגדרת המירכוב $V_{\mathbf{C}}$. ההעתקה הליניארית $T_{\mathbf{C}}$ נקראת **המירכוב** של ההעתקה הליניארית T.

3C Matrices

Representing a Linear Map by a Matrix

We know that if $v_1, \dots, v_n$ is a basis of V and $T : V \to W$ is linear, then the values of $Tv_1, \dots, Tv_n$ determine the values of T on arbitrary vectors in V—see the linear map lemma (3.4). As we will soon see, matrices provide an efficient method of recording the values of the Tv_k's in terms of a basis of W.

3.29 definition: *matrix*, $A_{j,k}$

Suppose m and n are nonnegative integers. An m-by-n *matrix* A is a rectangular array of elements of $\mathbf{F}$ with m rows and n columns:

$$A = \begin{pmatrix} A_{1,1} & \cdots & A_{1,n} \\ \vdots & & \vdots \\ A_{m,1} & \cdots & A_{m,n} \end{pmatrix}.$$

The notation $A_{j,k}$ denotes the entry in row j, column k of A.

3.30 example: $A_{j,k}$ *equals entry in row j, column k of A*

Suppose $A = \begin{pmatrix} 8 & 4 & 5-3i \\ 1 & 9 & 7 \end{pmatrix}$.

Thus $A_{2,3}$ refers to the entry in the second row, third column of A, which means that $A_{2,3} = 7$.

When dealing with matrices, the first index refers to the row number; the second index refers to the column number.

Now we come to the key definition in this section.

3.31 definition: *matrix of a linear map*, $\mathcal{M}(T)$

Suppose $T \in \mathcal{L}(V, W)$ and $v_1, \dots, v_n$ is a basis of V and $w_1, \dots, w_m$ is a basis of W. The *matrix of* T with respect to these bases is the m-by-n matrix $\mathcal{M}(T)$ whose entries $A_{j,k}$ are defined by

$$Tv_k = A_{1,k}w_1 + \cdots + A_{m,k}w_m.$$

If the bases $v_1, \dots, v_n$ and $w_1, \dots, w_m$ are not clear from the context, then the notation $\mathcal{M}(T, (v_1, \dots, v_n), (w_1, \dots, w_m))$ is used.

The matrix $\mathcal{M}(T)$ of a linear map $T \in \mathcal{L}(V, W)$ depends on the basis $v_1, \dots, v_n$ of V and the basis $w_1, \dots, w_m$ of W, as well as on T. However, the bases should be clear from the context, and thus they are often not included in the notation.

To remember how $\mathcal{M}(T)$ is constructed from T, you might write across the top of the matrix the basis vectors $v_1, \dots, v_n$ for the domain and along the left the basis vectors $w_1, \dots, w_m$ for the vector space into which T maps, as follows:

ג3 מטריצות

ייצוג העתקה ליניארית באמצעות מטריצה

אנו כבר יודעים שאם $T: V \to W$ ליניארית, אזי הערכים של T בסיס של V ו־ $v_1, \ldots, v_n$ קובעים את ערכי T מעל וקטורים שרירותיים ב־ V—ראו לְמָּת ההעתקה הליניארית (3.4). כפי שנראה מייד, מטריצות מספקות שיטה יעילה לרישום ערכי Tv_k במונחי בסיס של W.

3.29 הגדרה: מטריצה, $A_{j,k}$

נניח ש־ m ו־ n מספרים שלמים אי־שליליים. מטריצה בגודל $m \times n$ שנסמן ב־ A היא מערך מלבני של איברים ב־ $\mathbf{F}$ עם m שורות ו־ n עמודות:

$$A = \begin{pmatrix} A_{1,1} & \cdots & A_{1,n} \\ \vdots & & \vdots \\ A_{m,1} & \cdots & A_{m,n} \end{pmatrix}.$$

הסימון $A_{j,k}$ מסמן את האיבר בשורה j, בעמודה k של A.

3.30 דוגמה: $A_{j,k}$ הוא האיבר בשורה j, עמודה k של A

נניח ש־ $A = \begin{pmatrix} 8 & 4 & 5-3i \\ 1 & 9 & 7 \end{pmatrix}$.

כשאנו דנים במטריצות, האינדקס הראשון מתייחס למספר השורה; האינדקס השני מתייחס למספר העמודה.

לפיכך $A_{2,3}$ מתייחס לאיבר בשורה השנייה, בעמודה השלישית של A, ומשמעות הדבר היא ש־ $A_{2,3} = 7$.

כעת אנו מגיעים להגדרת יסוד ביחידה זו.

3.31 הגדרה: מטריצה של העתקה ליניארית, $\mathcal{M}(T)$

נניח ש־ $T \in \mathcal{L}(V, W)$ ו־ $v_1, \ldots, v_n$ בסיס של V ו־ $w_1, \ldots, w_m$ בסיס של W. המטריצה של T ביחס לבסיסים אלו היא $\mathcal{M}(T)$ מטריצה $m \times n$ שאיבריה $A_{j,k}$ מוגדרים על ידי

$$Tv_k = A_{1,k}w_1 + \cdots + A_{m,k}w_m.$$

אם הבסיסים $v_1, \ldots, v_n$ ו־ $w_1, \ldots, w_m$ אינם ברורים מההקשר, אזי נעשה שימוש בסימון $\mathcal{M}(T, (v_1, \ldots, v_n), (w_1, \ldots, w_m))$.

המטריצה $\mathcal{M}(T)$ של ההעתקה הליניארית $T \in \mathcal{L}(V, W)$ תלויה בבסיס $v_1, \ldots, v_n$ של V ובבסיס $w_1, \ldots, w_m$ של W, ובנוסף ב־ T. אולם, מההקשר אנו אמורים להבין את הגדרת הבסיסים, ולפיכך לרוב לא נכלול את הבסיסים בסימון.

כדי לזכור איך $\mathcal{M}(T)$ נִבְנֵית מתוך T, ייתכן שתרצו לכתוב בראש המטריצה לאורכה את וקטורי הבסיס $v_1, \ldots, v_n$ עבור התחום ולאורך צידה השמאלי את וקטורי הבסיס $w_1, \ldots, w_m$ עבור המרחב הווקטורי שאליו T מעתיקה, באופן הבא:

$$\mathcal{M}(T) = \begin{matrix} \\ w_1 \\ \vdots \\ w_m \end{matrix} \begin{matrix} \begin{matrix} v_1 & \cdots & v_k & \cdots & v_n \end{matrix} \\ \begin{pmatrix} & & A_{1,k} & & \\ & & \vdots & & \\ & & A_{m,k} & & \end{pmatrix} \end{matrix}.$$

In the matrix above only the kth column is shown. Thus the second index of each displayed entry of the matrix above is k. The picture above should remind you that Tv_k can be computed from $\mathcal{M}(T)$ by multiplying each entry in the kth column by the corresponding w_j from the left column, and then adding up the resulting vectors.

The kth column of $\mathcal{M}(T)$ consists of the scalars needed to write Tv_k as a linear combination of $w_1, \dots, w_m$:

$$Tv_k = \sum_{j=1}^{m} A_{j,k} w_j.$$

If T is a linear map from $\mathbf{F}^n$ to $\mathbf{F}^m$, then unless stated otherwise, assume the bases in question are the standard ones (where the kth basis vector is 1 in the kth slot and 0 in all other slots). If you think of elements of $\mathbf{F}^m$ as columns of m numbers, then you can think of the kth column of $\mathcal{M}(T)$ as T applied to the kth standard basis vector.

If T is a linear map from an n-dimensional vector space to an m-dimensional vector space, then $\mathcal{M}(T)$ is an m-by-n matrix.

3.32 example: *the matrix of a linear map from $\mathbf{F}^2$ to $\mathbf{F}^3$*

Suppose $T \in \mathcal{L}(\mathbf{F}^2, \mathbf{F}^3)$ is defined by

$$T(x, y) = (x + 3y, 2x + 5y, 7x + 9y).$$

Because $T(1, 0) = (1, 2, 7)$ and $T(0, 1) = (3, 5, 9)$, the matrix of T with respect to the standard bases is the 3-by-2 matrix below:

$$\mathcal{M}(T) = \begin{pmatrix} 1 & 3 \\ 2 & 5 \\ 7 & 9 \end{pmatrix}.$$

When working with $\mathcal{P}_m(\mathbf{F})$, use the standard basis $1, x, x^2, \dots, x^m$ unless the context indicates otherwise.

3.33 example: *matrix of the differentiation map from $\mathcal{P}_3(\mathbf{R})$ to $\mathcal{P}_2(\mathbf{R})$*

Suppose $D \in \mathcal{L}(\mathcal{P}_3(\mathbf{R}), \mathcal{P}_2(\mathbf{R}))$ is the differentiation map defined by $Dp = p'$. Because $(x^n)' = nx^{n-1}$, the matrix of D with respect to the standard bases is the 3-by-4 matrix below:

$$\mathcal{M}(D) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

$$\mathcal{M}(T) = \begin{matrix} \\ w_1 \\ \vdots \\ w_m \end{matrix} \begin{matrix} \begin{matrix} v_1 & \cdots & v_k & \cdots & v_n \end{matrix} \\ \begin{pmatrix} & & A_{1,k} & & \\ & & \vdots & & \\ & & A_{m,k} & & \end{pmatrix} \end{matrix}.$$

במטריצה לעיל הצגנו רק את העמודה ה־ k^{th}. לפיכך האינדקס השני של כל איבר שמוצג הוא k. השרטוט לעיל אמור להזכיר לנו שניתן לחשב את Tv_k מתוך $\mathcal{M}(T)$ בהכפלת כל איבר בעמודה ה־ k^{th} ב־ w_j התואם מהעמודה השמאלית, ואז חיבור של וקטורי התוצאה.

העמודה ה־ k^{th} של $\mathcal{M}(T)$ מורכבת מהסקלרים הנדרשים לבטא את Tv_k כצירוף ליניארי של $w_1, \dots, w_m$:

$$Tv_k = \sum_{j=1}^{m} A_{j,k} w_j.$$

אם T העתקה ליניארית ממרחב וקטורי n־ממדי אל מרחב וקטורי m־ממדי, אזי $\mathcal{M}(T)$ היא מטריצה $m \times n$.

אם T העתקה ליניארית מ־ $\mathbf{F}^n$ אל $\mathbf{F}^m$, אזי אם לא צוין אחרת, עלינו להניח כי הבסיסים הנדונים הם הבסיסים הסטנדרטיים (כאשר בווקטור הבסיס ה־ k^{th} הערך ברכיב ה־ k^{th} הוא 1 ובכל שאר הרכיבים 0). אם תחשבו על איברים ב־ $\mathbf{F}^m$ כעמודות עם m מספרים, אזי העמודה ה־ k^{th} של $\mathcal{M}(T)$ היא התוצאה כאשר T מופעלת על הווקטור ה־ k^{th} בבסיס הסטנדרטי.

3.32 דוגמה: מטריצת ההעתקה הליניארית מ־ $\mathbf{F}^2$ אל $\mathbf{F}^3$

נניח ש־ $T \in \mathcal{L}(\mathbf{F}^2, \mathbf{F}^3)$ מוגדרת על ידי

$$T(x, y) = (x + 3y, 2x + 5y, 7x + 9y).$$

מאחר ש־ $T(1, 0) = (1, 2, 7)$ ו־ $T(0, 1) = (3, 5, 9)$, המטריצה של T ביחס לבסיסים הסטנדרטיים היא המטריצה 3×2 להלן:

$$\mathcal{M}(T) = \begin{pmatrix} 1 & 3 \\ 2 & 5 \\ 7 & 9 \end{pmatrix}.$$

בעבודה עם $\mathcal{P}_m(\mathbf{F})$, השימוש יהיה בבסיס הסטנדרטי $1, x, x^2, \dots, x^m$ אלא אם כן עולה אחרת מההקשר.

3.33 דוגמה: המטריצה של העתקת הגזירה מ־ $\mathcal{P}_3(\mathbf{R})$ אל $\mathcal{P}_2(\mathbf{R})$

נניח ש־ $D \in \mathcal{L}(\mathcal{P}_3(\mathbf{R}), \mathcal{P}_2(\mathbf{R}))$ העתקת הגזירה המוגדרת על ידי $Dp = p'$. מאחר ש־ $(x^n)' = nx^{n-1}$, המטריצה של D ביחס לבסיסים הסטנדרטיים היא מטריצה 3×4 כדלהלן:

$$\mathcal{M}(D) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

Addition and Scalar Multiplication of Matrices

For the rest of this section, assume that U, V, and W are finite-dimensional and that a basis has been chosen for each of these vector spaces. Thus for each linear map from V to W, we can talk about its matrix (with respect to the chosen bases).

Is the matrix of the sum of two linear maps equal to the sum of the matrices of the two maps? Right now this question does not yet make sense because although we have defined the sum of two linear maps, we have not defined the sum of two matrices. Fortunately, the natural definition of the sum of two matrices has the right properties. Specifically, we make the following definition.

3.34 definition: *matrix addition*

The *sum of two matrices of the same size* is the matrix obtained by adding corresponding entries in the matrices:

$$\begin{pmatrix} A_{1,1} & \cdots & A_{1,n} \\ \vdots & & \vdots \\ A_{m,1} & \cdots & A_{m,n} \end{pmatrix} + \begin{pmatrix} C_{1,1} & \cdots & C_{1,n} \\ \vdots & & \vdots \\ C_{m,1} & \cdots & C_{m,n} \end{pmatrix}$$
$$= \begin{pmatrix} A_{1,1} + C_{1,1} & \cdots & A_{1,n} + C_{1,n} \\ \vdots & & \vdots \\ A_{m,1} + C_{m,1} & \cdots & A_{m,n} + C_{m,n} \end{pmatrix}.$$

In the next result, the assumption is that the same bases are used for all three linear maps $S + T$, S, and T.

3.35 *matrix of the sum of linear maps*

Suppose $S, T \in \mathcal{L}(V, W)$. Then $\mathcal{M}(S + T) = \mathcal{M}(S) + \mathcal{M}(T)$.

The verification of the result above follows from the definitions and is left to the reader.

Still assuming that we have some bases in mind, is the matrix of a scalar times a linear map equal to the scalar times the matrix of the linear map? Again, the question does not yet make sense because we have not defined scalar multiplication on matrices. Fortunately, the natural definition again has the right properties.

3.36 definition: *scalar multiplication of a matrix*

The product of a scalar and a matrix is the matrix obtained by multiplying each entry in the matrix by the scalar:

$$\lambda \begin{pmatrix} A_{1,1} & \cdots & A_{1,n} \\ \vdots & & \vdots \\ A_{m,1} & \cdots & A_{m,n} \end{pmatrix} = \begin{pmatrix} \lambda A_{1,1} & \cdots & \lambda A_{1,n} \\ \vdots & & \vdots \\ \lambda A_{m,1} & \cdots & \lambda A_{m,n} \end{pmatrix}.$$

חיבור וכפל בסקלר של מטריצות

מכאן ואילך עד לסוף יחידה זו, הניחו ש־ U, V, ו־ W מרחבים מממד סופי ובנוסף שנבחרו כבר בסיסים עבור מרחבים אלו. לפיכך, עבור כל העתקה ליניארית מ־ V ל־ W, נוכל לדון במטריצה שלה (ביחס לבסיסים שנבחרו).

האם מטריצת הסכום של שתי העתקות ליניאריות שווה לסכום המטריצות של שתי ההעתקות? בשלב זה לשאלה זו אין עדיין משמעות, משום שאף שכבר הגדרנו סכום של שתי העתקות ליניאריות, עדיין לא הגדרנו סכום של מטריצות. למזלנו, להגדרה הטבעית של סכום שתי מטריצות יש את התכונות המתאימות. בפרט, נגדיר כדלהלן:

3.34 הגדרה: חיבור מטריצות

הסכום של שתי מטריצות מאותו הגודל הוא המטריצה המתקבלת באמצעות חיבור איברים תואמים בין המטריצות:

$$\begin{pmatrix} A_{1,1} & \cdots & A_{1,n} \\ \vdots & & \vdots \\ A_{m,1} & \cdots & A_{m,n} \end{pmatrix} + \begin{pmatrix} C_{1,1} & \cdots & C_{1,n} \\ \vdots & & \vdots \\ C_{m,1} & \cdots & C_{m,n} \end{pmatrix}$$

$$= \begin{pmatrix} A_{1,1} + C_{1,1} & \cdots & A_{1,n} + C_{1,n} \\ \vdots & & \vdots \\ A_{m,1} + C_{m,1} & \cdots & A_{m,n} + C_{m,n} \end{pmatrix}.$$

במסקנה הבאה, ההנחה היא שנעשה שימוש באותם בסיסים עבור כל שלוש ההעתקות S, $S + T$, ו־ T.

3.35 המטריצה של סכום העתקות ליניאריות

נניח ש־ $S, T \in \mathcal{L}(V, W)$. אזי $\mathcal{M}(S + T) = \mathcal{M}(S) + \mathcal{M}(T)$.

הווידוא שהמסקנה לעיל נובעת מההגדרות מושאר בידי הקוראים.

עדיין תחת הנחת העבודה שהוגדרו בסיסים, נשאל את השאלה הבאה: האם המטריצה של העתקה המוכפלת בסקלר שווה למכפלת הסקלר במטריצה של ההעתקה? ושוב, בשלב זה לשאלה זו אין עדיין משמעות, משום שעדיין לא הגדרנו כפל מטריצה בסקלר. למזלנו, להגדרה הטבעית יש את התכונות המתאימות.

3.36 הגדרה: כפל סקלר במטריצה

מכפלת סקלר במטריצה היא המטריצה שמתקבלת בהכפלת כל איבר במטריצה בסקלר:

$$\lambda \begin{pmatrix} A_{1,1} & \cdots & A_{1,n} \\ \vdots & & \vdots \\ A_{m,1} & \cdots & A_{m,n} \end{pmatrix} = \begin{pmatrix} \lambda A_{1,1} & \cdots & \lambda A_{1,n} \\ \vdots & & \vdots \\ \lambda A_{m,1} & \cdots & \lambda A_{m,n} \end{pmatrix}.$$

3.37 example: *addition and scalar multiplication of matrices*

$$2\begin{pmatrix} 3 & 1 \\ -1 & 5 \end{pmatrix} + \begin{pmatrix} 4 & 2 \\ 1 & 6 \end{pmatrix} = \begin{pmatrix} 6 & 2 \\ -2 & 10 \end{pmatrix} + \begin{pmatrix} 4 & 2 \\ 1 & 6 \end{pmatrix} = \begin{pmatrix} 10 & 4 \\ -1 & 16 \end{pmatrix}$$

In the next result, the assumption is that the same bases are used for both the linear maps λT and T.

3.38 *the matrix of a scalar times a linear map*

Suppose $\lambda \in \mathbf{F}$ and $T \in \mathcal{L}(V, W)$. Then $\mathcal{M}(\lambda T) = \lambda \mathcal{M}(T)$.

The verification of the result above is also left to the reader.

Because addition and scalar multiplication have now been defined for matrices, you should not be surprised that a vector space is about to appear. First we introduce a bit of notation so that this new vector space has a name, and then we find the dimension of this new vector space.

3.39 notation: $\mathbf{F}^{m,n}$

For m and n positive integers, the set of all m-by-n matrices with entries in $\mathbf{F}$ is denoted by $\mathbf{F}^{m,n}$.

3.40 $\dim \mathbf{F}^{m,n} = mn$

Suppose m and n are positive integers. With addition and scalar multiplication defined as above, $\mathbf{F}^{m,n}$ is a vector space of dimension mn.

Proof The verification that $\mathbf{F}^{m,n}$ is a vector space is left to the reader. Note that the additive identity of $\mathbf{F}^{m,n}$ is the m-by-n matrix all of whose entries equal 0.

The reader should also verify that the list of distinct m-by-n matrices that have 0 in all entries except for a 1 in one entry is a basis of $\mathbf{F}^{m,n}$. There are mn such matrices, so the dimension of $\mathbf{F}^{m,n}$ equals mn. ∎

Matrix Multiplication

Suppose, as previously, that $v_1, \ldots, v_n$ is a basis of V and $w_1, \ldots, w_m$ is a basis of W. Suppose also that $u_1, \ldots, u_p$ is a basis of U.

Consider linear maps $T : U \to V$ and $S : V \to W$. The composition ST is a linear map from U to W. Does $\mathcal{M}(ST)$ equal $\mathcal{M}(S)\mathcal{M}(T)$? This question does not yet make sense because we have not defined the product of two matrices. We will choose a definition of matrix multiplication that forces this question to have a positive answer. Let's see how to do this.

3.37 דוגמה: חיבור וכפל בסקלר של מטריצות

$$2\begin{pmatrix} 3 & 1 \\ -1 & 5 \end{pmatrix} + \begin{pmatrix} 4 & 2 \\ 1 & 6 \end{pmatrix} = \begin{pmatrix} 6 & 2 \\ -2 & 10 \end{pmatrix} + \begin{pmatrix} 4 & 2 \\ 1 & 6 \end{pmatrix} = \begin{pmatrix} 10 & 4 \\ -1 & 16 \end{pmatrix}$$

במסקנה הבאה, ההנחה היא כי נעשה שימוש באותם בסיסים עבור שתי ההעתקות הליניאריות λT ו־ T.

3.38 המטריצה של מכפלת סקלר בהעתקה ליניארית

נניח ש־ $\lambda \in \mathbf{F}$ ו־ $T \in \mathcal{L}(V, W)$. אזי $\mathcal{M}(\lambda T) = \lambda \mathcal{M}(T)$.

גם וידוא המסקנה לעיל מושאר בידי הקוראים.

כעת, כשחיבור וכפל בסקלר מוגדרים עבור מטריצות, אל תופתעו שמופיע מרחב וקטורי. תחילה נכיר מספר סימונים כך שלמרחב הווקטורי הזה יהיה שם, ואז נמצא את ממדו.

3.39 סימון: $\mathbf{F}^{m,n}$

עבור m ו־ n מספרים שלמים חיוביים, קבוצת כל המטריצות $m \times n$ עם איברים ב־ $\mathbf{F}$ מסומנת ב־ $\mathbf{F}^{m,n}$.

3.40 $\dim \mathbf{F}^{m,n} = mn$

נניח ש־ m ו־ n מספרים שלמים חיוביים. עם פעולות החיבור והכפל בסקלר כפי שהוגדרו לעיל, $\mathbf{F}^{m,n}$ מרחב וקטורי מממד mn.

הוכחה ההוכחה לכך ש־ $\mathbf{F}^{m,n}$ הוא מרחב וקטורי מושארת בידי הקוראים. נשים לב כי איבר הזהות לחיבור של $\mathbf{F}^{m,n}$ הוא מטריצה $m \times n$ שכל איבריה 0.

עליכם לוודא בנוסף שהרשימה של המטריצות $m \times n$ השונות זו מזו, שכל איבריהן 0 למעט 1 באיבר אחד, היא בסיס של $\mathbf{F}^{m,n}$. יש mn מטריצות שכאלה ולכן הממד של $\mathbf{F}^{m,n}$ הוא mn. ■

כפל מטריצות

נניח, כמו קודם לכן, ש־ $v_1, \ldots, v_n$ בסיס של V ו־ $w_1, \ldots, w_m$ בסיס של W. נניח בנוסף ש־ $u_1, \ldots, u_p$ בסיס של U.

נתבונן בהעתקות ליניאריות $T: U \to V$ ו־ $S: V \to W$. ההרכבה ST היא העתקה ליניארית מ־ U ל־ W. האם $\mathcal{M}(ST)$ שווה ל־ $\mathcal{M}(S)\mathcal{M}(T)$? לשאלה זו אין עדיין משמעות, משום שעדיין לא הגדרנו מכפלת שתי מטריצות. אנו נבחר הגדרת כפל מטריצות כזו שהתשובה לשאלה זו תהיה חיובית. הבה ניגש למלאכה.

Suppose $\mathcal{M}(S) = A$ and $\mathcal{M}(T) = B$. For $1 \le k \le p$, we have

$$\begin{aligned}(ST)u_k &= S\Big(\sum_{r=1}^{n} B_{r,k} v_r\Big)\\ &= \sum_{r=1}^{n} B_{r,k} S v_r\\ &= \sum_{r=1}^{n} B_{r,k} \sum_{j=1}^{m} A_{j,r} w_j\\ &= \sum_{j=1}^{m} \Big(\sum_{r=1}^{n} A_{j,r} B_{r,k}\Big) w_j.\end{aligned}$$

Thus $\mathcal{M}(ST)$ is the m-by-p matrix whose entry in row j, column k, equals

$$\sum_{r=1}^{n} A_{j,r} B_{r,k}.$$

Now we see how to define matrix multiplication so that the desired equation $\mathcal{M}(ST) = \mathcal{M}(S)\mathcal{M}(T)$ holds.

3.41 definition: *matrix multiplication*

Suppose A is an m-by-n matrix and B is an n-by-p matrix. Then AB is defined to be the m-by-p matrix whose entry in row j, column k, is given by the equation

$$(AB)_{j,k} = \sum_{r=1}^{n} A_{j,r} B_{r,k}.$$

Thus the entry in row j, column k, of AB is computed by taking row j of A and column k of B, multiplying together corresponding entries, and then summing.

Note that we define the product of two matrices only when the number of columns of the first matrix equals the number of rows of the second matrix.

You may have learned this definition of matrix multiplication in an earlier course, although you may not have seen this motivation for it.

3.42 example: *matrix multiplication*

Here we multiply together a 3-by-2 matrix and a 2-by-4 matrix, obtaining a 3-by-4 matrix:

$$\begin{pmatrix} 1 & 2\\ 3 & 4\\ 5 & 6 \end{pmatrix} \begin{pmatrix} 6 & 5 & 4 & 3\\ 2 & 1 & 0 & -1 \end{pmatrix} = \begin{pmatrix} 10 & 7 & 4 & 1\\ 26 & 19 & 12 & 5\\ 42 & 31 & 20 & 9 \end{pmatrix}.$$

Matrix multiplication is not commutative—AB is not necessarily equal to BA even if both products are defined (see Exercise 10). Matrix multiplication is distributive and associative (see Exercises 11 and 12).

נניח ש־ $\mathcal{M}(S) = A$ ו־ $\mathcal{M}(T) = B$. עבור $1 \le k \le p$, נקבל

$$\begin{aligned}(ST)u_k &= S\Big(\sum_{r=1}^{n} B_{r,k} v_r\Big)\\ &= \sum_{r=1}^{n} B_{r,k} S v_r\\ &= \sum_{r=1}^{n} B_{r,k} \sum_{j=1}^{m} A_{j,r} w_j\\ &= \sum_{j=1}^{m} \Big(\sum_{r=1}^{n} A_{j,r} B_{r,k}\Big) w_j.\end{aligned}$$

לפיכך $\mathcal{M}(ST)$ היא מטריצה $m \times p$ שהאיבר שלה בשורה j, עמודה k, שווה ל־

$$\sum_{r=1}^{n} A_{j,r} B_{r,k}.$$

כעת נראה איך להגדיר כפל מטריצות כך שהמשוואה $\mathcal{M}(ST) = \mathcal{M}(S)\mathcal{M}(T)$ תתקיים.

3.41 הגדרה: כפל מטריצות

נניח ש־ A מטריצה $m \times n$ ו־ B מטריצה $n \times p$. אזי AB מוגדרת כמטריצה $m \times p$ שאיבריה בשורה j, עמודה k, נתונים על ידי המשוואה

$$(AB)_{j,k} = \sum_{r=1}^{n} A_{j,r} B_{r,k}.$$

לפיכך, האיבר בשורה j, עמודה k, של AB מחושב באופן שבו לוקחים את שורה j של A ואת עמודה k של B, כופלים איברים תואמים, ואז סוכמים.

שימו לב כי אנו מגדירים מכפלת שתי מטריצות רק כאשר מספר העמודות במטריצה הראשונה שווה למספר השורות במטריצה השנייה.

ייתכן מאוד שכבר למדתם על כפל מטריצות בקורס קודם, אך ייתכן גם שלא נחשפתם למוטיבציה שביסוד הגדרת הכפל באופן זה.

3.42 דוגמה: כפל מטריצות

כאן אנו כופלים מטריצה 3×2 במטריצה 2×4, ומקבלים מטריצה 3×4:

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \begin{pmatrix} 6 & 5 & 4 & 3 \\ 2 & 1 & 0 & -1 \end{pmatrix} = \begin{pmatrix} 10 & 7 & 4 & 1 \\ 26 & 19 & 12 & 5 \\ 42 & 31 & 20 & 9 \end{pmatrix}.$$

כפל מטריצות אינו פעולה קומוטטיבית—AB אינה בהכרח שווה ל־ BA גם אם שתי המכפלות מוגדרות (ראו תרגיל 10). כפל מטריצות הוא דיסטריבוטיבי ואסוציאטיבי (ראו תרגילים 11 ו־ 12).

In the next result, we assume that the same basis of V is used in considering $T \in \mathcal{L}(U,V)$ and $S \in \mathcal{L}(V,W)$, the same basis of W is used in considering $S \in \mathcal{L}(V,W)$ and $ST \in \mathcal{L}(U,W)$, and the same basis of U is used in considering $T \in \mathcal{L}(U,V)$ and $ST \in \mathcal{L}(U,W)$.

3.43 *matrix of product of linear maps*

If $T \in \mathcal{L}(U,V)$ and $S \in \mathcal{L}(V,W)$, then $\mathcal{M}(ST) = \mathcal{M}(S)\mathcal{M}(T)$.

The proof of the result above is the calculation that was done as motivation before the definition of matrix multiplication.

In the next piece of notation, note that as usual the first index refers to a row and the second index refers to a column, with a vertically centered dot used as a placeholder.

3.44 notation: $A_{j,\cdot}$, $A_{\cdot,k}$

Suppose A is an m-by-n matrix.

- If $1 \le j \le m$, then $A_{j,\cdot}$ denotes the 1-by-n matrix consisting of row j of A.
- If $1 \le k \le n$, then $A_{\cdot,k}$ denotes the m-by-1 matrix consisting of column k of A.

3.45 example: $A_{j,\cdot}$ *equals* j^{th} *row of* A *and* $A_{\cdot,k}$ *equals* k^{th} *column of* A

The notation $A_{2,\cdot}$ denotes the second row of A and $A_{\cdot,2}$ denotes the second column of A. Thus if $A = \begin{pmatrix} 8 & 4 & 5 \\ 1 & 9 & 7 \end{pmatrix}$, then

$$A_{2,\cdot} = \begin{pmatrix} 1 & 9 & 7 \end{pmatrix} \quad \text{and} \quad A_{\cdot,2} = \begin{pmatrix} 4 \\ 9 \end{pmatrix}.$$

The product of a 1-by-n matrix and an n-by-1 matrix is a 1-by-1 matrix. However, we will frequently identify a 1-by-1 matrix with its entry. For example,

$$\begin{pmatrix} 3 & 4 \end{pmatrix} \begin{pmatrix} 6 \\ 2 \end{pmatrix} = \begin{pmatrix} 26 \end{pmatrix}$$

because $3 \cdot 6 + 4 \cdot 2 = 26$. However, we can identify $\begin{pmatrix} 26 \end{pmatrix}$ with 26, writing $\begin{pmatrix} 3 & 4 \end{pmatrix} \begin{pmatrix} 6 \\ 2 \end{pmatrix} = 26$.

The next result uses the convention discussed in the paragraph above to give another way to think of matrix multiplication. For example, the next result and the calculation in the paragraph above explain why the entry in row 2, column 1, of the product in Example 3.42 equals 26.

במסקנה הבאה, אנו מניחים כי נעשה שימוש באותו בסיס של V כשאנו מתבוננים ב־ $T \in \mathcal{L}(U,V)$ ו־ $S \in \mathcal{L}(V,W)$, ושימוש באותו בסיס של W כשאנו מתבוננים ב־ $S \in \mathcal{L}(V,W)$ ו־ $ST \in \mathcal{L}(U,W)$, ושימוש באותו בסיס של U כשאנו מתבוננים ב־ $T \in \mathcal{L}(U,V)$ ו־ $ST \in \mathcal{L}(U,W)$.

3.43 מטריצת מכפלת העתקות ליניאריות

אם $T \in \mathcal{L}(U,V)$ ו־ $S \in \mathcal{L}(V,W)$, אזי $\mathcal{M}(ST) = \mathcal{M}(S)\mathcal{M}(T)$.

הוכחת המסקנה לעיל היא החישוב שביצענו כמוטיבציה טרם ההגדרה של כפל מטריצות.

בקטע הסימון הבא, שימו לב כי, כהרגלנו, האינדקס הראשון מתייחס לשורה והשני לעמודה, עם נקודה ממורכזת אנכית המשמשת כממלאת־מקום.

3.44 סימון: $A_{j,.}$, $A_{.,k}$

נניח ש־ A מטריצה $m \times n$.

- אם $1 \le j \le m$, אזי $A_{j,.}$ הוא הסימון למטריצה $1 \times n$ המורכבת משורה j של A.
- אם $1 \le k \le n$, אזי $A_{.,k}$ הוא הסימון למטריצה $m \times 1$ המורכבת מעמודה k של A.

3.45 דוגמה: $A_{j,.}$ שווה לשורה ה־ j^{th} של A ו־ $A_{.,k}$ שווה לעמודה ה־ k^{th} של A

הסימון $A_{2,.}$ מייצג את השורה השנייה של A ו־ $A_{.,2}$ מייצג את העמודה השנייה של A. לפיכך, אם $A = \begin{pmatrix} 8 & 4 & 5 \\ 1 & 9 & 7 \end{pmatrix}$, אזי

$$A_{2,.} = \begin{pmatrix} 1 & 9 & 7 \end{pmatrix} \quad \text{וגם} \quad A_{.,2} = \begin{pmatrix} 4 \\ 9 \end{pmatrix}.$$

מכפלת מטריצה $1 \times n$ במטריצה $n \times 1$ היא מטריצה 1×1. אולם, לעיתים קרובות נתייחס אל מטריצה 1×1 כאל האיבר שלה. לדוגמה,

$$\begin{pmatrix} 3 & 4 \end{pmatrix} \begin{pmatrix} 6 \\ 2 \end{pmatrix} = \begin{pmatrix} 26 \end{pmatrix}$$

מאחר ש־ $3 \cdot 6 + 4 \cdot 2 = 26$. עם זאת, אנו יכולים להתייחס ל־ $\begin{pmatrix} 26 \end{pmatrix}$ כ־ 26, כך שנכתוב $\begin{pmatrix} 3 & 4 \end{pmatrix} \begin{pmatrix} 6 \\ 2 \end{pmatrix} = 26$.

המסקנה הבאה משתמשת במוסכמה שתוארה בפסקה לעיל כדי לחשוב על כפל מטריצות באופן אחר. לדוגמה, המסקנה הבאה בשילוב עם החישוב בפסקה לעיל מסבירה מדוע האיבר בשורה 2, עמודה 1, של המכפלה בדוגמה 3.42 שווה 26.

3.46 *entry of matrix product equals row times column*

Suppose A is an m-by-n matrix and B is an n-by-p matrix. Then

$$(AB)_{j,k} = A_{j,\cdot} B_{\cdot,k}$$

if $1 \le j \le m$ and $1 \le k \le p$. In other words, the entry in row j, column k, of AB equals (row j of A) times (column k of B).

Proof Suppose $1 \le j \le m$ and $1 \le k \le p$. The definition of matrix multiplication states that

3.47 $$(AB)_{j,k} = A_{j,1}B_{1,k} + \cdots + A_{j,n}B_{n,k}.$$

The definition of matrix multiplication also implies that the product of the 1-by-n matrix $A_{j,\cdot}$ and the n-by-1 matrix $B_{\cdot,k}$ is the 1-by-1 matrix whose entry is the number on the right side of the equation above. Thus the entry in row j, column k, of AB equals (row j of A) times (column k of B). ∎

The next result gives yet another way to think of matrix multiplication. In the result below, $(AB)_{\cdot,k}$ is column k of the m-by-p matrix AB. Thus $(AB)_{\cdot,k}$ is an m-by-1 matrix. Also, $AB_{\cdot,k}$ is an m-by-1 matrix because it is the product of an m-by-n matrix and an n-by-1 matrix. Thus the two sides of the equation in the result below have the same size, making it reasonable that they might be equal.

3.48 *column of matrix product equals matrix times column*

Suppose A is an m-by-n matrix and B is an n-by-p matrix. Then

$$(AB)_{\cdot,k} = AB_{\cdot,k}$$

if $1 \le k \le p$. In other words, column k of AB equals A times column k of B.

Proof As discussed above, $(AB)_{\cdot,k}$ and $AB_{\cdot,k}$ are both m-by-1 matrices. If $1 \le j \le m$, then the entry in row j of $(AB)_{\cdot,k}$ is the left side of 3.47 and the entry in row j of $AB_{\cdot,k}$ is the right side of 3.47. Thus $(AB)_{\cdot,k} = AB_{\cdot,k}$. ∎

Our next result will give another way of thinking about the product of an m-by-n matrix and an n-by-1 matrix, motivated by the next example.

3.49 example: *product of a* 3*-by-*2 *matrix and a* 2*-by-*1 *matrix*

Use our definitions and basic arithmetic to verify that

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \begin{pmatrix} 5 \\ 1 \end{pmatrix} = \begin{pmatrix} 7 \\ 19 \\ 31 \end{pmatrix} = 5 \begin{pmatrix} 1 \\ 3 \\ 5 \end{pmatrix} + 1 \begin{pmatrix} 2 \\ 4 \\ 6 \end{pmatrix}.$$

Thus in this example, the product of a 3-by-2 matrix and a 2-by-1 matrix is a linear combination of the columns of the 3-by-2 matrix, with the scalars (5 and 1) that multiply the columns coming from the 2-by-1 matrix.

3.46 איבר במכפלת מטריצות שווה לכפל שורה בעמודה

נניח ש־ A היא מטריצה $m \times n$ ו־ B מטריצה $n \times p$. אזי
$$(AB)_{j,k} = A_{j,\cdot} B_{\cdot,k}$$
עבור $1 \le j \le m$ ו־ $1 \le k \le p$. במילים אחרות, האיבר בשורה j, עמודה k, של AB שווה ל־ (שורה j של A) כפול (עמודה k של B).

הוכחה נניח ש־ $1 \le j \le m$ ו־ $1 \le k \le p$. הגדרת כפל מטריצות קובעת ש־
$$(AB)_{j,k} = A_{j,1}B_{1,k} + \cdots + A_{j,n}B_{n,k}. \tag{3.47}$$
הגדרת כפל מטריצות קובעת בנוסף שמכפלת המטריצה $A_{j,\cdot}$ $1 \times n$ והמטריצה $B_{\cdot,k}$ $n \times 1$ היא מטריצה 1×1 שאיברה הוא המספר באגף הימני של המשוואה לעיל. לפיכך, האיבר בשורה j, עמודה k, של AB שווה ל־ (שורה j של A) כפול (עמודה k של B). ■

המסקנה הבאה מעניקה לנו נקודת מבט נוספת על כפל מטריצות. במסקנה להלן, $(AB)_{\cdot,k}$ היא עמודה k של AB מטריצה $m \times p$. לפיכך $(AB)_{\cdot,k}$ היא מטריצה $m \times 1$. בנוסף, $AB_{\cdot,k}$ היא מטריצה $m \times 1$. מאחר שהיא מכפלה של מטריצה $m \times n$ במטריצה $n \times 1$. לפיכך גודל המטריצות בשני אגפי המשוואה שווה, וסביר לשער שהן אף שוות.

3.48 עמודת מכפלת מטריצות שווה למטריצה כפול עמודה

נניח ש־ A מטריצה $m \times n$ ו־ B מטריצה $n \times p$. אזי
$$(AB)_{\cdot,k} = AB_{\cdot,k}$$
עבור $1 \le k \le p$. כלומר, עמודה k של AB שווה ל־ A כפול עמודה k של B.

הוכחה כפי שאמרנו לעיל, $(AB)_{\cdot,k}$ ו־ $AB_{\cdot,k}$ שתיהן מטריצות $m \times 1$. אם $1 \le j \le m$, אזי האיבר בשורה j של $(AB)_{\cdot,k}$ הוא האגף השמאלי של 3.47 והאיבר בשורה j של $AB_{\cdot,k}$ הוא האגף הימני של 3.47. לפיכך $(AB)_{\cdot,k} = AB_{\cdot,k}$. ■

המסקנה הבאה מעניקה נקודת מבט נוספת על מכפלת מטריצה $m \times n$ במטריצה $n \times 1$, כשאת המוטיבציה לכך נראה בדוגמה הבאה.

3.49 דוגמה: מכפלת מטריצה 2×3 במטריצה 3×1

השתמשו בהגדרות שלנו ובכללי החשבון הבסיסיים כדי לוודא שאכן
$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \begin{pmatrix} 5 \\ 1 \end{pmatrix} = \begin{pmatrix} 7 \\ 19 \\ 31 \end{pmatrix} = 5 \begin{pmatrix} 1 \\ 3 \\ 5 \end{pmatrix} + 1 \begin{pmatrix} 2 \\ 4 \\ 6 \end{pmatrix}.$$
לפיכך בדוגמה זו, מכפלת מטריצה 2×3 במטריצה 3×1 היא צירוף ליניארי של עמודות המטריצה 2×3, כאשר הסקלרים (5 ו־ 1) המכפילים את העמודות מגיעים מהמטריצה 2×1.

The next result generalizes the example above.

3.50 *linear combination of columns*

Suppose A is an m-by-n matrix and $b = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$ is an n-by-1 matrix. Then

$$Ab = b_1 A_{\cdot,1} + \cdots + b_n A_{\cdot,n}.$$

In other words, Ab is a linear combination of the columns of A, with the scalars that multiply the columns coming from b.

Proof If $k \in \{1, \ldots, m\}$, then the definition of matrix multiplication implies that the entry in row k of the m-by-1 matrix Ab is

$$A_{k,1}b_1 + \cdots + A_{k,n}b_n.$$

The entry in row k of $b_1 A_{\cdot,1} + \cdots + b_n A_{\cdot,n}$ also equals the number displayed above.

Because Ab and $b_1 A_{\cdot,1} + \cdots + b_n A_{\cdot,n}$ have the same entry in row k for each $k \in \{1, \ldots, m\}$, we conclude that $Ab = b_1 A_{\cdot,1} + \cdots + b_n A_{\cdot,n}$. ∎

Our two previous results focus on the columns of a matrix. Analogous results hold for the rows of a matrix. Specifically, see Exercises 8 and 9, which can be proved using appropriate modifications of the proofs of 3.48 and 3.50.

The next result is the main tool used in the next subsection to prove the column–row factorization (3.56) and to prove that the column rank of a matrix equals the row rank (3.57). To be consistent with the notation often used with the column–row factorization, including in the next subsection, the matrices in the next result are called C and R instead of A and B.

3.51 *matrix multiplication as linear combinations of columns or rows*

Suppose C is an m-by-c matrix and R is a c-by-n matrix.

(a) If $k \in \{1, \ldots, n\}$, then column k of CR is a linear combination of the columns of C, with the coefficients of this linear combination coming from column k of R.

(b) If $j \in \{1, \ldots, m\}$, then row j of CR is a linear combination of the rows of R, with the coefficients of this linear combination coming from row j of C.

Proof Suppose $k \in \{1, \ldots, n\}$. Then column k of CR equals $CR_{\cdot,k}$ (by 3.48), which equals the linear combination of the columns of C with coefficients coming from $R_{\cdot,k}$ (by 3.50). Thus (a) holds.

To prove (b), follow the pattern of the proof of (a) but use rows instead of columns and use Exercises 8 and 9 instead of 3.48 and 3.50. ∎

המסקנה הבאה מכלילה את הדוגמה לעיל.

3.50 צירוף ליניארי של עמודות

נניח ש־ A מטריצה $m \times n$ ו־ $b = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$ מטריצה $n \times 1$. אזי

$$Ab = b_1 A_{\cdot,1} + \cdots + b_n A_{\cdot,n}.$$

במילים אחרות, Ab היא צירוף ליניארי של עמודות A, כאשר הסקלרים המכפילים את העמודות מגיעים מ־ b.

הוכחה אם $k \in \{1, \ldots, m\}$, אזי הגדרת כפל מטריצות גוררת שהאיבר בשורה k של המטריצה Ab $m \times 1$ הוא

$$A_{k,1} b_1 + \cdots + A_{k,n} b_n.$$

האיבר בשורה k של $b_1 A_{\cdot,1} + \cdots + b_n A_{\cdot,n}$ גם הוא שווה למספר המוצג לעיל.

מאחר של־ Ab ולכל $k \in \{1, \ldots, m\}$ יש אותו איבר בשורה k של $b_1 A_{\cdot,1} + \cdots + b_n A_{\cdot,n}$, אנו מסיקים ש־ $Ab = b_1 A_{\cdot,1} + \cdots + b_n A_{\cdot,n}$. ■

שתי המסקנות האחרונות שלנו התמקדו בעמודות המטריצה. מסקנות דומות מתקיימות גם עבור שורות המטריצה. בפרט, ראו תרגילים 8 ו־ 9, שניתנים להוכחה באמצעות שינויים מתאימים שנבצע בהוכחות של 3.48 ו־ 3.50.

המסקנה הבאה היא הכלי העיקרי שישמש אותנו בתת־היחידה הבאה להוכחת פירוק עמודה־שורה (3.56) ולהוכחה כי דרגת העמודות של המטריצה שווה לדרגת השורות (3.57). כדי לשמור על עקביות הסימונים שבהם נשתמש לרוב בפירוק עמודה־שורה, כולל בתת־היחידה הבאה, המטריצות במסקנה הבאה תיקראנה C ו־ R במקום A ו־ B.

3.51 כפל מטריצות כצירוף ליניארי של עמודות או שורות

נניח ש־ C מטריצה $m \times c$ ו־ R מטריצה $c \times n$.

א. אם $k \in \{1, \ldots, n\}$, אזי עמודה k של CR היא צירוף ליניארי של עמודות C, כאשר מקדמי הצירוף מגיעים מעמודה k של R.

ב. אם $j \in \{1, \ldots, m\}$, אזי שורה j של CR היא צירוף ליניארי של שורות R, כאשר מקדמי הצירוף מגיעים משורה j של C.

הוכחה נניח ש־ $k \in \{1, \ldots, n\}$. אזי עמודה k של CR שווה ל־ $CR_{\cdot,k}$ (לפי 3.48), השווה לצירוף הליניארי של עמודות C עם מקדמים המגיעים מ־ $R_{\cdot,k}$ (לפי 3.50). לפיכך (א) מתקיים.

להוכחת (ב), השתמשו בתבנית ההוכחה של (א), החליפו עמודות בשורות והשתמשו בתרגילים 8 ו־ 9 במקום 3.48 ו־ 3.50. ■

Column–Row Factorization and Rank of a Matrix

We begin by defining two nonnegative integers associated with each matrix.

3.52 definition: *column rank, row rank*

Suppose A is an m-by-n matrix with entries in $\mathbf{F}$.

- The *column rank* of A is the dimension of the span of the columns of A in $\mathbf{F}^{m,1}$.
- The *row rank* of A is the dimension of the span of the rows of A in $\mathbf{F}^{1,n}$.

If A is an m-by-n matrix, then the column rank of A is at most n (because A has n columns) and the column rank of A is also at most m (because $\dim \mathbf{F}^{m,1} = m$). Similarly, the row rank of A is also at most $\min\{m, n\}$.

3.53 example: *column rank and row rank of a 2-by-4 matrix*

Suppose

$$A = \begin{pmatrix} 4 & 7 & 1 & 8 \\ 3 & 5 & 2 & 9 \end{pmatrix}.$$

The column rank of A is the dimension of

$$\operatorname{span}\left(\begin{pmatrix} 4 \\ 3 \end{pmatrix}, \begin{pmatrix} 7 \\ 5 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 8 \\ 9 \end{pmatrix} \right)$$

in $\mathbf{F}^{2,1}$. Neither of the first two vectors listed above in $\mathbf{F}^{2,1}$ is a scalar multiple of the other. Thus the span of this list of length four has dimension at least two. The span of this list of vectors in $\mathbf{F}^{2,1}$ cannot have dimension larger than two because $\dim \mathbf{F}^{2,1} = 2$. Thus the span of this list has dimension two, which means that the column rank of A is two.

The row rank of A is the dimension of

$$\operatorname{span}\big(\begin{pmatrix} 4 & 7 & 1 & 8 \end{pmatrix}, \begin{pmatrix} 3 & 5 & 2 & 9 \end{pmatrix} \big)$$

in $\mathbf{F}^{1,4}$. Neither of the two vectors listed above in $\mathbf{F}^{1,4}$ is a scalar multiple of the other. Thus the span of this list of length two has dimension two, which means that the row rank of A is two.

We now define the transpose of a matrix.

3.54 definition: *transpose*, A^t

The *transpose* of a matrix A, denoted by A^t, is the matrix obtained from A by interchanging rows and columns. Specifically, if A is an m-by-n matrix, then A^t is the n-by-m matrix whose entries are given by the equation

$$(A^t)_{k,j} = A_{j,k}.$$

פירוק עמודה־שורה ודרגת מטריצה

נפתח בהגדרת שני מספרים שלמים אי־שליליים המשויכים לכל מטריצה.

3.52 הגדרה: דרגת עמודות, דרגת שורות

נניח ש־ A מטריצה $m \times n$ עם איברים ב־ $\mathbf{F}$.

- דרגת העמודות של A היא ממד הפריסה של עמודות A ב־ $\mathbf{F}^{m,1}$.
- דרגת השורות של A היא ממד הפריסה של שורות A ב־ $\mathbf{F}^{1,n}$.

אם A מטריצה $m \times n$, אזי דרגת העמודות של A היא לכל היותר n (שכן יש n עמודות ב־ A) ודרגת העמודות של A היא גם לכל היותר m (שכן $\dim \mathbf{F}^{m,1} = m$). באותו האופן, דרגת השורות של A היא לכל היותר $\min\{m, n\}$.

3.53 דוגמה: דרגת עמודות ודרגת שורות של מטריצה 2×4

נניח ש־

$$A = \begin{pmatrix} 4 & 7 & 1 & 8 \\ 3 & 5 & 2 & 9 \end{pmatrix}.$$

דרגת העמודות של A היא הממד של

$$\operatorname{span}\left(\begin{pmatrix} 4 \\ 3 \end{pmatrix}, \begin{pmatrix} 7 \\ 5 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 8 \\ 9 \end{pmatrix} \right)$$

ב־ $\mathbf{F}^{2,1}$. אף אחד משני הווקטורים הראשונים ברשימה לעיל ב־ $\mathbf{F}^{2,1}$ אינו כפולת סקלר של האחר. לפיכך ממד הפריסה של רשימה זו באורך ארבע הוא לפחות שניים. ממד הפריסה של רשימת וקטורים זו ב־ $\mathbf{F}^{2,1}$ אינו יכול להיות גדול משניים מאחר ש־ $\dim \mathbf{F}^{2,1} = 2$. לפיכך, ממד הפריסה של רשימה זו הוא שניים, ומשמעות הדבר היא כי דרגת העמודות של A היא שתיים.

דרגת השורות של A היא הממד של

$$\operatorname{span}\big(\begin{pmatrix} 4 & 7 & 1 & 8 \end{pmatrix}, \begin{pmatrix} 3 & 5 & 2 & 9 \end{pmatrix} \big)$$

ב־ $\mathbf{F}^{1,4}$. אף אחד משני הווקטורים ברשימה לעיל ב־ $\mathbf{F}^{1,4}$ אינו כפולת סקלר של האחר. לפיכך, ממד הפריסה של רשימה זו בת שני וקטורים הוא שניים, ולכן דרגת השורות של A היא שתיים.

כעת נגדיר מטריצה משוחלפת.

3.54 הגדרה: מטריצה משוחלפת, A^t

המטריצה המשוחלפת של A, המסומנת ב־ A^t, היא המטריצה המתקבלת מ־ A על ידי הפיכת כל שורה לעמודה (שחלוף). ליתר דיוק, אם A מטריצה $m \times n$, אזי A^t היא מטריצה $n \times m$ שאיבריה נתונים על ידי המשוואה

$$(A^t)_{k,j} = A_{j,k}.$$

3.55 example: *transpose of a matrix*

If $A = \begin{pmatrix} 5 & -7 \\ 3 & 8 \\ -4 & 2 \end{pmatrix}$, then $A^t = \begin{pmatrix} 5 & 3 & -4 \\ -7 & 8 & 2 \end{pmatrix}$.

Note that here A is a 3-by-2 matrix and A^t is a 2-by-3 matrix.

The transpose has nice algebraic properties: $(A+B)^t = A^t + B^t$, $(\lambda A)^t = \lambda A^t$, and $(AC)^t = C^t A^t$ for all m-by-n matrices A, B, all $\lambda \in \mathbf{F}$, and all n-by-p matrices C (see Exercises 14 and 15).

The next result will be the main tool used to prove that the column rank equals the row rank (see 3.57).

3.56 *column–row factorization*

Suppose A is an m-by-n matrix with entries in $\mathbf{F}$ and column rank $c \geq 1$. Then there exist an m-by-c matrix C and a c-by-n matrix R, both with entries in $\mathbf{F}$, such that $A = CR$.

Proof Each column of A is an m-by-1 matrix. The list $A_{\cdot,1}, \ldots, A_{\cdot,n}$ of columns of A can be reduced to a basis of the span of the columns of A (by 2.30). This basis has length c, by the definition of the column rank. The c columns in this basis can be put together to form an m-by-c matrix C.

If $k \in \{1, \ldots, n\}$, then column k of A is a linear combination of the columns of C. Make the coefficients of this linear combination into column k of a c-by-n matrix that we call R. Then $A = CR$, as follows from 3.51(a). ∎

In Example 3.53, the column rank and row rank turned out to equal each other. The next result states that this happens for all matrices.

3.57 *column rank equals row rank*

Suppose $A \in \mathbf{F}^{m,n}$. Then the column rank of A equals the row rank of A.

Proof Let c denote the column rank of A. If $c = 0$, then $A = 0$ and hence the row rank of A also equals 0. Thus we can assume that $c \geq 1$.

Let $A = CR$ be the column–row factorization of A given by 3.56, where C is an m-by-c matrix and R is a c-by-n matrix. Then 3.51(b) tells us that every row of A is a linear combination of the rows of R. Because R has c rows, this implies that the row rank of A is less than or equal to the column rank c of A.

To prove the inequality in the other direction, apply the previous paragraph result to A^t, getting

$$\begin{aligned} \text{column rank of } A &= \text{row rank of } A^t \\ &\leq \text{column rank of } A^t \\ &= \text{row rank of } A. \end{aligned}$$

Thus the column rank of A equals the row rank of A. ∎

3.55 דוגמה: מטריצה משוחלפת

אם $A = \begin{pmatrix} 5 & -7 \\ 3 & 8 \\ -4 & 2 \end{pmatrix}$, אזי $A^t = \begin{pmatrix} 5 & 3 & -4 \\ -7 & 8 & 2 \end{pmatrix}$.

שימו לב שכאן A מטריצה 3×2 ו־ A^t מטריצה 2×3.

למטריצה המשוחלפת תכונות אלגבריות נחמדות: $(A+B)^t = A^t+B^t$, $(\lambda A)^t = \lambda A^t$, ו־ $(AC)^t = C^t A^t$ לכל A, B מטריצות $m \times n$, לכל $\lambda \in \mathbf{F}$, ולכל C מטריצה $n \times p$ (ראו תרגילים 14 ו־ 15). המסקנה הבאה תשמש ככלי העיקרי להוכחת הטענה שדרגת העמודות שווה לדרגת השורות (ראו 3.57).

3.56 פירוק עמודה־שורה

נניח ש־ A מטריצה $m \times n$ עם איברים ב־ $\mathbf{F}$ ודרגת העמודות $c \geq 1$. אזי, קיימות C מטריצה $m \times c$ ו־ R מטריצה $c \times n$, שתיהן עם איברים ב־ $\mathbf{F}$, כך ש־ $A = CR$.

הוכחה כל עמודה של A היא מטריצה $m \times 1$. רשימת העמודות $A_{\cdot,1}, \ldots, A_{\cdot,n}$ של A יכולה להצטמצם לבסיס של פריסת העמודות של A (לפי 2.30). אורכו של בסיס זה הוא c, לפי הגדרת דרגת העמודות. נוכל לצרף את c העמודות בבסיס זה לכדי מטריצה $m \times c$ שנסמן ב־ C.

אם $k \in \{1, \ldots, n\}$, אזי עמודה k של A היא צירוף ליניארי של עמודות C. את מקדמי הצירוף הליניארי נציב כעמודה k של מטריצה $c \times n$ שאותה נסמן ב־ R ונקבל $A = CR$ כפי שנובע מ־ 3.51(א). ■

בדוגמה 3.53, קיבלנו שדרגת העמודות ודרגת השורות שוות זו לזו. המסקנה הבאה קובעת ששוויון זה מתקיים בכל המטריצות.

3.57 דרגת העמודות שווה לדרגת השורות

נניח ש־ $A \in \mathbf{F}^{m,n}$. אזי דרגת העמודות של A שווה לדרגת השורות של A.

הוכחה נסמן ב־ c את דרגת העמודות של A. אם $c = 0$, אזי $A = 0$ ולכן גם דרגת השורות של A שווה ל־ 0. לפיכך אנו יכולים להניח ש־ $c \geq 1$.

נסמן ב־ $A = CR$ את פירוק עמודה־שורה של A שניתן על ידי 3.56, כאשר C היא מטריצה $m \times c$ ו־ R היא מטריצה $c \times n$. אזי מ־ 3.51(ב) נקבל שכל שורה של A היא צירוף ליניארי של שורות R. מאחר שמספר השורות של R הוא c, נסיק שדרגת השורות של A היא קטנה או שווה לדרגת העמודות c של A.

להוכחת אי השוויון בכיוון השני, נפעיל את הפסקה הקודמת על A^t ונקבל

$$\begin{aligned} \text{דרגת העמודות של } A &= \text{דרגת השורות של } A^t \\ &\geq \text{דרגת העמודות של } A^t \\ &= \text{דרגת השורות של } A. \end{aligned}$$

לפיכך דרגת העמודות של A שווה לדרגת השורות של A. ■

Because the column rank equals the row rank, the last result allows us to dispense with the terms "column rank" and "row rank" and just use the simpler term "rank".

3.58 definition: *rank*

The *rank* of a matrix $A \in \mathbf{F}^{m,n}$ is the column rank of A.

See 3.133 and Exercise 8 in Section 7A for alternative proofs that the column rank equals the row rank.

Exercises 3C

1 Suppose $T \in \mathcal{L}(V, W)$. Show that with respect to each choice of bases of V and W, the matrix of T has at least $\dim \operatorname{range} T$ nonzero entries.

2 Suppose $T \in \mathcal{L}(V, W)$, where V and W are finite-dimensional and nonzero. Prove that $\dim \operatorname{range} T = 1$ if and only if there exist a basis of V and a basis of W such that with respect to these bases, all entries of $\mathcal{M}(T)$ equal 1.

3 Suppose $v_1, \dots, v_n$ is a basis of V and $w_1, \dots, w_m$ is a basis of W.

(a) Show that if $S, T \in \mathcal{L}(V, W)$, then $\mathcal{M}(S + T) = \mathcal{M}(S) + \mathcal{M}(T)$.

(b) Show that if $\lambda \in \mathbf{F}$ and $T \in \mathcal{L}(V, W)$, then $\mathcal{M}(\lambda T) = \lambda \mathcal{M}(T)$.

This exercise asks you to verify 3.35 and 3.38.

4 Suppose that $D \in \mathcal{L}(\mathcal{P}_3(\mathbf{R}), \mathcal{P}_2(\mathbf{R}))$ is the differentiation map defined by $Dp = p'$. Find a basis of $\mathcal{P}_3(\mathbf{R})$ and a basis of $\mathcal{P}_2(\mathbf{R})$ such that the matrix of D with respect to these bases is

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

Compare with Example 3.33. The next exercise generalizes this exercise.

5 Suppose V and W are finite-dimensional and $T \in \mathcal{L}(V, W)$. Prove that there exist a basis of V and a basis of W such that with respect to these bases, all entries of $\mathcal{M}(T)$ are 0 except that the entries in row k, column k, equal 1 if $1 \le k \le \dim \operatorname{range} T$.

6 Suppose $v_1, \dots, v_m$ is a basis of V and W is finite-dimensional. Suppose $T \in \mathcal{L}(V, W)$. Prove that there exists a basis $w_1, \dots, w_n$ of W such that all entries in the first column of $\mathcal{M}(T)$ [with respect to the bases $v_1, \dots, v_m$ and $w_1, \dots, w_n$] are 0 except for possibly a 1 in the first row, first column.

In this exercise, unlike Exercise 5, you are given the basis of V instead of being able to choose a basis of V.

מאחר שדרגת העמודות שווה לדרגת השורות, המסקנה האחרונה מאפשרת לנו לחדול מהשימוש במונחים "דרגת העמודות" ו"דרגת השורות" ולהשתמש במונח הפשוט יותר "דרגה".

3.58 הגדרה: דרגה

דרגת המטריצה $A \in \mathbf{F}^{m,n}$ היא דרגת העמודות של A.

ראו 3.133 ותרגיל 8 ביחידה 1ג להוכחות חלופיות לכך שדרגת העמודות שווה לדרגת השורות.

תרגילים ג3

1 נניח ש־ $T \in \mathcal{L}(V,W)$. הראו כי ביחס לכל בחירה של בסיסים עבור V ו־ W, למטריצה של T יש לכל הפחות $\dim \operatorname{range} T$ איברים שאינם אפס.

2 נניח ש־ $T \in \mathcal{L}(V,W)$, כאשר V ו־ W מממד סופי ושונים מאפס. הוכיחו כי $\dim \operatorname{range} T = 1$ אם־ורק־אם קיימים בסיס של V ובסיס של W כך שביחס לבסיסים אלו, כל האיברים של $\mathcal{M}(T)$ שווים 1.

3 נניח ש־ $v_1, \dots, v_n$ בסיס של V ו־ $w_1, \dots, w_m$ בסיס של W.

א. הראו שאם $S, T \in \mathcal{L}(V,W)$, אזי $\mathcal{M}(S+T) = \mathcal{M}(S) + \mathcal{M}(T)$.

ב. הראו שאם $\lambda \in \mathbf{F}$ ו־ $T \in \mathcal{L}(V,W)$, אזי $\mathcal{M}(\lambda T) = \lambda \mathcal{M}(T)$.

תרגיל זה מבקש מכם לווודא את 3.35 ואת 3.38.

4 נניח ש־ $D \in \mathcal{L}(\mathcal{P}_3(\mathbf{R}), \mathcal{P}_2(\mathbf{R}))$ היא העתקת הגזירה המוגדרת על ידי $Dp = p'$. מצאו בסיס של $\mathcal{P}_3(\mathbf{R})$ ובסיס של $\mathcal{P}_2(\mathbf{R})$ כך שהמטריצה של D ביחס לבסיסים אלו היא

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

השוו עם דוגמה 3.33. התרגיל הבא מכליל תרגיל זה.

5 נניח ש־ V ו־ W מממד סופי ו־ $T \in \mathcal{L}(V,W)$. הוכיחו שקיימים בסיס של V ובסיס של W כך שביחס לבסיסים אלו, כל איברי $\mathcal{M}(T)$ הם 0 מלבד אלו שבשורה k, עמודה k, השווים 1 אם $1 \le k \le \dim \operatorname{range} T$.

6 נניח ש־ $v_1, \dots, v_m$ בסיס של V ו־ W מממד סופי. נניח ש־ $T \in \mathcal{L}(V,W)$. הוכיחו שקיים בסיס $w_1, \dots, w_n$ של W כך שכל איברי העמודה הראשונה של $\mathcal{M}(T)$ [ביחס לבסיסים $v_1, \dots, v_m$ ו־ $w_1, \dots, w_n$] הם 0, למעט אולי איבר שערכו 1 בשורה הראשונה, בעמודה הראשונה.

בתרגיל זה, בשונה מתרגיל 5, אתם מקבלים בסיס נתון של V במקום האפשרות לבחור בסיס עבור V.

7 Suppose $w_1, \dots, w_n$ is a basis of W and V is finite-dimensional. Suppose $T \in \mathcal{L}(V, W)$. Prove that there exists a basis $v_1, \dots, v_m$ of V such that all entries in the first row of $\mathcal{M}(T)$ [with respect to the bases $v_1, \dots, v_m$ and $w_1, \dots, w_n$] are 0 except for possibly a 1 in the first row, first column.

In this exercise, unlike Exercise 5, you are given the basis of W instead of being able to choose a basis of W.

8 Suppose A is an m-by-n matrix and B is an n-by-p matrix. Prove that

$$(AB)_{j,\cdot} = A_{j,\cdot} B$$

for each $1 \le j \le m$. In other words, show that row j of AB equals (row j of A) times B.

This exercise gives the row version of 3.48.

9 Suppose $a = \begin{pmatrix} a_1 & \cdots & a_n \end{pmatrix}$ is a 1-by-n matrix and B is an n-by-p matrix. Prove that

$$aB = a_1 B_{1,\cdot} + \cdots + a_n B_{n,\cdot}.$$

In other words, show that aB is a linear combination of the rows of B, with the scalars that multiply the rows coming from a.

This exercise gives the row version of 3.50.

10 Give an example of 2-by-2 matrices A and B such that $AB \ne BA$.

11 Prove that the distributive property holds for matrix addition and matrix multiplication. In other words, suppose A, B, C, D, E, and F are matrices whose sizes are such that $A(B + C)$ and $(D + E)F$ make sense. Explain why $AB + AC$ and $DF + EF$ both make sense and prove that

$$A(B + C) = AB + AC \quad \text{and} \quad (D + E)F = DF + EF.$$

12 Prove that matrix multiplication is associative. In other words, suppose A, B, and C are matrices whose sizes are such that $(AB)C$ makes sense. Explain why $A(BC)$ makes sense and prove that

$$(AB)C = A(BC).$$

Try to find a clean proof that illustrates the following quote from Emil Artin: "It is my experience that proofs involving matrices can be shortened by 50% if one throws the matrices out."

13 Suppose A is an n-by-n matrix and $1 \le j, k \le n$. Show that the entry in row j, column k, of A^3 (which is defined to mean AAA) is

$$\sum_{p=1}^{n} \sum_{r=1}^{n} A_{j,p} A_{p,r} A_{r,k}.$$

14 Suppose m and n are positive integers. Prove that the function $A \mapsto A^t$ is a linear map from $\mathbf{F}^{m,n}$ to $\mathbf{F}^{n,m}$.

7 נניח ש־ $T \in \mathcal{L}(V,W)$. נניח ש־ $w_1,\dots,w_n$ בסיס של W ו־ V מממד סופי. הוכיחו שקיים בסיס $v_1,\dots,v_m$ של V כך שכל האיברים בשורה הראשונה של $\mathcal{M}(T)$ [ביחס לבסיסים $v_1,\dots,v_m$ ו־ $w_1,\dots,w_n$] הם 0, למעט אולי איבר שערכו 1 בשורה הראשונה, בעמודה הראשונה.

בתרגיל זה, בשונה מתרגיל 5, אתם מקבלים בסיס נתון של W במקום האפשרות לבחור בסיס עבור W.

8 נניח ש־ A מטריצה $m \times n$ ו־ B מטריצה $n \times p$. הוכיחו ש־

$$(AB)_{j,\cdot} = A_{j,\cdot} B$$

לכל $1 \le j \le m$. כלומר, הראו ששורה j של AB שווה ל־ (שורה j של A) כפול B.

תרגיל זה הוא גרסת השורה של 3.48.

9 נניח ש־ $a = \begin{pmatrix} a_1 & \cdots & a_n \end{pmatrix}$ מטריצה $1 \times n$ ו־ B מטריצה $n \times p$. הוכיחו ש־

$$aB = a_1 B_{1,\cdot} + \cdots + a_n B_{n,\cdot} .$$

כלומר, הראו ש־ aB היא צירוף ליניארי של שורות B, כאשר הסקלרים המכפילים את השורות מגיעים מ־ a.

תרגיל זה הוא גרסת השורה של 3.50.

10 הביאו דוגמה ל־ A ו־ B, מטריצות 2×2, כך ש־ $AB \ne BA$.

11 הוכיחו שתכונת הדיסטריבוטיביות מתקיימת עבור חיבור וכפל מטריצות. במילים אחרות, נניח ש־ A, B, C, D, E, ו־ F מטריצות שגודלן כזה ש־ $A(B+C)$ ו־ $(D+E)F$ מוגדרות. הסבירו מדוע $AB+AC$ ו־ $DF+EF$ שתיהן מוגדרות ובנוסף הוכיחו כי

$$A(B+C) = AB + AC \quad \text{וגם} \quad (D+E)F = DF + EF.$$

12 הוכיחו שכפל מטריצות הוא אסוציאטיבי. במילים אחרות, נניח ש־ A, B, ו־ C מטריצות שגודלן כזה ש־ $(AB)C$ מוגדרת. הסבירו מדוע $A(BC)$ מוגדרת והוכיחו כי

$$(AB)C = A(BC).$$

נסו למצוא הוכחה אלגנטית המפחישה את הציטוט הבא של אמיל ארטין: "מניסיוני, ניתן לקצר הוכחות העוסקות במטריצות ב־ 50% אם משליכים את המטריצות החוצה."

13 נניח ש־ A מטריצה $n \times n$ ו־ $1 \le j,k \le n$. הראו שהאיבר בשורה j, עמודה k, של A^3 (המוגדרת כ־ AAA) הוא

$$\sum_{p=1}^{n} \sum_{r=1}^{n} A_{j,p} A_{p,r} A_{r,k}.$$

14 נניח ש־ m ו־ n מספרים שלמים חיוביים. הוכיחו שהפונקציה $A \mapsto A^t$ היא העתקה ליניארית מ־ $\mathbf{F}^{m,n}$ אל $\mathbf{F}^{n,m}$.

15 Prove that if A is an m-by-n matrix and C is an n-by-p matrix, then

$$(AC)^t = C^t A^t.$$

This exercise shows that the transpose of the product of two matrices is the product of the transposes in the opposite order.

16 Suppose A is an m-by-n matrix with $A \neq 0$. Prove that the rank of A is 1 if and only if there exist $(c_1, \dots, c_m) \in \mathbf{F}^m$ and $(d_1, \dots, d_n) \in \mathbf{F}^n$ such that $A_{j,k} = c_j d_k$ for every $j = 1, \dots, m$ and every $k = 1, \dots, n$.

17 Suppose $T \in \mathcal{L}(V)$, and $u_1, \dots, u_n$ and $v_1, \dots, v_n$ are bases of V. Prove that the following are equivalent.

(a) T is injective.

(b) The columns of $\mathcal{M}(T)$ are linearly independent in $\mathbf{F}^{n,1}$.

(c) The columns of $\mathcal{M}(T)$ span $\mathbf{F}^{n,1}$.

(d) The rows of $\mathcal{M}(T)$ span $\mathbf{F}^{1,n}$.

(e) The rows of $\mathcal{M}(T)$ are linearly independent in $\mathbf{F}^{1,n}$.

Here $\mathcal{M}(T)$ means $\mathcal{M}(T, (u_1, \dots, u_n), (v_1, \dots, v_n))$.

15 הוכיחו שאם A מטריצה $m \times n$ ו־ C מטריצה $n \times p$, אזי
$$(AC)^t = C^t A^t.$$
תרגיל זה מראה שהמשוחלפת של מכפלת שתי מטריצות היא מכפלת המשוחלפות בסדר הפוך.

16 נניח ש־ A מטריצה $m \times n$ ו־ $A \neq 0$. הוכיחו שדרגת A היא 1 אם־ורק־אם קיימים $(c_1, \dots, c_m) \in \mathbf{F}^m$ ו־ $(d_1, \dots, d_n) \in \mathbf{F}^n$ כך ש־ $A_{j,k} = c_j d_k$ לכל $j = 1, \dots, m$ ולכל $k = 1, \dots, n$.

17 נניח ש־ $T \in \mathcal{L}(V)$, ו־ $u_1, \dots, u_n$ ו־ $v_1, \dots, v_n$ בסיסים של V. הוכיחו שהתנאים הבאים שקולים.

א. T חד־חד־ערכית.

ב. עמודות $\mathcal{M}(T)$ בלתי תלויות ליניארית ב־ $\mathbf{F}^{n,1}$.

ג. עמודות $\mathcal{M}(T)$ פורסות את $\mathbf{F}^{n,1}$.

ד. שורות $\mathcal{M}(T)$ פורסות את $\mathbf{F}^{1,n}$.

ה. שורות $\mathcal{M}(T)$ בלתי תלויות ליניארית ב־ $\mathbf{F}^{1,n}$.

הסימון $\mathcal{M}(T)$ כאן משמעותו $\mathcal{M}(T, (u_1, \dots, u_n), (v_1, \dots, v_n))$.

אלגברה ליניארית בדרך הישר, מהדורה רביעית, מחבר: שלדון אקסלר
Linear Algebra Done Right, fourth edition, by Sheldon Axler

3D Invertibility and Isomorphisms

Invertible Linear Maps

We begin this section by defining the notions of invertible and inverse in the context of linear maps.

3.59 definition: *invertible, inverse*

- A linear map $T \in \mathcal{L}(V, W)$ is called *invertible* if there exists a linear map $S \in \mathcal{L}(W, V)$ such that ST equals the identity operator on V and TS equals the identity operator on W.
- A linear map $S \in \mathcal{L}(W, V)$ satisfying $ST = I$ and $TS = I$ is called an *inverse* of T (note that the first I is the identity operator on V and the second I is the identity operator on W).

The definition above mentions "an inverse". However, the next result shows that we can change this terminology to "the inverse".

3.60 *inverse is unique*

An invertible linear map has a unique inverse.

Proof Suppose $T \in \mathcal{L}(V, W)$ is invertible and S_1 and S_2 are inverses of T. Then

$$S_1 = S_1 I = S_1(TS_2) = (S_1 T)S_2 = IS_2 = S_2.$$

Thus $S_1 = S_2$. ∎

Now that we know that the inverse is unique, we can give it a notation.

3.61 notation: T^{-1}

If T is invertible, then its inverse is denoted by T^{-1}. In other words, if $T \in \mathcal{L}(V, W)$ is invertible, then T^{-1} is the unique element of $\mathcal{L}(W, V)$ such that $T^{-1}T = I$ and $TT^{-1} = I$.

3.62 example: *inverse of a linear map from* $\mathbf{R}^3$ *to* $\mathbf{R}^3$

Suppose $T \in \mathcal{L}(\mathbf{R}^3)$ is defined by $T(x, y, z) = (-y, x, 4z)$. Thus T is a counterclockwise rotation by 90° in the xy-plane and a stretch by a factor of 4 in the direction of the z-axis.

Hence the inverse map $T^{-1} \in \mathcal{L}(\mathbf{R}^3)$ is the clockwise rotation by 90° in the xy-plane and a stretch by a factor of $\frac{1}{4}$ in the direction of the z-axis:

$$T^{-1}(x, y, z) = (y, -x, \tfrac{1}{4}z).$$

ג4 הַפִּיכוּת ואיזומורפיזמים

העתקות ליניאריות הַפִּיכוֹת

אנו פותחים יחידה זו בהגדרת הרעיונות של הַפִּיכוּת והוֹפְכִיוּת בהקשר של העתקות ליניאריות.

3.59 הגדרה: הַפִּיכָה, הוֹפְכִית

- העתקה ליניארית $T \in \mathcal{L}(V,W)$ נקראת הַפִּיכָה אם קיימת העתקה ליניארית $S \in \mathcal{L}(W,V)$ כך ש־ ST שווה לאופרטור הזהות מעל V ו־ TS שווה לאופרטור הזהות מעל W.
- העתקה ליניארית $S \in \mathcal{L}(W,V)$ המקיימת $ST = I$ ו־ $TS = I$ נקראת הוֹפְכִית של T (שימו לב כי ה־ I הראשון הוא אופרטור הזהות מעל V ואילו ה־ I השני הוא אופרטור הזהות מעל W).

ההגדרה לעיל מציינת "הופכית". אך לפי מה שנראה במסקנה הבאה, נוכל לשנות את המינוח ל"ההופכית".

3.60 הופכית היא יחידה

להעתקה ליניארית הפיכה יש הופכית יחידה.

הוכחה נניח ש־ $T \in \mathcal{L}(V,W)$ הפיכה ו־ S_1 ו־ S_2 הופכיות של T. אזי

$$S_1 = S_1 I = S_1(TS_2) = (S_1T)S_2 = IS_2 = S_2.$$

לפיכך $S_1 = S_2$. ■

כעת כשאנו יודעים שההופכית יחידה נוכל להעניק לה סימון.

3.61 סימון: T^{-1}

אם T הפיכה, אזי ההופכית שלה מסומנת ב־ T^{-1}. במילים אחרות, אם $T \in \mathcal{L}(V,W)$ הפיכה, אזי T^{-1} היא האיבר היחיד של $\mathcal{L}(W,V)$ כך ש־ $T^{-1}T = I$ ו־ $TT^{-1} = I$.

3.62 דוגמה: הופכית של העתקה ליניארית מ־ $\mathbf{R}^3$ ל־ $\mathbf{R}^3$

נניח ש־ $T \in \mathcal{L}(\mathbf{R}^3)$ מוגדרת על ידי $T(x,y,z) = (-y,x,4z)$. לפיכך, T היא סיבוב של $90°$ נגד כיוון השעון במישור xy ומתיחה פי 4 בכיוון ציר ה־ z.

לכן, ההעתקה ההופכית $T^{-1} \in \mathcal{L}(\mathbf{R}^3)$ היא סיבוב של $90°$ בכיוון השעון במישור xy ומתיחה פי $\frac{1}{4}$ בכיוון ציר ה־ z:

$$T^{-1}(x,y,z) = (y,-x,\tfrac{1}{4}z).$$

The next result shows that a linear map is invertible if and only if it is one-to-one and onto.

3.63 *invertibility ⟺ injectivity and surjectivity*

A linear map is invertible if and only if it is injective and surjective.

Proof Suppose $T \in \mathcal{L}(V, W)$. We need to show that T is invertible if and only if it is injective and surjective.

First suppose T is invertible. To show that T is injective, suppose $u, v \in V$ and $Tu = Tv$. Then

$$u = T^{-1}(Tu) = T^{-1}(Tv) = v,$$

so $u = v$. Hence T is injective.

We are still assuming that T is invertible. Now we want to prove that T is surjective. To do this, let $w \in W$. Then $w = T(T^{-1}w)$, which shows that w is in the range of T. Thus $\text{range}\, T = W$. Hence T is surjective, completing this direction of the proof.

Now suppose T is injective and surjective. We want to prove that T is invertible. For each $w \in W$, define $S(w)$ to be the unique element of V such that $T(S(w)) = w$ (the existence and uniqueness of such an element follow from the surjectivity and injectivity of T). The definition of S implies that $T \circ S$ equals the identity operator on W.

To prove that $S \circ T$ equals the identity operator on V, let $v \in V$. Then

$$T((S \circ T)v) = (T \circ S)(Tv) = I(Tv) = Tv.$$

This equation implies that $(S \circ T)v = v$ (because T is injective). Thus $S \circ T$ equals the identity operator on V.

To complete the proof, we need to show that S is linear. To do this, suppose $w_1, w_2 \in W$. Then

$$T(S(w_1) + S(w_2)) = T(S(w_1)) + T(S(w_2)) = w_1 + w_2.$$

Thus $S(w_1) + S(w_2)$ is the unique element of V that T maps to $w_1 + w_2$. By the definition of S, this implies that $S(w_1 + w_2) = S(w_1) + S(w_2)$. Hence S satisfies the additive property required for linearity.

The proof of homogeneity is similar. Specifically, if $w \in W$ and $\lambda \in \mathbf{F}$, then

$$T(\lambda S(w)) = \lambda T(S(w)) = \lambda w.$$

Thus $\lambda S(w)$ is the unique element of V that T maps to λw. By the definition of S, this implies that $S(\lambda w) = \lambda S(w)$. Hence S is linear, as desired. ∎

For a linear map from a vector space to itself, you might wonder whether injectivity alone, or surjectivity alone, is enough to imply invertibility. On infinite-dimensional vector spaces, neither condition alone implies invertibility, as illustrated by the next example, which uses two familiar linear maps from Example 3.3.

המסקנה הבאה מראה כי העתקה ליניארית היא הפיכה אם־ורק־אם היא העתקה חד־חד־ערכית ועל.

3.63 הַפִּיכוּת $\iff$ חד־חד־ערכיות ותכונת העל

העתקה ליניארית היא הפיכה אם־ורק־אם היא חד־חד־ערכית ועל.

הוכחה נניח ש־ $T \in \mathcal{L}(V,W)$. עלינו להראות ש־ T הפיכה אם־ורק־אם היא חד־חד־ערכית ועל.

ראשית נניח כי T הפיכה. כדי להראות ש־ T חד־חד־ערכית, נניח ש־ $u, v \in V$ ו־ $Tu = Tv$. אזי

$$u = T^{-1}(Tu) = T^{-1}(Tv) = v,$$

ומכאן ש־ $u = v$. לכן T חד־חד־ערכית.

אנו עדיין מניחים ש־ T הפיכה. נרצה כעת להוכיח ש־ T על. לצורך כך, יהי $w \in W$. אזי $w = T(T^{-1}w)$, מה שמראה ש־ w בטווח של T. לפיכך $\operatorname{range} T = W$. לכן T על ובכך השלמנו כיוון זה של ההוכחה.

כעת נניח ש־ T חד־חד־ערכית ועל. ברצוננו להוכיח כי T הפיכה. לכל $w \in W$, נגדיר את $S(w)$ כאיבר היחיד של V כך ש־ $T(S(w)) = w$ (קיומו ויחידותו של איבר זה הם תוצאה ישירה של היות T על וחד־חד־ערכית). מהגדרת S נסיק ש־ $T \circ S$ שווה לאופרטור הזהות מעל W.

כדי להוכיח ש־ $S \circ T$ שווה לאופרטור הזהות מעל V, יהי $v \in V$. אזי

$$T((S \circ T)v) = (T \circ S)(Tv) = I(Tv) = Tv.$$

משוואה זו גוררת כי $(S \circ T)v = v$ (מאחר ש־ T חד־חד־ערכית). לפיכך $S \circ T$ שווה לאופרטור הזהות מעל V.

להשלמת ההוכחה, עלינו להראות ש־ S ליניארית. לשם כך, נניח כי $w_1, w_2 \in W$. אזי

$$T(S(w_1) + S(w_2)) = T(S(w_1)) + T(S(w_2)) = w_1 + w_2.$$

לפיכך $S(w_1) + S(w_2)$ הוא האיבר היחיד של V שאותו T מעתיקה ל־ $w_1 + w_2$. מהגדרת S נסיק ש־ $S(w_1 + w_2) = S(w_1) + S(w_2)$. לכן S מקיימת את האדיטיביות הנדרשת לליניאריות.

ההוכחה להומוגניות היא דומה. ליתר דיוק, אם $w \in W$ ו־ $\lambda \in \mathbf{F}$, אזי

$$T(\lambda S(w)) = \lambda T(S(w)) = \lambda w.$$

לפיכך $\lambda S(w)$ הוא האיבר היחיד של V שאותו T מעתיקה ל־ λw. לפי הגדרת S, מכאן נסיק ש־ $S(\lambda w) = \lambda S(w)$. לכן S ליניארית כפי שרצינו להוכיח. ∎

עבור העתקה ליניארית ממרחב וקטורי לעצמו, ייתכן שתתהו האם חד־חד־ערכיות בלבד, או על בלבד, מספיקה כדי לקבוע שההעתקה הפיכה. מעל מרחב וקטורי מממד אינסופי, מקיום אחד מהתנאים לא נוכל להסיק הפיכות, כפי שנראה בדוגמה הבאה, המשתמשת בשתי העתקות מֻכָּרוֹת מדוגמה 3.3.

3.64 example: *neither injectivity nor surjectivity implies invertibility*

- The multiplication by x^2 linear map from $\mathcal{P}(\mathbf{R})$ to $\mathcal{P}(\mathbf{R})$ (see 3.3) is injective but it is not invertible because it is not surjective (the polynomial 1 is not in the range).
- The backward shift linear map from $\mathbf{F}^\infty$ to $\mathbf{F}^\infty$ (see 3.3) is surjective but it is not invertible because it is not injective [the vector $(1, 0, 0, 0, \dots)$ is in the null space].

In view of the example above, the next result is remarkable—it states that for a linear map from a finite-dimensional vector space to a vector space of the same dimension, either injectivity or surjectivity alone implies the other condition. Note that the hypothesis below that $\dim V = \dim W$ is automatically satisfied in the important special case where V is finite-dimensional and $W = V$.

3.65 *injectivity is equivalent to surjectivity* (*if* $\dim V = \dim W < \infty$)

Suppose that V and W are finite-dimensional vector spaces, $\dim V = \dim W$, and $T \in \mathcal{L}(V, W)$. Then

$$T \text{ is invertible} \iff T \text{ is injective} \iff T \text{ is surjective.}$$

Proof The fundamental theorem of linear maps (3.21) states that

3.66 $$\dim V = \dim \operatorname{null} T + \dim \operatorname{range} T.$$

If T is injective (which by 3.15 is equivalent to the condition $\dim \operatorname{null} T = 0$), then the equation above implies that

$$\dim \operatorname{range} T = \dim V - \dim \operatorname{null} T = \dim V = \dim W,$$

which implies that T is surjective (by 2.39).

Conversely, if T is surjective, then 3.66 implies that

$$\dim \operatorname{null} T = \dim V - \dim \operatorname{range} T = \dim V - \dim W = 0,$$

which implies that T is injective.

Thus we have shown that T is injective if and only if T is surjective. Thus if T is either injective or surjective, then T is both injective and surjective, which implies that T is invertible. Hence T is invertible if and only if T is injective if and only if T is surjective. ∎

The next example illustrates the power of the previous result. Although it is possible to prove the result in the example below without using linear algebra, the proof using linear algebra is cleaner and easier.

3.64 דוגמה: חד־חד־ערכיות או על אינה גוררת הפיכות

- ההעתקה הליניארית של ההכפלה ב־ x^2 מ־ $\mathcal{P}(\mathbf{R})$ אל $\mathcal{P}(\mathbf{R})$ (ראו 3.3) היא חד־חד־ערכית אך אינה הפיכה, מאחר שאינה על (הפולינום 1 אינו בטווח).
- ההעתקה הליניארית של ההסטה לאחור מ־ $\mathbf{F}^\infty$ ל־ $\mathbf{F}^\infty$ (ראו 3.3) היא על אך אינה הפיכה, מאחר שאינה חד־חד־ערכית [הווקטור $(1, 0, 0, 0, \dots)$ נמצא במרחב האפס].

לאור הדוגמה לעיל, המסקנה הבאה ראויה לציון—היא קובעת כי להעתקה ליניארית ממרחב וקטורי מממד סופי למרחב וקטורי מאותו הממד, די בקיום אחת התכונות — חד־חד־ערכיות או על — כדי להסיק את קיום האחרת. נשים לב כי ההנחה להלן ש־ $\dim V = \dim W$ מתקיימת באופן אוטומטי במקרה המיוחד והחשוב שבו V מממד סופי ו־ $W = V$.

3.65 חד־חד־ערכיות שקולה לעל (אם $\dim V = \dim W < \infty$)

נניח ש־ V ו־ W מרחבים וקטוריים מממד סופי, $\dim V = \dim W$, ו־ $T \in \mathcal{L}(V, W)$. אזי T הפיכה אם־ורק־אם T חד־חד־ערכית אם־ורק־אם T על.

הוכחה המשפט היסודי של ההעתקות הליניאריות (3.21) קובע כי

$$\dim V = \dim \operatorname{null} T + \dim \operatorname{range} T. \tag{3.66}$$

אם T חד־חד־ערכית (מה שלפי 3.15 שקול לכך ש־ $\dim \operatorname{null} T = 0$), אזי מהמשוואה לעיל נסיק ש־

$$\dim \operatorname{range} T = \dim V - \dim \operatorname{null} T = \dim V = \dim W,$$

מה שגורר כי T על (לפי 2.39).

בכיוון ההפוך, אם T על, אזי לפי 3.66 נסיק ש־

$$\dim \operatorname{null} T = \dim V - \dim \operatorname{range} T = \dim V - \dim W = 0,$$

מה שגורר כי T חד־חד־ערכית.

לפיכך הראינו ש־ T חד־חד־ערכית אם־ורק־אם T על. לפיכך אם T חד־חד־ערכית או על, אזי T מקיימת את שתי התכונות, חד־חד־ערכית ועל, מה שמוביל למסקנה ש־ T הפיכה. לכן T הפיכה אם־ורק־אם T חד־חד־ערכית אם־ורק־אם T על. ■

הדוגמה הבאה ממחישה את עוצמת המסקנה הקודמת. אמנם ניתן להוכיח את התוצאה שבדוגמה להלן ללא שימוש באלגברה ליניארית, אך ההוכחה המשתמשת באלגברה ליניארית היא אלגנטית ופשוטה יותר.

3.67 example: *there exists a polynomial p such that $((x^2 + 5x + 7)p)'' = q$*

The linear map

$$p \mapsto ((x^2 + 5x + 7)p)''$$

from $\mathcal{P}(\mathbf{R})$ to itself is injective, as you can show. Thus we are tempted to use 3.65 to show that this map is surjective. However, Example 3.64 shows that the magic of 3.65 does not apply to the infinite-dimensional vector space $\mathcal{P}(\mathbf{R})$. We will get around this problem by restricting attention to the finite-dimensional vector space $\mathcal{P}_m(\mathbf{R})$.

Suppose $q \in \mathcal{P}(\mathbf{R})$. There exists a nonnegative integer m such that $q \in \mathcal{P}_m(\mathbf{R})$. Define $T : \mathcal{P}_m(\mathbf{R}) \to \mathcal{P}_m(\mathbf{R})$ by

$$Tp = ((x^2 + 5x + 7)p)''.$$

Multiplying a nonzero polynomial by $(x^2 + 5x + 7)$ increases the degree by 2, and then differentiating twice reduces the degree by 2. Thus T is indeed a linear map from $\mathcal{P}_m(\mathbf{R})$ to itself.

Every polynomial whose second derivative equals 0 is of the form $ax + b$, where $a, b \in \mathbf{R}$. Thus $\operatorname{null} T = \{0\}$. Hence T is injective.

Thus T is surjective (by 3.65), which means that there exists a polynomial $p \in \mathcal{P}_m(\mathbf{R})$ such that $((x^2 + 5x + 7)p)'' = q$, as claimed in the title of this example.

Exercise 35 in Section 6A gives a similar but more spectacular example of using 3.65.

The hypothesis in the result below that $\dim V = \dim W$ holds in the important special case in which V is finite-dimensional and $W = V$. Thus in that case, the equation $ST = I$ implies that $ST = TS$, even though we do not have multiplicative commutativity of arbitrary linear maps from V to V.

3.68 $ST = I \iff TS = I$ (*on vector spaces of the same dimension*)

Suppose V and W are finite-dimensional vector spaces of the same dimension, $S \in \mathcal{L}(W, V)$, and $T \in \mathcal{L}(V, W)$. Then $ST = I$ if and only if $TS = I$.

Proof First suppose $ST = I$. If $v \in V$ and $Tv = 0$, then

$$v = Iv = (ST)v = S(Tv) = S(0) = 0.$$

Thus T is injective (by 3.15). Because V and W have the same dimension, this implies that T is invertible (by 3.65).

Now multiply both sides of the equation $ST = I$ by T^{-1} on the right, getting

$$S = T^{-1}.$$

Thus $TS = TT^{-1} = I$, as desired.

To prove the implication in the other direction, simply reverse the roles of S and T (and V and W) in the direction we have already proved, showing that if $TS = I$, then $ST = I$. ∎

3.67 דוגמה: קיים פולינום p כך ש־ $((x^2+5x+7)p)'' = q$

ההעתקה הליניארית

$$p \mapsto ((x^2+5x+7)p)''$$

מ־ $\mathcal{P}(\mathbf{R})$ לעצמו היא חד־חד־ערכית, כפי שאתם יכולים להראות. לפיכך, נתפתה להשתמש ב־ 3.65 כדי להראות שהעתקה זו היא על. אולם דוגמה 3.64 מראה שהקסם של 3.65 אינו חל על המרחב הווקטורי מממד אינסופי $\mathcal{P}(\mathbf{R})$. את מגבלה זו נעקוף בכך שנתמקד במרחב הווקטורי מממד סופי $\mathcal{P}_m(\mathbf{R})$.

נניח ש־ $q \in \mathcal{P}(\mathbf{R})$. קיים מספר שלם אי־שלילי m כך ש־ $q \in \mathcal{P}_m(\mathbf{R})$. נגדיר $T : \mathcal{P}_m(\mathbf{R}) \to \mathcal{P}_m(\mathbf{R})$ על ידי

$$Tp = ((x^2+5x+7)p)''.$$

הכפלת פולינום השונה מפולינום האפס ב־ (x^2+5x+7) מגדילה את המעלה ב־ 2, ואז גזירה כפולה מצמצמת את המעלה ב־ 2. לפיכך T היא בהחלט העתקה ליניארית מ־ $\mathcal{P}_m(\mathbf{R})$ לעצמו.

כל פולינום שנגזרתו השנייה היא 0 הוא מהצורה $ax+b$, כאשר $a, b \in \mathbf{R}$. לפיכך $\operatorname{null} T = \{0\}$. לכן T חד־חד־ערכית.

לפיכך T היא העתקה על (לפי 3.65), כלומר, קיים פולינום $p \in \mathcal{P}_m(\mathbf{R})$ כך ש־ $((x^2+5x+7)p)'' = q$, כפי שנטען בכותרת הדוגמה.

תרגיל 35 ביחידה 1I נותן דוגמה דומה אך מרהיבה יותר של שימוש ב־ 3.65.

ההנחה במסקנה להלן ש־ $\dim V = \dim W$ מתקיימת במקרה המיוחד והחשוב שבו V מממד סופי ו־ $W = V$. לפיכך במקרה זה, המשוואה $ST = I$ מובילה למסקנה כי $ST = TS$, אף שתכונת הקומוטטיביות של הכפל אינה מתקיימת בהעתקות ליניאריות שרירותיות מ־ V אל V.

3.68 $ST = I$ אם־ורק־אם $TS = I$ (מעל מרחבים וקטוריים מאותו הממד)

נניח ש־ V ו־ W מרחבים וקטוריים מממד סופי ומאותו הממד, $S \in \mathcal{L}(W, V)$, ו־ $T \in \mathcal{L}(V, W)$. אזי $ST = I$ אם־ורק־אם $TS = I$.

הוכחה תחילה נניח ש־ $ST = I$. אם $v \in V$ ו־ $Tv = 0$, אזי

$$v = Iv = (ST)v = S(Tv) = S(0) = 0.$$

לפיכך T חד־חד־ערכית (לפי 3.15). מאחר שממדי V ו־ W זהים, נסיק ש־ T הפיכה (לפי 3.65).

כעת נכפיל את שני אגפי המשוואה $ST = I$ ב־ T^{-1} מימין ונקבל

$$S = T^{-1}.$$

לפיכך $TS = TT^{-1} = I$, כנדרש.

להוכחת התוצאה בכיוון ההפוך, פשוט נהפוך את תפקידי S ו־ T (ו־ V ו־ W) בכיוון שכבר הוכחנו, ובכך נראה שאם $TS = I$ אזי $ST = I$. ■

Isomorphic Vector Spaces

The next definition captures the idea of two vector spaces that are essentially the same, except for the names of their elements.

3.69 definition: *isomorphism, isomorphic*

- An *isomorphism* is an invertible linear map.
- Two vector spaces are called *isomorphic* if there is an isomorphism from one vector space onto the other one.

Think of an isomorphism $T : V \to W$ as relabeling $v \in V$ as $Tv \in W$. This viewpoint explains why two isomorphic vector spaces have the same vector space properties. The terms "isomorphism" and "invertible linear map" mean the same thing. Use "isomorphism" when you want to emphasize that the two spaces are essentially the same.

It can be difficult to determine whether two mathematical structures (such as groups or topological spaces) are essentially the same, differing only in the names of the elements of underlying sets. However, the next result shows that we need to look at only a single number (the dimension) to determine whether two vector spaces are isomorphic.

3.70 *dimension shows whether vector spaces are isomorphic*

Two finite-dimensional vector spaces over $\mathbf{F}$ are isomorphic if and only if they have the same dimension.

Proof First suppose V and W are isomorphic finite-dimensional vector spaces. Thus there exists an isomorphism T from V onto W. Because T is invertible, we have $\operatorname{null} T = \{0\}$ and $\operatorname{range} T = W$. Thus

$$\dim \operatorname{null} T = 0 \quad \text{and} \quad \dim \operatorname{range} T = \dim W.$$

The formula

$$\dim V = \dim \operatorname{null} T + \dim \operatorname{range} T$$

(the fundamental theorem of linear maps, which is 3.21) thus becomes the equation $\dim V = \dim W$, completing the proof in one direction.

To prove the other direction, suppose V and W are finite-dimensional vector spaces of the same dimension. Let $v_1, \dots, v_n$ be a basis of V and $w_1, \dots, w_n$ be a basis of W. Let $T \in \mathcal{L}(V, W)$ be defined by

$$T(c_1 v_1 + \cdots + c_n v_n) = c_1 w_1 + \cdots + c_n w_n.$$

Then T is a well-defined linear map because $v_1, \dots, v_n$ is a basis of V. Also, T is surjective because $w_1, \dots, w_n$ spans W. Furthermore, $\operatorname{null} T = \{0\}$ because $w_1, \dots, w_n$ is linearly independent. Thus T is injective. Because T is injective and surjective, it is an isomorphism (see 3.63). Hence V and W are isomorphic. ∎

מרחבים וקטוריים איזומורפיים

ההגדרה הבאה לוכדת את הרעיון של שני מרחבים וקטוריים שהם זהים במהותם, מלבד השוני בשמות איבריהם.

3.69 הגדרה: אִיזוֹמוֹרְפִיזְם, אִיזוֹמוֹרְפִי

- אִיזוֹמוֹרְפִיזְם הוא העתקה ליניארית הפיכה.
- שני מרחבים וקטוריים נקראים אִיזוֹמוֹרְפִיִים אם קיים אִיזוֹמוֹרְפִיזְם ממרחב וקטורי אחד על האחר.

נחשוב על איזומורפיזם $T: V \to W$ כעל תיוג מחדש של $v \in V$ כ־ $Tv \in W$. נקודת מבט זו מסבירה מדוע לשני מרחבים וקטוריים איזומורפיים יש אותן תכונות מרחב וקטורי. המונח "איזומורפיזם" והמונח "העתקה ליניארית הפיכה" משמעותם זהה. נשתמש ב"איזומורפיזם" כשנרצה להדגיש את העובדה ששני המרחבים הם במהותם אותו מרחב.

עשוי להיות קשה לקבוע האם שני מבנים מתמטיים (כגון חבורות או מרחבים טופולוגיים) הם שווים במהותם, ושונים רק בשמות איברי קבוצות הנושא. ובכל זאת, במסקנה הבאה נראה שעלינו לבחון רק מספר אחד (הממד) כדי לקבוע האם שני מרחבים וקטוריים הם איזומורפיים.

3.70 הממד מורה האם שני מרחבים וקטוריים הם איזומורפיים

שני מרחבים וקטוריים מממד סופי מעל $\mathbf{F}$ הם איזומורפיים אם־ורק־אם לשניהם אותו ממד.

הוכחה ראשית נניח ש־ V ו־ W מרחבים וקטוריים איזומורפיים מממד סופי. לפיכך קיים איזומורפיזם T מ־ V על W. מאחר ש־ T הפיכה, מתקיים $\operatorname{null} T = \{0\}$ וגם $\operatorname{range} T = W$. לפיכך

$$\dim \operatorname{null} T = 0 \quad \text{וגם} \quad \dim \operatorname{range} T = \dim W.$$

הנוסחה

$$\dim V = \dim \operatorname{null} T + \dim \operatorname{range} T$$

(המשפט היסודי של ההעתקות הליניאריות המופיע ב־ 3.21) הופכת אפוא למשוואה $\dim V = \dim W$, ובכך השלמנו את ההוכחה בכיוון אחד.

להוכחת הכיוון השני, נניח ש־ V ו־ W מרחבים וקטוריים מממד סופי זהה. יהי $v_1, \dots, v_n$ בסיס של V ו־ $w_1, \dots, w_n$ בסיס של W. תהא $T \in \mathcal{L}(V, W)$ המוגדרת על ידי

$$T(c_1 v_1 + \cdots + c_n v_n) = c_1 w_1 + \cdots + c_n w_n.$$

אז T היא העתקה ליניארית מוגדרת היטב שכן $v_1, \dots, v_n$ בסיס של V. בנוסף, T על, מאחר ש־ $w_1, \dots, w_n$ פורסת את W. יתר על כן, $\operatorname{null} T = \{0\}$ שכן $w_1, \dots, w_n$ בלתי תלויה ליניארית. לפיכך T חד־חד־ערכית. מכיוון ש־ T חד־חד־ערכית ועל, היא איזומורפיזם (ראו 3.63). לכן V ו־ W איזומורפיים. ∎

The previous result implies that each finite-dimensional vector space V is isomorphic to $\mathbf{F}^n$, where $n = \dim V$. For example, if m is a nonnegative integer, then $\mathcal{P}_m(\mathbf{F})$ is isomorphic to $\mathbf{F}^{m+1}$.

Recall that the notation $\mathbf{F}^{m,n}$ denotes the vector space of m-by-n matrices with entries in $\mathbf{F}$. If $v_1, \dots, v_n$ is a basis of V and $w_1, \dots, w_m$ is a basis of W, then for each $T \in \mathcal{L}(V, W)$, we have a matrix $\mathcal{M}(T) \in \mathbf{F}^{m,n}$. Thus once bases have been fixed for V and W, $\mathcal{M}$ becomes a function from $\mathcal{L}(V, W)$ to $\mathbf{F}^{m,n}$. Notice that 3.35 and 3.38 show that $\mathcal{M}$ is a linear map. This linear map is actually an isomorphism, as we now show.

Every finite-dimensional vector space is isomorphic to some $\mathbf{F}^n$. Thus why not just study $\mathbf{F}^n$ instead of more general vector spaces? To answer this question, note that an investigation of $\mathbf{F}^n$ would soon lead to other vector spaces. For example, we would encounter the null space and range of linear maps. Although each of these vector spaces is isomorphic to some $\mathbf{F}^m$, thinking of them that way often adds complexity but no new insight.

3.71 *$\mathcal{L}(V, W)$ and $\mathbf{F}^{m,n}$ are isomorphic*

Suppose $v_1, \dots, v_n$ is a basis of V and $w_1, \dots, w_m$ is a basis of W. Then $\mathcal{M}$ is an isomorphism between $\mathcal{L}(V, W)$ and $\mathbf{F}^{m,n}$.

Proof We already noted that $\mathcal{M}$ is linear. We need to prove that $\mathcal{M}$ is injective and surjective.

We begin with injectivity. If $T \in \mathcal{L}(V, W)$ and $\mathcal{M}(T) = 0$, then $Tv_k = 0$ for each $k = 1, \dots, n$. Because $v_1, \dots, v_n$ is a basis of V, this implies $T = 0$. Thus $\mathcal{M}$ is injective (by 3.15).

To prove that $\mathcal{M}$ is surjective, suppose $A \in \mathbf{F}^{m,n}$. By the linear map lemma (3.4), there exists $T \in \mathcal{L}(V, W)$ such that

$$Tv_k = \sum_{j=1}^{m} A_{j,k} w_j$$

for each $k = 1, \dots, n$. Because $\mathcal{M}(T)$ equals A, the range of $\mathcal{M}$ equals $\mathbf{F}^{m,n}$, as desired. ∎

Now we can determine the dimension of the vector space of linear maps from one finite-dimensional vector space to another.

3.72 $\dim \mathcal{L}(V, W) = (\dim V)(\dim W)$

Suppose V and W are finite-dimensional. Then $\mathcal{L}(V, W)$ is finite-dimensional and

$$\dim \mathcal{L}(V, W) = (\dim V)(\dim W).$$

Proof The desired result follows from 3.71, 3.70, and 3.40. ∎

מהמסקנה הקודמת נסיק שכל מרחב וקטורי מממד סופי V הוא איזומורפי ל־ $\mathbf{F}^n$, כאשר $n = \dim V$. לדוגמה, אם m מספר שלם אי־שלילי, אזי $\mathcal{P}_m(\mathbf{F})$ איזומורפי ל־ $\mathbf{F}^{m+1}$.

כל מרחב וקטורי מממד סופי הוא איזומורפי ל־ $\mathbf{F}^n$ כלשהו. לפיכך, מזוע שלא נתמקד ב־ $\mathbf{F}^n$ במקום במרחבים וקטוריים כלליים יותר? כתשובה לשאלה זו, נזכיר שהעיסוק ב־ $\mathbf{F}^n$ יוביל אותנו למרחבים וקטוריים אחרים. לדוגמה, נגיע מהר מאוד לעסוק במרחב האפס ובטווח של העתקות ליניאריות. אמנם מרחבים אלו הם איזומורפיים ל־ $\mathbf{F}^m$ כלשהם, אך לעסוק בהם ככאלה, לרוב, יוסיף מורכבות מבלי לספק תובנות חדשות.

נזכיר כי הסימון $\mathbf{F}^{m,n}$ מסמן את המרחב הווקטורי של מטריצות $m \times n$ עם איברים ב־ $\mathbf{F}$. אם $v_1, \dots, v_n$ בסיס של V ו־ $w_1, \dots, w_m$ בסיס של W, אזי לכל $T \in \mathcal{L}(V, W)$, יש לנו מטריצה $\mathcal{M}(T) \in \mathbf{F}^{m,n}$. לפיכך משנקבעו בסיסים עבור V ו־ W, המטריצה M הופכת לפונקציה מ־ $\mathcal{L}(V, W)$ אל $\mathbf{F}^{m,n}$. שימו לב כי 3.35 ו־ 3.38 מראים ש־ $\mathcal{M}$ העתקה ליניארית. העתקה ליניארית זו היא למעשה איזומורפיזם, כפי שנראה כעת.

3.71 $\mathcal{L}(V, W)$ ו־ $\mathbf{F}^{m,n}$ איזומורפיים

נניח ש־ $v_1, \dots, v_n$ בסיס של V ו־ $w_1, \dots, w_m$ בסיס של W. אזי $\mathcal{M}$ היא איזומורפיזם בין $\mathcal{L}(V, W)$ ל־ $\mathbf{F}^{m,n}$.

הוכחה כבר ציינו את העובדה ש־ $\mathcal{M}$ ליניארית. עלינו להוכיח ש־ $\mathcal{M}$ חד־חד־ערכית ועל.

נתחיל בחד־חד־ערכיות. אם $T \in \mathcal{L}(V, W)$ ו־ $\mathcal{M}(T) = 0$, אזי $Tv_k = 0$ לכל $k = 1, \dots, n$. מאחר ש־ $v_1, \dots, v_n$ בסיס של V, מכאן נסיק ש־ $T = 0$. לפיכך $\mathcal{M}$ חד־חד־ערכית (לפי 3.15).

להוכחה ש־ $\mathcal{M}$ על, נניח ש־ $A \in \mathbf{F}^{m,n}$. לפי למת ההעתקה הליניארית (3.4), קיימת $T \in \mathcal{L}(V, W)$ כך ש־

$$Tv_k = \sum_{j=1}^{m} A_{j,k} w_j$$

לכל $k = 1, \dots, n$. מאחר ש־ $\mathcal{M}(T)$ שווה ל־ A, הטווח של $\mathcal{M}$ שווה ל־ $\mathbf{F}^{m,n}$, כפי שרצינו להוכיח. ■

כעת ביכולתנו לקבוע את ממד המרחב הווקטורי של העתקות ליניאריות ממרחב וקטורי מממד סופי אחד לאחר.

3.72 $\dim \mathcal{L}(V, W) = (\dim V)(\dim W)$

נניח ש־ V ו־ W מממד סופי. אזי $\mathcal{L}(V, W)$ הוא מממד סופי, ומתקיים
$$\dim \mathcal{L}(V, W) = (\dim V)(\dim W).$$

הוכחה המסקנה הרצויה נובעת מ־ 3.71, 3.70, ו־ 3.40. ■

Linear Maps Thought of as Matrix Multiplication

Previously we defined the matrix of a linear map. Now we define the matrix of a vector.

3.73 definition: *matrix of a vector*, $\mathcal{M}(v)$

Suppose $v \in V$ and $v_1, \dots, v_n$ is a basis of V. The *matrix of* v with respect to this basis is the n-by-1 matrix

$$\mathcal{M}(v) = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix},$$

where $b_1, \dots, b_n$ are the scalars such that

$$v = b_1 v_1 + \cdots + b_n v_n.$$

The matrix $\mathcal{M}(v)$ of a vector $v \in V$ depends on the basis $v_1, \dots, v_n$ of V, as well as on v. However, the basis should be clear from the context and thus it is not included in the notation.

3.74 example: *matrix of a vector*

- The matrix of the polynomial $2 - 7x + 5x^3 + x^4$ with respect to the standard basis of $\mathcal{P}_4(\mathbf{R})$ is

$$\begin{pmatrix} 2 \\ -7 \\ 0 \\ 5 \\ 1 \end{pmatrix}.$$

- The matrix of a vector $x \in \mathbf{F}^n$ with respect to the standard basis is obtained by writing the coordinates of x as the entries in an n-by-1 matrix. In other words, if $x = (x_1, \dots, x_n) \in \mathbf{F}^n$, then

$$\mathcal{M}(x) = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Occasionally we want to think of elements of V as relabeled to be n-by-1 matrices. Once a basis $v_1, \dots, v_n$ is chosen, the function $\mathcal{M}$ that takes $v \in V$ to $\mathcal{M}(v)$ is an isomorphism of V onto $\mathbf{F}^{n,1}$ that implements this relabeling.

Recall that if A is an m-by-n matrix, then $A_{\cdot,k}$ denotes the kth column of A, thought of as an m-by-1 matrix. In the next result, $\mathcal{M}(Tv_k)$ is computed with respect to the basis $w_1, \dots, w_m$ of W.

העתקות ליניאריות ככפל מטריצות

הגדרנו לעיל את המטריצה של העתקה ליניארית. כעת נגדיר מטריצה של וקטור.

3.73 הגדרה: מטריצה של וקטור, $\mathcal{M}(v)$

נניח ש־ $v \in V$ ו־ $v_1, \dots, v_n$ בסיס של V. המטריצה של v ביחס לבסיס זה היא המטריצה $n \times 1$

$$\mathcal{M}(v) = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix},$$

כאשר $b_1, \dots, b_n$ הם הסקלרים כך ש־

$$v = b_1 v_1 + \cdots + b_n v_n.$$

המטריצה $\mathcal{M}(v)$ של וקטור $v \in V$ תלויה בבסיס $v_1, \dots, v_n$ של V וגם ב־ v. עם זאת, הבסיס אמור להיות ידוע מההקשר, ולפיכך אינו נכלל בסימון.

3.74 דוגמה: מטריצה של וקטור

- המטריצה של הפולינום $2 - 7x + 5x^3 + x^4$ ביחס לבסיס הסטנדרטי של $\mathcal{P}_4(\mathbf{R})$ היא

$$\begin{pmatrix} 2 \\ -7 \\ 0 \\ 5 \\ 1 \end{pmatrix}.$$

- המטריצה של וקטור $x \in \mathbf{F}^n$ ביחס לבסיס הסטנדרטי מושגת באמצעות כתיבת הקואורדינטות של x כאיברים במטריצה $n \times 1$. כלומר, אם $x = (x_1, \dots, x_n) \in \mathbf{F}^n$, אזי

$$\mathcal{M}(x) = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

לעיתים נרצה לחשוב על איברים של V כמתויגים מחדש כמטריצות $n \times 1$. משנבחר בסיס $v_1, \dots, v_n$, הפונקציה $\mathcal{M}$ המעתיקה $v \in V$ אל $\mathcal{M}(v)$ היא איזומורפיזם של V על $\mathbf{F}^{n,1}$ המממשת תיוג זה.

נזכיר שאם A היא מטריצה $m \times n$, אזי $A_{\cdot,k}$ מסמנת את העמודה ה־ k^{th} של A, כשאנו חושבים עליה כמטריצה $m \times 1$. במסקנה הבאה, $\mathcal{M}(Tv_k)$ מחושבת ביחס לבסיס $w_1, \dots, w_m$ של W.

3.75 $\mathcal{M}(T)_{\cdot,k} = \mathcal{M}(Tv_k)$

Suppose $T \in \mathcal{L}(V, W)$ and $v_1, \dots, v_n$ is a basis of V and $w_1, \dots, w_m$ is a basis of W. Let $1 \le k \le n$. Then the kth column of $\mathcal{M}(T)$, which is denoted by $\mathcal{M}(T)_{\cdot,k}$, equals $\mathcal{M}(Tv_k)$.

Proof The desired result follows immediately from the definitions of $\mathcal{M}(T)$ and $\mathcal{M}(Tv_k)$. ∎

The next result shows how the notions of the matrix of a linear map, the matrix of a vector, and matrix multiplication fit together.

3.76 *linear maps act like matrix multiplication*

Suppose $T \in \mathcal{L}(V, W)$ and $v \in V$. Suppose $v_1, \dots, v_n$ is a basis of V and $w_1, \dots, w_m$ is a basis of W. Then

$$\mathcal{M}(Tv) = \mathcal{M}(T)\mathcal{M}(v).$$

Proof Suppose $v = b_1 v_1 + \cdots + b_n v_n$, where $b_1, \dots, b_n \in \mathbf{F}$. Thus

3.77 $$Tv = b_1 Tv_1 + \cdots + b_n Tv_n.$$

Hence

$$\begin{aligned}\mathcal{M}(Tv) &= b_1 \mathcal{M}(Tv_1) + \cdots + b_n \mathcal{M}(Tv_n)\\ &= b_1 \mathcal{M}(T)_{\cdot,1} + \cdots + b_n \mathcal{M}(T)_{\cdot,n}\\ &= \mathcal{M}(T)\mathcal{M}(v),\end{aligned}$$

where the first equality follows from 3.77 and the linearity of $\mathcal{M}$, the second equality comes from 3.75, and the last equality comes from 3.50. ∎

Each m-by-n matrix A induces a linear map from $\mathbf{F}^{n,1}$ to $\mathbf{F}^{m,1}$, namely the matrix multiplication function that takes $x \in \mathbf{F}^{n,1}$ to $Ax \in \mathbf{F}^{m,1}$. The result above can be used to think of every linear map (from a finite-dimensional vector space to another finite-dimensional vector space) as a matrix multiplication map after suitable relabeling via the isomorphisms given by $\mathcal{M}$. Specifically, if $T \in \mathcal{L}(V, W)$ and we identify $v \in V$ with $\mathcal{M}(v) \in \mathbf{F}^{n,1}$, then the result above says that we can identify Tv with $\mathcal{M}(T)\mathcal{M}(v)$.

Because the result above allows us to think (via isomorphisms) of each linear map as multiplication on $\mathbf{F}^{n,1}$ by some matrix A, keep in mind that the specific matrix A depends not only on the linear map but also on the choice of bases. One of the themes of many of the most important results in later chapters will be the choice of a basis that makes the matrix A as simple as possible.

In this book, we concentrate on linear maps rather than on matrices. However, sometimes thinking of linear maps as matrices (or thinking of matrices as linear maps) gives important insights that we will find useful.

3.75 $$\mathcal{M}(T)_{\cdot,k} = \mathcal{M}(Tv_k)$$

נניח ש־ $T \in \mathcal{L}(V,W)$ ו־ $v_1,\dots,v_n$ בסיס של V ו־ $w_1,\dots,w_m$ בסיס של W. יהי $1 \le k \le n$. אזי העמודה ה־ k^{th} של המטריצה $\mathcal{M}(T)$, המסומנת על ידי $\mathcal{M}(T)_{\cdot,k}$, שווה ל־ $\mathcal{M}(Tv_k)$.

הוכחה המסקנה הרצויה נובעת באופן מיידי מההגדרות של $\mathcal{M}(T)$ ו־ $\mathcal{M}(Tv_k)$.

המסקנה הבאה מראה כיצד המושגים של המטריצה של העתקה ליניארית, המטריצה של וקטור וכפל מטריצות משתלבים יחד.

3.76 העתקות ליניאריות פועלות כמו כפל מטריצות

נניח ש־ $T \in \mathcal{L}(V,W)$ ו־ $v \in V$. נניח ש־ $v_1,\dots,v_n$ בסיס של V ו־ $w_1,\dots,w_m$ בסיס של W. אזי

$$\mathcal{M}(Tv) = \mathcal{M}(T)\mathcal{M}(v).$$

הוכחה נניח ש־ $v = b_1v_1 + \cdots + b_nv_n$, כאשר $b_1,\dots,b_n \in \mathbf{F}$. לפיכך

$$Tv = b_1Tv_1 + \cdots + b_nTv_n. \tag{3.77}$$

ולכן

$$\begin{aligned}\mathcal{M}(Tv) &= b_1\mathcal{M}(Tv_1) + \cdots + b_n\mathcal{M}(Tv_n)\\ &= b_1\mathcal{M}(T)_{\cdot,1} + \cdots + b_n\mathcal{M}(T)_{\cdot,n}\\ &= \mathcal{M}(T)\mathcal{M}(v),\end{aligned}$$

כאשר את השוויון הראשון אנו מסיקים מ־ 3.77 ומהליניאריות של $\mathcal{M}$, השוויון השני נובע מ־ 3.75, והשוויון האחרון נובע מ־ 3.50.

כל A, מטריצה $m \times n$, משרה העתקה ליניארית מ־ $\mathbf{F}^{n,1}$ אל $\mathbf{F}^{m,1}$, כלומר, פונקציית כפל המטריצות המעתיקה $x \in \mathbf{F}^{n,1}$ אל $Ax \in \mathbf{F}^{m,1}$. נוכל להשתמש במסקנה לעיל כדי לחשוב מחדש על כל העתקה ליניארית (ממרחב וקטורי מממד סופי אל מרחב וקטורי מממד סופי) כהעתקת כפל מטריצות אחרי תיוג מחדש מתאים באמצעות האיזומורפיזמים הנתונים על ידי $\mathcal{M}$. בפרט, אם $T \in \mathcal{L}(V,W)$ ואנו נזהה את $v \in V$ עם $\mathcal{M}(v) \in \mathbf{F}^{n,1}$, אזי לפי המסקנה לעיל, נוכל לזהות את Tv עם $\mathcal{M}(T)\mathcal{M}(v)$.

מאחר שהמסקנה לעיל מאפשרת לנו (באמצעות איזומורפיזמים) לחשוב על כל העתקה ליניארית ככפל על $\mathbf{F}^{n,1}$ במטריצה A כלשהי, עלינו לזכור כי המטריצה המסוימת A תלויה לא רק בהגדרת ההעתקה כי אם גם בבחירת הבסיסים. נושא החוזר ברבות מהמסקנות החשובות ביותר בפרקים הבאים יהיה בחירת בסיס שיהפוך את המטריצה A לפשוטה ככל האפשר.

בספר זה אנו מתמקדים בעיקר בהעתקות ליניאריות, יותר מאשר במטריצות. אולם, לעיתים, תפיסת העתקות ליניאריות כמטריצות (ומטריצות כהעתקות ליניאריות) תעניק לנו תובנות חשובות ושימושיות.

Notice that no bases are in sight in the statement of the next result. Although $\mathcal{M}(T)$ in the next result depends on a choice of bases of V and W, the next result shows that the column rank of $\mathcal{M}(T)$ is the same for all such choices (because range T does not depend on a choice of basis).

3.78 *dimension of* range T *equals column rank of* $\mathcal{M}(T)$

Suppose V and W are finite-dimensional and $T \in \mathcal{L}(V, W)$. Then dim range T equals the column rank of $\mathcal{M}(T)$.

Proof Suppose $v_1, \dots, v_n$ is a basis of V and $w_1, \dots, w_m$ is a basis of W. The linear map that takes $w \in W$ to $\mathcal{M}(w)$ is an isomorphism from W onto the space $\mathbf{F}^{m,1}$ of m-by-1 column vectors. The restriction of this isomorphism to range T [which equals $\operatorname{span}(Tv_1, \dots, Tv_n)$ by Exercise 10 in Section 3B] is an isomorphism from range T onto $\operatorname{span}(\mathcal{M}(Tv_1), \dots, \mathcal{M}(Tv_n))$. For each $k \in \{1, \dots, n\}$, the m-by-1 matrix $\mathcal{M}(Tv_k)$ equals column k of $\mathcal{M}(T)$. Thus

$$\dim \text{range } T = \text{ the column rank of } \mathcal{M}(T),$$

as desired. ∎

Change of Basis

In Section 3C we defined the matrix

$$\mathcal{M}(T, (v_1, \dots, v_n), (w_1, \dots, w_m))$$

of a linear map T from V to a possibly different vector space W, where $v_1, \dots, v_n$ is a basis of V and $w_1, \dots, w_m$ is a basis of W. For linear maps from a vector space to itself, we usually use the same basis for both the domain vector space and the target vector space. When using a single basis in both capacities, we often write the basis only once. In other words, if $T \in \mathcal{L}(V)$ and $v_1, \dots, v_n$ is a basis of V, then the notation $\mathcal{M}(T, (v_1, \dots, v_n))$ is defined by the equation

$$\mathcal{M}(T, (v_1, \dots, v_n)) = \mathcal{M}(T, (v_1, \dots, v_n), (v_1, \dots, v_n)).$$

If the basis $v_1, \dots, v_n$ is clear from the context, then we can write just $\mathcal{M}(T)$.

3.79 definition: *identity matrix, I*

Suppose n is a positive integer. The n-by-n matrix

$$\begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}$$

with 1's on the diagonal (the entries where the row number equals the column number) and 0's elsewhere is called the *identity matrix* and is denoted by I.

שימו לב כי בניסוח המסקנה להלן בסיסים אינם מוזכרים כלל. אף ש־ $\mathcal{M}(T)$ במסקנה הבאה תלויה בבחירת הבסיסים של V ו־ W, מסקנה זו מראה כי דרגת העמודות של $\mathcal{M}(T)$ היא זהה בכל בחירה של בסיסים (שכן range T אינו תלוי בבחירת בסיסים).

3.78 ממד range T שווה לדרגת העמודות של $\mathcal{M}(T)$

נניח ש־ V ו־ W מממד סופי ו־ $T \in \mathcal{L}(V,W)$. אזי $\dim \operatorname{range} T$ שווה לדרגת העמודות של $\mathcal{M}(T)$.

הוכחה נניח ש־ $v_1, \dots, v_n$ בסיס של V ו־ $w_1, \dots, w_m$ בסיס של W. ההעתקה הליניארית המעתיקה את $w \in W$ אל $\mathcal{M}(w)$ היא איזומורפיזם מ־ W על המרחב $\mathbf{F}^{m,1}$ של וקטורי עמודה $m \times 1$. צמצום איזומורפיזם זה ל־ range T [השווה ל־ $\operatorname{span}(Tv_1, \dots, Tv_n)$ לפי תרגיל 10 ביחידה ג2] הוא איזומורפיזם מ־ range T על $\operatorname{span}(\mathcal{M}(Tv_1), \dots, \mathcal{M}(Tv_n))$. לכל $k \in \{1, \dots, n\}$, המטריצה $m \times 1$ $\mathcal{M}(Tv_k)$ שווה לעמודה k של $\mathcal{M}(T)$. לפיכך

$$\dim \operatorname{range} T = \text{דרגת העמודות של } \mathcal{M}(T),$$

כנדרש. ■

שינוי בסיס

ביחידה ג3 הגדרנו את המטריצה

$$\mathcal{M}(T, (v_1, \dots, v_n), (w_1, \dots, w_m))$$

של ההעתקה הליניארית T מ־ V אל מרחב וקטורי זהה או אחר W, כאשר $v_1, \dots, v_n$ בסיס של V ו־ $w_1, \dots, w_m$ בסיס של W. במקרה של העתקה ליניארית ממרחב וקטורי לעצמו, בדרך כלל נשתמש באותו הבסיס עבור המרחב הווקטורי של התחום ועבור המרחב הווקטורי של היעד. כשאנו משתמשים בבסיס יחיד בשני התפקידים, לרוב נציין את הבסיס פעם אחת בלבד. כלומר, אם $T \in \mathcal{L}(V)$ ו־ $v_1, \dots, v_n$ בסיס של V, אזי הסימון $\mathcal{M}(T, (v_1, \dots, v_n))$ מוגדר על ידי המשוואה

$$\mathcal{M}(T, (v_1, \dots, v_n)) = \mathcal{M}(T, (v_1, \dots, v_n), (v_1, \dots, v_n)).$$

אם הבסיס $v_1, \dots, v_n$ ברור מההקשר, נוכל לרשום רק $\mathcal{M}(T)$.

3.79 הגדרה: מטריצת הזהות (היחידה), I

נניח ש־ n מספר שלם חיובי. המטריצה $n \times n$

$$\begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}$$

עם 1־ים באלכסון (האיברים שעבורם מספר השורה ומספר העמודה שווים) ואפסים בכל האיברים האחרים נקראת מטריצת הזהות ומסומנת ב־ I.

In the definition above, the 0 in the lower left corner of the matrix indicates that all entries below the diagonal are 0, and the 0 in the upper right corner indicates that all entries above the diagonal are 0.

With respect to each basis of V, the matrix of the identity operator $I \in \mathcal{L}(V)$ is the identity matrix I. Note that the symbol I is used to denote both the identity operator and the identity matrix. The context indicates which meaning of I is intended. For example, consider the equation $\mathcal{M}(I) = I$; on the left side I denotes the identity operator, and on the right side I denotes the identity matrix.

If A is a square matrix (meaning it has the same number of rows as columns) with the same size as I, then $AI = IA = A$, as you should verify.

3.80 definition: *invertible, inverse, A^{-1}*

A square matrix A is called *invertible* if there is a square matrix B of the same size such that $AB = BA = I$; we call B the *inverse* of A and denote it by A^{-1}.

The same proof as used in 3.60 shows that if A is an invertible square matrix, then there is a unique matrix B such that $AB = BA = I$ (and thus the notation $B = A^{-1}$ is justified).

*Some mathematicians use the terms **nonsingular** and **singular**, which mean the same as invertible and non-invertible.*

If A is an invertible matrix, then $(A^{-1})^{-1} = A$ because

$$A^{-1}A = AA^{-1} = I.$$

Also, if A and C are invertible square matrices of the same size, then AC is invertible and $(AC)^{-1} = C^{-1}A^{-1}$ because

$$\begin{aligned}(AC)(C^{-1}A^{-1}) &= A(CC^{-1})A^{-1}\\ &= AIA^{-1}\\ &= AA^{-1}\\ &= I,\end{aligned}$$

and similarly $(C^{-1}A^{-1})(AC) = I$.

The next result holds because we defined matrix multiplication to make it true—see 3.43 and the material preceding it. Now we are just being more explicit about the bases involved.

3.81 *matrix of product of linear maps*

Suppose $T \in \mathcal{L}(U, V)$ and $S \in \mathcal{L}(V, W)$. If $u_1, \dots, u_m$ is a basis of U, $v_1, \dots, v_n$ is a basis of V, and $w_1, \dots, w_p$ is a basis of W, then

$$\mathcal{M}(ST, (u_1, \dots, u_m), (w_1, \dots, w_p)) = \mathcal{M}(S, (v_1, \dots, v_n), (w_1, \dots, w_p))\mathcal{M}(T, (u_1, \dots, u_m), (v_1, \dots, v_n)).$$

בהגדרה לעיל, ה־ 0 בפינה השמאלית התחתונה מצביע על כך שכל האיברים תחת האלכסון הם 0, וה־ 0 בפינה הימנית העליונה מצביע על כך שכל האיברים מעל האלכסון הם 0.

ביחס לכל בסיס של V, המטריצה של אופרטור הזהות $I \in \mathcal{L}(V)$ היא מטריצת הזהות I. שימו לב שהסימון I משמש לסימון הן של אופרטור הזהות והן של מטריצת הזהות. מן ההקשר ברור לאיזה I אנו מתכוונים. לדוגמה, ניקח את המשוואה הבאה $\mathcal{M}(I) = I$; באגף השמאלי I מסמן את אופרטור הזהות, ובאגף הימני I מסמן את מטריצת הזהות.

אם A מטריצה ריבועית (כלומר מספר שורותיה ועמודותיה שווים) ובאותו הגודל כשל I, אזי $AI = IA = A$, כפי שעליכם לוודא.

3.80 הגדרה: הפיכה, הופכית, A^{-1}

מטריצה ריבועית A נקראת הפיכה אם קיימת מטריצה ריבועית B באותו הגודל כך שמתקיים $AB = BA = I$; אנו קוראים ל־ B ההופכית של A ומסמנים אותה ב־ A^{-1}.

ישנם מתמטיקאים המשתמשים במונחים **אי־סינגולרית** ו־ **סינגולרית**, שמשמעותם זהה להפיכה ולבלתי הפיכה.

אותה הוכחה שבה השתמשנו ב־ 3.60 מראה שאם A מטריצה ריבועית הפיכה, אזי קיימת מטריצה יחידה B כך שמתקיים $AB = BA = I$ (ולפיכך הסימון $B = A^{-1}$ מוצדק).

אם A מטריצה הפיכה, אזי $(A^{-1})^{-1} = A$ מאחר ש־

$$A^{-1}A = AA^{-1} = I.$$

בנוסף, אם A ו־ C מטריצות ריבועיות הפיכות באותו הגודל, אזי AC הפיכה וגם $(AC)^{-1} = C^{-1}A^{-1}$ מאחר ש־

$$\begin{aligned}(AC)(C^{-1}A^{-1}) &= A(CC^{-1})A^{-1}\\ &= AIA^{-1}\\ &= AA^{-1}\\ &= I,\end{aligned}$$

כמו כן $(C^{-1}A^{-1})(AC) = I$.

המסקנה הבאה מתקיימת מאחר שכפל מטריצות תוכנן והוגדר כך שנגיע לתוצאה זו—ראו 3.43 והחומר שקדם לו. כעת אנו פשוט מתייחסים במפורש לבסיסים הקשורים להעתקות.

3.81 מטריצת מכפלת העתקות ליניאריות

נניח ש־ $T \in \mathcal{L}(U, V)$ ו־ $S \in \mathcal{L}(V, W)$. אם $u_1, \dots, u_m$ בסיס של U, ו־ $v_1, \dots, v_n$ בסיס של V, ו־ $w_1, \dots, w_p$ בסיס של W, אזי

$$\mathcal{M}(ST, (u_1, \dots, u_m), (w_1, \dots, w_p)) = \mathcal{M}(S, (v_1, \dots, v_n), (w_1, \dots, w_p))\mathcal{M}(T, (u_1, \dots, u_m), (v_1, \dots, v_n)).$$

The next result deals with the matrix of the identity operator I with respect to two different bases. Note that the kth column of $\mathcal{M}(I, (u_1, \dots, u_n), (v_1, \dots, v_n))$ consists of the scalars needed to write u_k as a linear combination of the basis $v_1, \dots, v_n$.

In the statement of the next result, I denotes the identity operator from V to V. In the proof, I also denotes the n-by-n identity matrix.

3.82 *matrix of identity operator with respect to two bases*

Suppose that $u_1, \dots, u_n$ and $v_1, \dots, v_n$ are bases of V. Then the matrices

$$\mathcal{M}(I, (u_1, \dots, u_n), (v_1, \dots, v_n)) \quad \text{and} \quad \mathcal{M}(I, (v_1, \dots, v_n), (u_1, \dots, u_n))$$

are invertible, and each is the inverse of the other.

Proof In 3.81, replace w_k with u_k, and replace S and T with I, getting

$$I = \mathcal{M}(I, (v_1, \dots, v_n), (u_1, \dots, u_n))\mathcal{M}(I, (u_1, \dots, u_n), (v_1, \dots, v_n)).$$

Now interchange the roles of the u's and v's, getting

$$I = \mathcal{M}(I, (u_1, \dots, u_n), (v_1, \dots, v_n))\mathcal{M}(I, (v_1, \dots, v_n), (u_1, \dots, u_n)).$$

These two equations above give the desired result. ∎

3.83 example: *matrix of identity operator on* $\mathbf{F}^2$ *with respect to two bases*

Consider the bases $(4, 2), (5, 3)$ and $(1, 0), (0, 1)$ of $\mathbf{F}^2$. Because $I(4, 2) = 4(1, 0) + 2(0, 1)$ and $I(5, 3) = 5(1, 0) + 3(0, 1)$, we have

$$\mathcal{M}\big(I, ((4, 2), (5, 3)), ((1, 0), (0, 1))\big) = \begin{pmatrix} 4 & 5 \\ 2 & 3 \end{pmatrix}.$$

The inverse of the matrix above is

$$\begin{pmatrix} \frac{3}{2} & -\frac{5}{2} \\ -1 & 2 \end{pmatrix},$$

as you should verify. Thus 3.82 implies that

$$\mathcal{M}\big(I, ((1, 0), (0, 1)), ((4, 2), (5, 3))\big) = \begin{pmatrix} \frac{3}{2} & -\frac{5}{2} \\ -1 & 2 \end{pmatrix}.$$

Our next result shows how the matrix of T changes when we change bases. In the next result, we have two different bases of V, each of which is used as a basis for the domain space and as a basis for the target space. Recall our shorthand notation that allows us to display a basis only once when it is used in both capacities:

$$\mathcal{M}(T, (u_1, \dots, u_n)) = \mathcal{M}(T, (u_1, \dots, u_n), (u_1, \dots, u_n)).$$

המסקנה הבאה עוסקת במטריצה של אופרטור הזהות I ביחס לשני בסיסים שונים. שימו לב כי העמודה ה־ k^{th} של $\mathcal{M}(I,(u_1,\dots,u_n),(v_1,\dots,v_n))$ מורכבת מהסקלרים הנדרשים לכתיבת u_k כצירוף ליניארי של הבסיס $v_1,\dots,v_n$.

בניסוח המסקנה הבאה, I מסמן את אופרטור הזהות מ־ V אל V. בהוכחה, I מסמן גם את מטריצת הזהות $n \times n$.

3.82 המטריצה של אופרטור הזהות ביחס לשני בסיסים

נניח ש־ $u_1,\dots,u_n$ ו־ $v_1,\dots,v_n$ בסיסים של V. אזי המטריצות
$\mathcal{M}(I,(v_1,\dots,v_n),(u_1,\dots,u_n))$ וגם $\mathcal{M}(I,(u_1,\dots,u_n),(v_1,\dots,v_n))$
הפיכות, וכל אחת היא ההופכית של האחרת.

הוכחה ב־ 3.81, נחליף w_k ב־ u_k, ונחליף S ו־ T ב־ I, ונקבל
$$I = \mathcal{M}(I,(v_1,\dots,v_n),(u_1,\dots,u_n))\mathcal{M}(I,(u_1,\dots,u_n),(v_1,\dots,v_n)).$$
כעת נחליף בין התפקידים של ה־ u־ים ושל ה־ v־ים, ונקבל
$$I = \mathcal{M}(I,(u_1,\dots,u_n),(v_1,\dots,v_n))\mathcal{M}(I,(v_1,\dots,v_n),(u_1,\dots,u_n)).$$
שתי משוואות אלו מספקות את התוצאה הנדרשת. ■

3.83 דוגמה: מטריצה של אופרטור הזהות מעל $\mathbf{F}^2$ ביחס לשני בסיסים

ניקח את הבסיסים $(4,2),(5,3)$ ו־ $(1,0),(0,1)$ של $\mathbf{F}^2$. מאחר ש־ $I(4,2) = 4(1,0) + 2(0,1)$ ו־ $I(5,3) = 5(1,0) + 3(0,1)$, נקבל
$$\mathcal{M}\big(I,((4,2),(5,3)),((1,0),(0,1))\big) = \begin{pmatrix} 4 & 5 \\ 2 & 3 \end{pmatrix}.$$

ההופכית של המטריצה לעיל היא
$$\begin{pmatrix} \frac{3}{2} & -\frac{5}{2} \\ -1 & 2 \end{pmatrix},$$
כפי שעליכם לוודא. לפיכך, מ־ 3.82 נסיק ש־
$$\mathcal{M}\big(I,((1,0),(0,1)),((4,2),(5,3))\big) = \begin{pmatrix} \frac{3}{2} & -\frac{5}{2} \\ -1 & 2 \end{pmatrix}.$$

המסקנה הבאה מראה כיצד המטריצה של T משתנה כאשר אנו משנים בסיסים. במסקנה הבאה, יש לנו שני בסיסים של V, כל אחד מהם משמש כבסיס עבור מרחב התחום וכבסיס עבור מרחב היעד. זכרו את הסימון המקוצר שמאפשר לנו להציג את הבסיס פעם אחת בלבד כאשר הבסיס משמש לשני המרחבים:
$$\mathcal{M}(T,(u_1,\dots,u_n)) = \mathcal{M}(T,(u_1,\dots,u_n),(u_1,\dots,u_n)).$$

3.84 *change-of-basis formula*

Suppose $T \in \mathcal{L}(V)$. Suppose $u_1, \dots, u_n$ and $v_1, \dots, v_n$ are bases of V. Let

$$A = \mathcal{M}(T, (u_1, \dots, u_n)) \quad \text{and} \quad B = \mathcal{M}(T, (v_1, \dots, v_n))$$

and $C = \mathcal{M}(I, (u_1, \dots, u_n), (v_1, \dots, v_n))$. Then

$$A = C^{-1}BC.$$

Proof In 3.81, replace w_k with u_k and replace S with I, getting

3.85 $$A = C^{-1}\mathcal{M}(T, (u_1, \dots, u_n), (v_1, \dots, v_n)),$$

where we have used 3.82.

Again use 3.81, this time replacing w_k with v_k. Also replace T with I and replace S with T, getting

$$\mathcal{M}(T, (u_1, \dots, u_n), (v_1, \dots, v_n)) = BC.$$

Substituting the equation above into 3.85 gives the equation $A = C^{-1}BC$. ∎

The proof of the next result is left as an exercise.

3.86 *matrix of inverse equals inverse of matrix*

Suppose that $v_1, \dots, v_n$ is a basis of V and $T \in \mathcal{L}(V)$ is invertible. Then $\mathcal{M}(T^{-1}) = (\mathcal{M}(T))^{-1}$, where both matrices are with respect to the basis $v_1, \dots, v_n$.

Exercises 3D

1 Suppose $T \in \mathcal{L}(V, W)$ is invertible. Show that T^{-1} is invertible and

$$(T^{-1})^{-1} = T.$$

2 Suppose $T \in \mathcal{L}(U, V)$ and $S \in \mathcal{L}(V, W)$ are both invertible linear maps. Prove that $ST \in \mathcal{L}(U, W)$ is invertible and that $(ST)^{-1} = T^{-1}S^{-1}$.

3 Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Prove that the following are equivalent.

(a) T is invertible.

(b) $Tv_1, \dots, Tv_n$ is a basis of V for every basis $v_1, \dots, v_n$ of V.

(c) $Tv_1, \dots, Tv_n$ is a basis of V for some basis $v_1, \dots, v_n$ of V.

4 Suppose V is finite-dimensional and $\dim V > 1$. Prove that the set of noninvertible linear maps from V to itself is not a subspace of $\mathcal{L}(V)$.

3.84 נוסחת שינוי בסיס

נניח ש־ $T \in \mathcal{L}(V)$. נניח ש־ $u_1, \dots, u_n$ ו־ $v_1, \dots, v_n$ בסיסים של V. תהיינה
$$B = \mathcal{M}(T, (v_1, \dots, v_n)) \quad \text{וגם} \quad A = \mathcal{M}(T, (u_1, \dots, u_n))$$
ו־ $C = \mathcal{M}(I, (u_1, \dots, u_n), (v_1, \dots, v_n))$. אזי
$$A = C^{-1}BC.$$

הוכחה ב־ 3.81, נחליף w_k ב־ u_k ונחליף S ב־ I, ונקבל

$$A = C^{-1}\mathcal{M}(T, (u_1, \dots, u_n), (v_1, \dots, v_n)), \tag{3.85}$$

כאשר השתמשנו ב־ 3.82.

שוב נשתמש ב־ 3.81, הפעם נחליף w_k ב־ v_k. בנוסף נחליף T ב־ I ונחליף S ב־ T, ונקבל

$$\mathcal{M}(T, (u_1, \dots, u_n), (v_1, \dots, v_n)) = BC.$$

נציב את המשוואה לעיל ב־ 3.85 ונקבל את המשוואה $A = C^{-1}BC$. ■

ההוכחה למסקנה הבאה מושארת לקוראים כתרגיל.

3.86 המטריצה של ההופכית שווה להופכית של המטריצה

נניח ש־ $v_1, \dots, v_n$ בסיס של V ו־ $T \in \mathcal{L}(V)$ הפיכה. אזי $\mathcal{M}(T^{-1}) = (\mathcal{M}(T))^{-1}$, כאשר שתי המטריצות מוגדרות ביחס לבסיס $v_1, \dots, v_n$.

תרגילים 3ג

1 נניח ש־ $T \in \mathcal{L}(V, W)$ הפיכה. הראו ש־ T^{-1} הפיכה וגם
$$(T^{-1})^{-1} = T.$$

2 נניח ש־ $T \in \mathcal{L}(U, V)$ ו־ $S \in \mathcal{L}(V, W)$ שתיהן העתקות ליניאריות הפיכות. הוכיחו ש־ $ST \in \mathcal{L}(U, W)$ הפיכה וש־ $(ST)^{-1} = T^{-1}S^{-1}$.

3 נניח ש־ V מממד סופי ו־ $T \in \mathcal{L}(V)$. הוכיחו שהתנאים הבאים שקולים.

א. T הפיכה.

ב. $Tv_1, \dots, Tv_n$ בסיס של V לכל בסיס $v_1, \dots, v_n$ של V.

ג. $Tv_1, \dots, Tv_n$ בסיס של V עבור בסיס $v_1, \dots, v_n$ כלשהו של V.

4 נניח ש־ V מממד סופי ו־ $\dim V > 1$. הוכיחו שקבוצת ההעתקות הליניאריות הבלתי־הפיכות מ־ V לעצמו אינה תת־מרחב של $\mathcal{L}(V)$.

5 Suppose V is finite-dimensional, U is a subspace of V, and $S \in \mathcal{L}(U, V)$. Prove that there exists an invertible linear map T from V to itself such that $Tu = Su$ for every $u \in U$ if and only if S is injective.

6 Suppose that W is finite-dimensional and $S, T \in \mathcal{L}(V, W)$. Prove that $\operatorname{null} S = \operatorname{null} T$ if and only if there exists an invertible $E \in \mathcal{L}(W)$ such that $S = ET$.

7 Suppose that V is finite-dimensional and $S, T \in \mathcal{L}(V, W)$. Prove that $\operatorname{range} S = \operatorname{range} T$ if and only if there exists an invertible $E \in \mathcal{L}(V)$ such that $S = TE$.

8 Suppose V and W are finite-dimensional and $S, T \in \mathcal{L}(V, W)$. Prove that there exist invertible $E_1 \in \mathcal{L}(V)$ and $E_2 \in \mathcal{L}(W)$ such that $S = E_2 T E_1$ if and only if $\dim \operatorname{null} S = \dim \operatorname{null} T$.

9 Suppose V is finite-dimensional and $T \colon V \to W$ is a surjective linear map of V onto W. Prove that there is a subspace U of V such that $T|_U$ is an isomorphism of U onto W.

Here $T|_U$ means the function T restricted to U. Thus $T|_U$ is the function whose domain is U, with $T|_U$ defined by $T|_U(u) = Tu$ for every $u \in U$.

10 Suppose V and W are finite-dimensional and U is a subspace of V. Let

$$\mathcal{E} = \{T \in \mathcal{L}(V, W) : U \subseteq \operatorname{null} T\}.$$

(a) Show that $\mathcal{E}$ is a subspace of $\mathcal{L}(V, W)$.
(b) Find a formula for $\dim \mathcal{E}$ in terms of $\dim V$, $\dim W$, and $\dim U$.

Hint: Define $\Phi \colon \mathcal{L}(V, W) \to \mathcal{L}(U, W)$ by $\Phi(T) = T|_U$. What is $\operatorname{null} \Phi$? What is $\operatorname{range} \Phi$?

11 Suppose V is finite-dimensional and $S, T \in \mathcal{L}(V)$. Prove that

$$ST \text{ is invertible} \iff S \text{ and } T \text{ are invertible.}$$

12 Suppose V is finite-dimensional and $S, T, U \in \mathcal{L}(V)$ and $STU = I$. Show that T is invertible and that $T^{-1} = US$.

13 Show that the result in Exercise 12 can fail without the hypothesis that V is finite-dimensional.

14 Prove or give a counterexample: If V is a finite-dimensional vector space and $R, S, T \in \mathcal{L}(V)$ are such that RST is surjective, then S is injective.

15 Suppose $T \in \mathcal{L}(V)$ and $v_1, \dots, v_m$ is a list in V such that $Tv_1, \dots, Tv_m$ spans V. Prove that $v_1, \dots, v_m$ spans V.

16 Prove that every linear map from $\mathbf{F}^{n,1}$ to $\mathbf{F}^{m,1}$ is given by a matrix multiplication. In other words, prove that if $T \in \mathcal{L}(\mathbf{F}^{n,1}, \mathbf{F}^{m,1})$, then there exists an m-by-n matrix A such that $Tx = Ax$ for every $x \in \mathbf{F}^{n,1}$.

5 נניח ש־ V מממד סופי, U תת־מרחב של V, ו־ $S \in \mathcal{L}(U,V)$. הוכיחו שקיימת העתקה ליניארית הפיכה T מ־ V לעצמו כך ש־ $Tu = Su$ לכל $u \in U$ אם־ורק־אם S חד־חד־ערכית.

6 נניח ש־ W מממד סופי ו־ $S, T \in \mathcal{L}(V,W)$. הוכיחו ש־ $\operatorname{null} S = \operatorname{null} T$ אם־ורק־אם קיימת $E \in \mathcal{L}(W)$ הפיכה, כך ש־ $S = ET$.

7 נניח ש־ V מממד סופי ו־ $S, T \in \mathcal{L}(V,W)$. הוכיחו ש־ $\operatorname{range} S = \operatorname{range} T$ אם־ורק־אם קיימת $E \in \mathcal{L}(V)$ הפיכה כך ש־ $S = TE$.

8 נניח ש־ V ו־ W מממד סופי ו־ $S, T \in \mathcal{L}(V,W)$. הוכיחו שקיימות העתקות הפיכות $E_1 \in \mathcal{L}(V)$ ו־ $E_2 \in \mathcal{L}(W)$ כך ש־ $S = E_2 T E_1$ אם־ורק־אם $\dim \operatorname{null} S = \dim \operatorname{null} T$.

9 נניח ש־ V מממד סופי ו־ $T : V \to W$ העתקה ליניארית מ־ V על W. הוכיחו שקיים תת־מרחב U של V כך ש־ $T|_U$ היא איזומורפיזם של U על W.

כאן $T|_U$ היא הפונקציה T מצומצמת ל־ U. לפיכך $T|_U$ היא הפונקציה שתחומה U, כאשר $T|_U$ מוגדרת על ידי $T|_U(u) = Tu$ לכל $u \in U$.

10 נניח ש־ V ו־ W מממד סופי ו־ U תת־מרחב של V. יהי

$$\mathcal{E} = \{T \in \mathcal{L}(V,W) : U \subseteq \operatorname{null} T\}.$$

א. הראו ש־ $\mathcal{E}$ תת־מרחב של $\mathcal{L}(V,W)$.

ב. מצאו נוסחה עבור $\dim \mathcal{E}$ במונחים של $\dim V$, $\dim W$, ו־ $\dim U$.

רמז: הגדירו $\Phi : \mathcal{L}(V,W) \to \mathcal{L}(U,W)$ על ידי $\Phi(T) = T|_U$. מהו $\operatorname{null} \Phi$? מהו $\operatorname{range} \Phi$?

11 נניח ש־ V מממד סופי ו־ $S, T \in \mathcal{L}(V)$. הוכיחו ש־

$$ST \text{ הפיכה} \iff S \text{ ו־ } T \text{ הפיכות.}$$

12 נניח ש־ V מממד סופי ו־ $S, T, U \in \mathcal{L}(V)$ ו־ $STU = I$. הראו ש־ T הפיכה וגם ש־ $T^{-1} = US$.

13 הראו שהמסקנה בתרגיל 12 אינה מתקיימת בהכרח ללא ההנחה ש־ V מממד סופי.

14 הוכיחו או הביאו דוגמה נגדית: אם V מרחב וקטורי מממד סופי ו־ $R, S, T \in \mathcal{L}(V)$ כך ש־ RST על, אזי S חד־חד־ערכית.

15 נניח ש־ $T \in \mathcal{L}(V)$ ו־ $v_1, \dots, v_m$ רשימה ב־ V כך ש־ $Tv_1, \dots, Tv_m$ פורסת את V. הוכיחו ש־ $v_1, \dots, v_m$ פורסת את V.

16 הוכיחו שכל העתקה ליניארית מ־ $\mathbf{F}^{n,1}$ אל $\mathbf{F}^{m,1}$ נתונה על ידי כפל מטריצות. כלומר, הוכיחו שאם $T \in \mathcal{L}(\mathbf{F}^{n,1}, \mathbf{F}^{m,1})$, אזי קיימת A מטריצה $m \times n$ כך ש־ $Tx = Ax$ לכל $x \in \mathbf{F}^{n,1}$.

17 Suppose V is finite-dimensional and $S \in \mathcal{L}(V)$. Define $\mathcal{A} \in \mathcal{L}(\mathcal{L}(V))$ by

$$\mathcal{A}(T) = ST$$

for $T \in \mathcal{L}(V)$.

(a) Show that $\dim \operatorname{null} \mathcal{A} = (\dim V)(\dim \operatorname{null} S)$.

(b) Show that $\dim \operatorname{range} \mathcal{A} = (\dim V)(\dim \operatorname{range} S)$.

18 Show that V and $\mathcal{L}(\mathbf{F}, V)$ are isomorphic vector spaces.

19 Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Prove that T has the same matrix with respect to every basis of V if and only if T is a scalar multiple of the identity operator.

20 Suppose $q \in \mathcal{P}(\mathbf{R})$. Prove that there exists a polynomial $p \in \mathcal{P}(\mathbf{R})$ such that

$$q(x) = (x^2 + x)p''(x) + 2xp'(x) + p(3)$$

for all $x \in \mathbf{R}$.

21 Suppose n is a positive integer and $A_{j,k} \in \mathbf{F}$ for all $j, k = 1, \dots, n$. Prove that the following are equivalent (note that in both parts below, the number of equations equals the number of variables).

(a) The trivial solution $x_1 = \cdots = x_n = 0$ is the only solution to the homogeneous system of equations

$$\begin{aligned} \sum_{k=1}^{n} A_{1,k} x_k &= 0 \\ &\vdots \\ \sum_{k=1}^{n} A_{n,k} x_k &= 0. \end{aligned}$$

(b) For every $c_1, \dots, c_n \in \mathbf{F}$, there exists a solution to the system of equations

$$\begin{aligned} \sum_{k=1}^{n} A_{1,k} x_k &= c_1 \\ &\vdots \\ \sum_{k=1}^{n} A_{n,k} x_k &= c_n. \end{aligned}$$

22 Suppose $T \in \mathcal{L}(V)$ and $v_1, \dots, v_n$ is a basis of V. Prove that

$$\mathcal{M}(T, (v_1, \dots, v_n)) \text{ is invertible} \iff T \text{ is invertible.}$$

23 Suppose that $u_1, \dots, u_n$ and $v_1, \dots, v_n$ are bases of V. Let $T \in \mathcal{L}(V)$ be such that $Tv_k = u_k$ for each $k = 1, \dots, n$. Prove that

$$\mathcal{M}(T, (v_1, \dots, v_n)) = \mathcal{M}(I, (u_1, \dots, u_n), (v_1, \dots, v_n)).$$

24 Suppose A and B are square matrices of the same size and $AB = I$. Prove that $BA = I$.

17 נניח ש־ V מממד סופי ו־ $S \in \mathcal{L}(V)$. הגדירו $\mathcal{A} \in \mathcal{L}(\mathcal{L}(V))$ על ידי

$$\mathcal{A}(T) = ST$$

לכל $T \in \mathcal{L}(V)$.

א. הראו ש־ $\dim \operatorname{null} \mathcal{A} = (\dim V)(\dim \operatorname{null} S)$.

ב. הראו ש־ $\dim \operatorname{range} \mathcal{A} = (\dim V)(\dim \operatorname{range} S)$.

18 הראו ש־ V ו־ $\mathcal{L}(\mathbf{F}, V)$ מרחבים וקטוריים איזומורפיים.

19 נניח ש־ V מממד סופי ו־ $T \in \mathcal{L}(V)$. הוכיחו שיש ל־ T אותה מטריצה ביחס לכל בסיס של V אם־ורק־אם T כפולת סקלר של אופרטור הזהות.

20 נניח ש־ $q \in \mathcal{P}(\mathbf{R})$. הוכיחו שקיים פולינום $p \in \mathcal{P}(\mathbf{R})$ כך ש־

$$q(x) = (x^2 + x)p''(x) + 2xp'(x) + p(3)$$

לכל $x \in \mathbf{R}$.

21 נניח ש־ n מספר שלם חיובי ו־ $A_{j,k} \in \mathbf{F}$ לכל $j, k = 1, \dots, n$. הוכיחו שהתנאים הבאים שקולים (שימו לב כי בשני החלקים להלן, מספר המשוואות שווה למספר המשתנים).

א. הפתרון הטריוויאלי $x_1 = \cdots = x_n = 0$ הוא הפתרון היחיד למערכת ההומוגנית של המשוואות הליניאריות

$$\begin{aligned} \sum_{k=1}^{n} A_{1,k} x_k &= 0 \\ &\vdots \\ \sum_{k=1}^{n} A_{n,k} x_k &= 0. \end{aligned}$$

ב. לכל $c_1, \dots, c_n \in \mathbf{F}$, קיים פתרון למערכת המשוואות

$$\begin{aligned} \sum_{k=1}^{n} A_{1,k} x_k &= c_1 \\ &\vdots \\ \sum_{k=1}^{n} A_{n,k} x_k &= c_n. \end{aligned}$$

22 נניח ש־ $T \in \mathcal{L}(V)$ ו־ $v_1, \dots, v_n$ בסיס של V. הוכיחו ש־

$$\mathcal{M}(T, (v_1, \dots, v_n)) \text{ הפיכה} \iff T \text{ הפיכה}.$$

23 נניח ש־ $u_1, \dots, u_n$ ו־ $v_1, \dots, v_n$ בסיסים של V. תהי $T \in \mathcal{L}(V)$ העתקה כך ש־ $Tv_k = u_k$ לכל $k = 1, \dots, n$. הוכיחו ש־

$$\mathcal{M}(T, (v_1, \dots, v_n)) = \mathcal{M}(I, (u_1, \dots, u_n), (v_1, \dots, v_n)).$$

24 נניח ש־ A ו־ B מטריצות ריבועיות בגודל זהה ו־ $AB = I$. הוכיחו ש־ $BA = I$.

אלגברה ליניארית בדרך הישר, מהדורה רביעית, מחבר: שלדון אקסלר
Linear Algebra Done Right, fourth edition, by Sheldon Axler

3E Products and Quotients of Vector Spaces

Products of Vector Spaces

As usual when dealing with more than one vector space, all vector spaces in use should be over the same field.

3.87 definition: *product of vector spaces*

Suppose $V_1, \dots, V_m$ are vector spaces over $\mathbf{F}$.

- The *product* $V_1 \times \cdots \times V_m$ is defined by

$$V_1 \times \cdots \times V_m = \{(v_1, \dots, v_m) : v_1 \in V_1, \dots, v_m \in V_m\}.$$

- Addition on $V_1 \times \cdots \times V_m$ is defined by

$$(u_1, \dots, u_m) + (v_1, \dots, v_m) = (u_1 + v_1, \dots, u_m + v_m).$$

- Scalar multiplication on $V_1 \times \cdots \times V_m$ is defined by

$$\lambda(v_1, \dots, v_m) = (\lambda v_1, \dots, \lambda v_m).$$

3.88 example: *product of the vector spaces $\mathcal{P}_5(\mathbf{R})$ and $\mathbf{R}^3$*

Elements of $\mathcal{P}_5(\mathbf{R}) \times \mathbf{R}^3$ are lists of length two, with the first item in the list an element of $\mathcal{P}_5(\mathbf{R})$ and the second item in the list an element of $\mathbf{R}^3$.

For example, $(5 - 6x + 4x^2, (3, 8, 7))$ and $(x + 9x^5, (2, 2, 2))$ are elements of $\mathcal{P}_5(\mathbf{R}) \times \mathbf{R}^3$. Their sum is defined by

$$\begin{aligned}(5 - 6x + 4x^2, (3, 8, 7)) + (x + 9x^5, (2, 2, 2)) \\ = (5 - 5x + 4x^2 + 9x^5, (5, 10, 9)).\end{aligned}$$

Also, $2(5 - 6x + 4x^2, (3, 8, 7)) = (10 - 12x + 8x^2, (6, 16, 14))$.

The next result should be interpreted to mean that the product of vector spaces is a vector space with the operations of addition and scalar multiplication as defined by 3.87.

3.89 *product of vector spaces is a vector space*

Suppose $V_1, \dots, V_m$ are vector spaces over $\mathbf{F}$. Then $V_1 \times \cdots \times V_m$ is a vector space over $\mathbf{F}$.

The proof of the result above is left to the reader. Note that the additive identity of $V_1 \times \cdots \times V_m$ is $(0, \dots, 0)$, where the 0 in the kth slot is the additive identity of V_k. The additive inverse of $(v_1, \dots, v_m) \in V_1 \times \cdots \times V_m$ is $(-v_1, \dots, -v_m)$.

ג5 מכפלות ומנות של מרחבים וקטוריים

מכפלות של מרחבים וקטוריים

כהרגלנו, כאשר עוסקים ביותר ממרחב וקטורי אחד, כל המרחבים הווקטוריים שבהם נעסוק צריכים להיות מעל אותו שדה.

3.87 הגדרה: מכפלת מרחבים וקטוריים

נניח ש־ $V_1, \dots, V_m$ מרחבים וקטוריים מעל $\mathbf{F}$.

- המכפלה $V_1 \times \cdots \times V_m$ מוגדרת על ידי
$$V_1 \times \cdots \times V_m = \{(v_1, \dots, v_m) : v_1 \in V_1, \dots, v_m \in V_m\}.$$
- פעולת החיבור מעל $V_1 \times \cdots \times V_m$ מוגדרת על ידי
$$(u_1, \dots, u_m) + (v_1, \dots, v_m) = (u_1 + v_1, \dots, u_m + v_m).$$
- כפל בסקלר מעל $V_1 \times \cdots \times V_m$ מוגדר על ידי
$$\lambda(v_1, \dots, v_m) = (\lambda v_1, \dots, \lambda v_m).$$

3.88 דוגמה: מכפלת המרחבים הווקטוריים $\mathcal{P}_5(\mathbf{R})$ ו־ $\mathbf{R}^3$

איברי $\mathcal{P}_5(\mathbf{R}) \times \mathbf{R}^3$ הם רשימות בנות שני איברים, כאשר האיבר הראשון ברשימה הוא איבר ב־ $\mathcal{P}_5(\mathbf{R})$ והאיבר השני ברשימה הוא איבר ב־ $\mathbf{R}^3$.

לדוגמה, $(5-6x+4x^2, (3,8,7))$ ו־ $(x+9x^5, (2,2,2))$ הם איברים של $\mathcal{P}_5(\mathbf{R}) \times \mathbf{R}^3$. סכומם מוגדר על ידי

$$(5 - 6x + 4x^2, (3,8,7)) + (x + 9x^5, (2,2,2))$$
$$= (5 - 5x + 4x^2 + 9x^5, (5,10,9)).$$

בנוסף, $2(5 - 6x + 4x^2, (3,8,7)) = (10 - 12x + 8x^2, (6,16,14))$.

את המסקנה הבאה יש לפרש כך: מכפלת מרחבים וקטוריים בצירוף פעולות החיבור והכפל בסקלר כפי שהוגדרו ב־ 3.87, היא מרחב וקטורי.

3.89 מכפלת מרחבים וקטוריים היא מרחב וקטורי

נניח ש־ $V_1, \dots, V_m$ מרחבים וקטוריים מעל $\mathbf{F}$. אזי $V_1 \times \cdots \times V_m$ מרחב וקטורי מעל $\mathbf{F}$.

ההוכחה למסקנה לעיל מושארת לקוראים. שימו לב כי איבר הזהות לחיבור של $V_1 \times \cdots \times V_m$ הוא $(0, \dots, 0)$, כאשר ה־ 0 ברכיב ה־ k^{th} הוא איבר הזהות לחיבור של V_k. האיבר הנגדי לחיבור של $(v_1, \dots, v_m) \in V_1 \times \cdots \times V_m$ הוא $(-v_1, \dots, -v_m)$.

3.90 example: $\mathbf{R}^2 \times \mathbf{R}^3 \neq \mathbf{R}^5$ *but* $\mathbf{R}^2 \times \mathbf{R}^3$ *is isomorphic to* $\mathbf{R}^5$

Elements of the vector space $\mathbf{R}^2 \times \mathbf{R}^3$ are lists

$$((x_1, x_2), (x_3, x_4, x_5)),$$

where $x_1, x_2, x_3, x_4, x_5 \in \mathbf{R}$. Elements of $\mathbf{R}^5$ are lists

$$(x_1, x_2, x_3, x_4, x_5),$$

where $x_1, x_2, x_3, x_4, x_5 \in \mathbf{R}$.

Although elements of $\mathbf{R}^2 \times \mathbf{R}^3$ and $\mathbf{R}^5$ look similar, they are not the same kind of object. Elements of $\mathbf{R}^2 \times \mathbf{R}^3$ are lists of length two (with the first item itself a list of length two and the second item a list of length three), and elements of $\mathbf{R}^5$ are lists of length five. Thus $\mathbf{R}^2 \times \mathbf{R}^3$ does not equal $\mathbf{R}^5$.

The linear map

$$((x_1, x_2), (x_3, x_4, x_5)) \mapsto (x_1, x_2, x_3, x_4, x_5)$$

is an isomorphism of the vector space $\mathbf{R}^2 \times \mathbf{R}^3$ onto the vector space $\mathbf{R}^5$. Thus these two vector spaces are isomorphic, although they are not equal.

This isomorphism is so natural that we should think of it as a relabeling. Some people informally say that $\mathbf{R}^2 \times \mathbf{R}^3$ *equals* $\mathbf{R}^5$*, which is not technically correct but which captures the spirit of identification via relabeling.*

The next example illustrates the idea that we will use in the proof of 3.92.

3.91 example: *a basis of* $\mathcal{P}_2(\mathbf{R}) \times \mathbf{R}^2$

Consider this list of length five of elements of $\mathcal{P}_2(\mathbf{R}) \times \mathbf{R}^2$:

$$(1, (0, 0)), (x, (0, 0)), (x^2, (0, 0)), (0, (1, 0)), (0, (0, 1)).$$

The list above is linearly independent and it spans $\mathcal{P}_2(\mathbf{R}) \times \mathbf{R}^2$. Thus it is a basis of $\mathcal{P}_2(\mathbf{R}) \times \mathbf{R}^2$.

3.92 *dimension of a product is the sum of dimensions*

Suppose $V_1, \ldots, V_m$ are finite-dimensional vector spaces. Then $V_1 \times \cdots \times V_m$ is finite-dimensional and

$$\dim(V_1 \times \cdots \times V_m) = \dim V_1 + \cdots + \dim V_m.$$

Proof Choose a basis of each V_k. For each basis vector of each V_k, consider the element of $V_1 \times \cdots \times V_m$ that equals the basis vector in the kth slot and 0 in the other slots. The list of all such vectors is linearly independent and spans $V_1 \times \cdots \times V_m$. Thus it is a basis of $V_1 \times \cdots \times V_m$. The length of this basis is $\dim V_1 + \cdots + \dim V_m$, as desired.

3.90 דוגמה: $\mathbf{R}^2 \times \mathbf{R}^3 \neq \mathbf{R}^5$ אולם $\mathbf{R}^2 \times \mathbf{R}^3$ איזומורפי ל־ $\mathbf{R}^5$

איברי המרחב הווקטורי $\mathbf{R}^2 \times \mathbf{R}^3$ הם רשימות

$$((x_1, x_2), (x_3, x_4, x_5)),$$

כאשר $x_1, x_2, x_3, x_4, x_5 \in \mathbf{R}$. האיברים של $\mathbf{R}^5$ הם רשימות

$$(x_1, x_2, x_3, x_4, x_5),$$

כאשר $x_1, x_2, x_3, x_4, x_5 \in \mathbf{R}$.

למרות הדמיון בין איברי $\mathbf{R}^2 \times \mathbf{R}^3$ לאיברי $\mathbf{R}^5$, האיברים הם אובייקטים מטיפוסים שונים. איברי $\mathbf{R}^2 \times \mathbf{R}^3$ הם רשימות בנות שני איברים (כאשר האיבר הראשון הוא עצמו רשימה בת שני איברים והאיבר השני הוא רשימה בת שלושה איברים), ואילו האיברים ב־ $\mathbf{R}^5$ הם רשימות בנות חמישה איברים. לפיכך $\mathbf{R}^2 \times \mathbf{R}^3$ אינו שווה ל־ $\mathbf{R}^5$.

ההעתקה הליניארית

$$((x_1, x_2), (x_3, x_4, x_5)) \mapsto (x_1, x_2, x_3, x_4, x_5)$$

היא איזומורפיזם של המרחב הווקטורי $\mathbf{R}^2 \times \mathbf{R}^3$ על המרחב הווקטורי $\mathbf{R}^5$. לפיכך שני מרחבים וקטוריים אלו הם איזומורפיים, אף שאינם שווים.

איזומורפיזם זה הוא כה טבעי שנוכל לחשוב עליו כתיוג מחדש. יש האומרים באופן בלתי רשמי כי $\mathbf{R}^2 \times \mathbf{R}^3$ שווה ל־ $\mathbf{R}^5$, אמירה שאינה נכונה טכנית, אבל יש בה כדי להמחיש את רוח הדברים של הזיהוי באמצעות תיוג מחדש.

הדוגמה הבאה ממחישה את הרעיון שבו נשתמש בהוכחת 3.92.

3.91 דוגמה: בסיס של $\mathcal{P}_2(\mathbf{R}) \times \mathbf{R}^2$

נתבונן ברשימה זו בת חמישה איברים של $\mathcal{P}_2(\mathbf{R}) \times \mathbf{R}^2$:

$$(1, (0, 0)), (x, (0, 0)), (x^2, (0, 0)), (0, (1, 0)), (0, (0, 1)).$$

רשימה זו בלתי תלויה ליניארית ופורסת את $\mathcal{P}_2(\mathbf{R}) \times \mathbf{R}^2$. לפיכך היא בסיס של $\mathcal{P}_2(\mathbf{R}) \times \mathbf{R}^2$.

3.92 ממד המכפלה הוא סכום הממדים

נניח ש־ $V_1, \dots, V_m$ מרחבים וקטוריים מממד סופי. אזי $V_1 \times \cdots \times V_m$ מממד סופי ו־

$$\dim(V_1 \times \cdots \times V_m) = \dim V_1 + \cdots + \dim V_m.$$

הוכחה נבחר בסיס לכל V_k. עבור כל וקטור בסיס של כל V_k, נתבונן באיבר של $V_1 \times \cdots \times V_m$ השווה לווקטור הבסיס ברכיב ה־ k^{th} ו־ 0 בשאר הרכיבים. רשימת כל הווקטורים הללו בלתי תלויה ליניארית ופורסת את $V_1 \times \cdots \times V_m$. לפיכך היא בסיס של $V_1 \times \cdots \times V_m$. אורכו של בסיס זה הוא $\dim V_1 + \cdots + \dim V_m$, כפי שרצינו להוכיח. ■

In the next result, the map Γ is surjective by the definition of $V_1 + \cdots + V_m$. Thus the last word in the result below could be changed from "injective" to "invertible".

3.93 *products and direct sums*

Suppose that $V_1, \ldots, V_m$ are subspaces of V. Define a linear map $\Gamma : V_1 \times \cdots \times V_m \to V_1 + \cdots + V_m$ by

$$\Gamma(v_1, \ldots, v_m) = v_1 + \cdots + v_m.$$

Then $V_1 + \cdots + V_m$ is a direct sum if and only if Γ is injective.

Proof By 3.15, Γ is injective if and only if the only way to write 0 as a sum $v_1 + \cdots + v_m$, where each v_k is in V_k, is by taking each v_k equal to 0. Thus 1.45 shows that Γ is injective if and only if $V_1 + \cdots + V_m$ is a direct sum, as desired. ∎

3.94 *a sum is a direct sum if and only if dimensions add up*

Suppose V is finite-dimensional and $V_1, \ldots, V_m$ are subspaces of V. Then $V_1 + \cdots + V_m$ is a direct sum if and only if

$$\dim(V_1 + \cdots + V_m) = \dim V_1 + \cdots + \dim V_m.$$

Proof The map Γ in 3.93 is surjective. Thus by the fundamental theorem of linear maps (3.21), Γ is injective if and only if

$$\dim(V_1 + \cdots + V_m) = \dim(V_1 \times \cdots \times V_m).$$

Combining 3.93 and 3.92 now shows that $V_1 + \cdots + V_m$ is a direct sum if and only if

$$\dim(V_1 + \cdots + V_m) = \dim V_1 + \cdots + \dim V_m,$$

as desired. ∎

In the special case $m = 2$, an alternative proof that $V_1 + V_2$ is a direct sum if and only if $\dim(V_1 + V_2) = \dim V_1 + \dim V_2$ can be obtained by combining 1.46 and 2.43.

Quotient Spaces

We begin our approach to quotient spaces by defining the sum of a vector and a subset.

3.95 notation: $v + U$

Suppose $v \in V$ and $U \subseteq V$. Then $v + U$ is the subset of V defined by

$$v + U = \{v + u : u \in U\}.$$

במסקנה הבאה, ההעתקה Γ היא על, לפי הגדרת $V_1 + \cdots + V_m$. לפיכך, בסוף המסקנה להלן נוכל להחליף את הביטוי "חד־חד־ערכית" ב"הפיכה".

3.93 מכפלות וסכומים ישרים

נניח ש־ $V_1, \ldots, V_m$ תתי־מרחב של V. נגדיר העתקה ליניארית $\Gamma : V_1 \times \cdots \times V_m \to V_1 + \cdots + V_m$ על ידי

$$\Gamma(v_1, \ldots, v_m) = v_1 + \cdots + v_m.$$

אזי $V_1 + \cdots + V_m$ סכום ישר אם־ורק־אם Γ חד־חד־ערכית.

הוכחה לפי 3.15, Γ חד־חד־ערכית אם־ורק־אם האופן היחיד שבו ניתן לבטא 0 כסכום $v_1 + \cdots + v_m$, כאשר כל v_k ב־ V_k, הוא שכל v_k שווה 0. לפיכך, לפי 1.45 נקבל ש־ Γ חד־חד־ערכית אם־ורק־אם $V_1 + \cdots + V_m$ הוא סכום ישר, כפי שרצינו. ■

3.94 סכום הוא סכום ישר אם־ורק־אם המימדים מסתכמים

נניח ש־ V מממד סופי ו־ $V_1, \ldots, V_m$ תתי־מרחב של V. אזי $V_1 + \cdots + V_m$ הוא סכום ישר אם־ורק־אם

$$\dim(V_1 + \cdots + V_m) = \dim V_1 + \cdots + \dim V_m.$$

הוכחה ההעתקה Γ ב־ 3.93 היא על. לפיכך, לפי המשפט היסודי של ההעתקות הליניאריות (3.21), Γ חד־חד־ערכית אם־ורק־אם

$$\dim(V_1 + \cdots + V_m) = \dim(V_1 \times \cdots \times V_m).$$

בשילוב 3.93 עם 3.92 אנו מקבלים ש־ $V_1 + \cdots + V_m$ סכום ישר אם־ורק־אם

$$\dim(V_1 + \cdots + V_m) = \dim V_1 + \cdots + \dim V_m,$$

כפי שרצינו. ■

עבור המקרה הפרטי $m = 2$, ניתן להפיק הוכחה חלופית לכך ש־ $V_1 + V_2$ סכום ישר אם־ורק־אם $\dim(V_1 + V_2) = \dim V_1 + \dim V_2$ על ידי שילוב 1.46 ו־ 2.43.

מרחבי מנה

בבואנו לעסוק במרחבי מנה, נפתח בהגדרת סכום של וקטור ותת־קבוצה.

3.95 סימון: $v + U$

נניח ש־ $v \in V$ ו־ $U \subseteq V$. אזי $v + U$ היא התת־קבוצה של V המוגדרת על ידי

$$v + U = \{v + u : u \in U\}.$$

3.96 example: *sum of a vector and a one-dimensional subspace of* $\mathbf{R}^2$

Suppose

$$U = \{(x, 2x) \in \mathbf{R}^2 : x \in \mathbf{R}\}.$$

Hence U is the line in $\mathbf{R}^2$ through the origin with slope 2. Thus

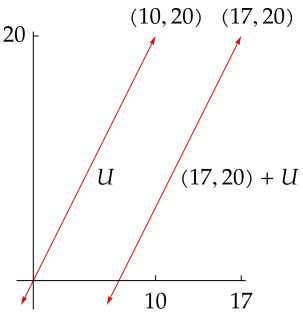
$$(17, 20) + U$$

is the line in $\mathbf{R}^2$ that contains the point $(17, 20)$ and has slope 2.

Because

$$(10, 20) \in U \quad \text{and} \quad (17, 20) \in (17, 20) + U,$$

we see that $(17, 20) + U$ is obtained by moving U to the right by 7 units.



$(17, 20) + U$ *is parallel to the subspace* U.

3.97 definition: *translate*

For $v \in V$ and U a subset of V, the set $v + U$ is said to be a *translate* of U.

3.98 example: *translates*

- If U is the line in $\mathbf{R}^2$ defined by $U = \{(x, 2x) \in \mathbf{R}^2 : x \in \mathbf{R}\}$, then all lines in $\mathbf{R}^2$ with slope 2 are translates of U. See Example 3.96 above for a drawing of U and one of its translates.
- More generally, if U is a line in $\mathbf{R}^2$, then the set of all translates of U is the set of all lines in $\mathbf{R}^2$ that are parallel to U.
- If $U = \{(x, y, 0) \in \mathbf{R}^3 : x, y \in \mathbf{R}\}$, then the translates of U are the planes in $\mathbf{R}^3$ that are parallel to the xy-plane U.
- More generally, if U is a plane in $\mathbf{R}^3$, then the set of all translates of U is the set of all planes in $\mathbf{R}^3$ that are parallel to U (see, for example, Exercise 7).

3.99 definition: *quotient space,* V/U

Suppose U is a subspace of V. Then the *quotient space* V/U is the set of all translates of U. Thus

$$V/U = \{v + U : v \in V\}.$$

3.96 דוגמה: סכום של וקטור ותת־מרחב חז־ממדי של $\mathbf{R}^2$

נניח ש־

$$U = \{(x, 2x) \in \mathbf{R}^2 : x \in \mathbf{R}\}.$$

לכן U הוא הישר ב־ $\mathbf{R}^2$ העובר דרך הראשית ששיפועו 2. לפיכך

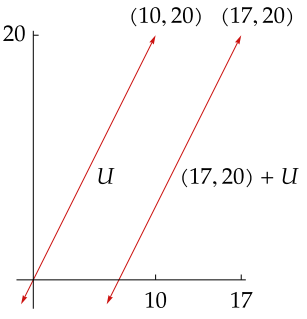
$$(17, 20) + U$$

הוא הישר ב־ $\mathbf{R}^2$ המכיל את הנקודה $(17, 20)$ ושיפועו 2.

מאחר ש־

$$(17, 20) \in (17, 20) + U, \quad \text{וגם} \quad (10, 20) \in U$$

אנו רואים ש־ $(17, 20) + U$ מושג באמצעות הזזת U ב־ 7 יחידות לימין.



$(17, 20) + U$ מקביל לתת־מרחב U.

3.97 הגדרה: הזזה

עבור $v \in V$ ו־ U תת־קבוצה של V, הקבוצה $v + U$ נקראת הזזה של U.

3.98 דוגמה: הזזות

- אם U הוא הישר ב־ $\mathbf{R}^2$ המוגדר על ידי $U = \{(x, 2x) \in \mathbf{R}^2 : x \in \mathbf{R}\}$, אזי כל הישרים ב־ $\mathbf{R}^2$ ששיפועם 2 הם הזזות של U. ראו דוגמה 3.96 לעיל לשרטוט של U ואחת מההזזות שלו.
- ובאופן כללי, אם U הוא ישר ב־ $\mathbf{R}^2$, אזי, קבוצת כל ההזזות של U היא קבוצת כל הישרים ב־ $\mathbf{R}^2$ המקבילים ל־ U.
- אם $U = \{(x, y, 0) \in \mathbf{R}^3 : x, y \in \mathbf{R}\}$, אזי, ההזזות של U הן המישורים ב־ $\mathbf{R}^3$ המקבילים ל־ U, מישור ה־ xy.
- ובאופן כללי, אם U הוא מישור ב־ $\mathbf{R}^3$, אזי, קבוצת כל ההזזות של U היא קבוצת כל המישורים ב־ $\mathbf{R}^3$ המקבילים ל־ U (ראו, לדוגמה, תרגיל 7).

3.99 הגדרה: מרחב מנה, V/U

נניח ש־ U תת־מרחב של V. אזי מרחב המנה V/U הוא קבוצת כל ההזזות של U. לפיכך

$$V/U = \{v + U : v \in V\}.$$

3.100 example: *quotient spaces*

- If $U = \{(x, 2x) \in \mathbf{R}^2 : x \in \mathbf{R}\}$, then $\mathbf{R}^2/U$ is the set of all lines in $\mathbf{R}^2$ that have slope 2.
- If U is a line in $\mathbf{R}^3$ containing the origin, then $\mathbf{R}^3/U$ is the set of all lines in $\mathbf{R}^3$ parallel to U.
- If U is a plane in $\mathbf{R}^3$ containing the origin, then $\mathbf{R}^3/U$ is the set of all planes in $\mathbf{R}^3$ parallel to U.

Our next goal is to make V/U into a vector space. To do this, we will need the next result.

3.101 *two translates of a subspace are equal or disjoint*

Suppose U is a subspace of V and $v, w \in V$. Then

$$v - w \in U \iff v + U = w + U \iff (v + U) \cap (w + U) \neq \emptyset.$$

Proof First suppose $v - w \in U$. If $u \in U$, then

$$v + u = w + ((v - w) + u) \in w + U.$$

Thus $v + U \subseteq w + U$. Similarly, $w + U \subseteq v + U$. Thus $v + U = w + U$, completing the proof that $v - w \in U$ implies $v + U = w + U$.

The equation $v + U = w + U$ implies that $(v + U) \cap (w + U) \neq \emptyset$.

Now suppose $(v + U) \cap (w + U) \neq \emptyset$. Thus there exist $u_1, u_2 \in U$ such that

$$v + u_1 = w + u_2.$$

Thus $v - w = u_2 - u_1$. Hence $v - w \in U$, showing that $(v + U) \cap (w + U) \neq \emptyset$ implies $v - w \in U$, which completes the proof.

Now we can define addition and scalar multiplication on V/U.

3.102 definition: *addition and scalar multiplication on V/U*

Suppose U is a subspace of V. Then *addition* and *scalar multiplication* are defined on V/U by

$$(v + U) + (w + U) = (v + w) + U$$
$$\lambda(v + U) = (\lambda v) + U$$

for all $v, w \in V$ and all $\lambda \in \mathbf{F}$.

As part of the proof of the next result, we will show that the definitions above make sense.

3.100 דוגמה: מרחבי מנה

- אם $U = \{(x, 2x) \in \mathbf{R}^2 : x \in \mathbf{R}\}$, אזי $\mathbf{R}^2/U$ הוא קבוצת כל הישרים ב־ $\mathbf{R}^2$ ששיפועם 2.
- אם U הוא ישר ב־ $\mathbf{R}^3$ העובר דרך הראשית, אזי $\mathbf{R}^3/U$ הוא קבוצת כל הישרים ב־ $\mathbf{R}^3$ המקבילים ל־ U.
- אם U הוא מישור ב־ $\mathbf{R}^3$ העובר דרך הראשית, אזי $\mathbf{R}^3/U$ הוא קבוצת כל המישורים ב־ $\mathbf{R}^3$ המקבילים ל־ U.

יעדנו הבא הוא להפוך את V/U למרחב וקטורי. לשם כך, נזדקק למסקנה הבאה.

3.101 שתי הזזות של תת־מרחב הן שוות או זרות

נניח ש־ U תת־מרחב של V ו־ $v, w \in V$. אזי

$$v - w \in U \iff v + U = w + U \iff (v + U) \cap (w + U) \neq \emptyset.$$

הוכחה תחילה נניח ש־ $v - w \in U$. אם $u \in U$, אזי

$$v + u = w + ((v - w) + u) \in w + U.$$

לפיכך, $v + U \subseteq w + U$. באופן דומה, $w + U \subseteq v + U$. לפיכך $v + U = w + U$, ובכך השלמנו את ההוכחה ש־ $v - w \in U$ גורר ש־ $v + U = w + U$.

מהמשוואה $v + U = w + U$ נובע ש־ $(v + U) \cap (w + U) \neq \emptyset$.

כעת נניח ש־ $(v + U) \cap (w + U) \neq \emptyset$. לפיכך קיימים $u_1, u_2 \in U$ כך ש־

$$v + u_1 = w + u_2.$$

לפיכך, $v - w = u_2 - u_1$. לכן $v - w \in U$, ובכך הראינו ש־ $(v + U) \cap (w + U) \neq \emptyset$ גורר ש־ $v - w \in U$, והשלמנו את ההוכחה. ■

כעת נוכל להגדיר חיבור וכפל בסקלר מעל V/U.

3.102 הגדרה: חיבור וכפל בסקלר מעל V/U

נניח ש־ U תת־מרחב של V. אזי חיבור וכפל בסקלר מוגדרים מעל V/U על ידי

$$(v + U) + (w + U) = (v + w) + U$$

$$\lambda(v + U) = (\lambda v) + U$$

לכל $v, w \in V$ ולכל $\lambda \in \mathbf{F}$.

כחלק מהוכחת המסקנה הבאה, נצטרך להראות שההגדרות לעיל הגיוניות.

3.103 *quotient space is a vector space*

Suppose U is a subspace of V. Then V/U, with the operations of addition and scalar multiplication as defined above, is a vector space.

Proof The potential problem with the definitions above of addition and scalar multiplication on V/U is that the representation of a translate of U is not unique. Specifically, suppose $v_1, v_2, w_1, w_2 \in V$ are such that

$$v_1 + U = v_2 + U \quad \text{and} \quad w_1 + U = w_2 + U.$$

To show that the definition of addition on V/U given above makes sense, we must show that $(v_1 + w_1) + U = (v_2 + w_2) + U$.

By 3.101, we have

$$v_1 - v_2 \in U \quad \text{and} \quad w_1 - w_2 \in U.$$

Because U is a subspace of V and thus is closed under addition, this implies that $(v_1 - v_2) + (w_1 - w_2) \in U$. Thus $(v_1 + w_1) - (v_2 + w_2) \in U$. Using 3.101 again, we see that

$$(v_1 + w_1) + U = (v_2 + w_2) + U,$$

as desired. Thus the definition of addition on V/U makes sense.

Similarly, suppose $\lambda \in \mathbf{F}$. We are still assuming that $v_1 + U = v_2 + U$. Because U is a subspace of V and thus is closed under scalar multiplication, we have $\lambda(v_1 - v_2) \in U$. Thus $\lambda v_1 - \lambda v_2 \in U$. Hence 3.101 implies that

$$(\lambda v_1) + U = (\lambda v_2) + U.$$

Thus the definition of scalar multiplication on V/U makes sense.

Now that addition and scalar multiplication have been defined on V/U, the verification that these operations make V/U into a vector space is straightforward and is left to the reader. Note that the additive identity of V/U is $0 + U$ (which equals U) and that the additive inverse of $v + U$ is $(-v) + U$.

The next concept will lead to a computation of the dimension of V/U.

3.104 definition: *quotient map, π*

Suppose U is a subspace of V. The *quotient map* $\pi : V \to V/U$ is the linear map defined by

$$\pi(v) = v + U$$

for each $v \in V$.

The reader should verify that π is indeed a linear map. Although π depends on U as well as V, these spaces are left out of the notation because they should be clear from the context.

3.103 מרחב מנה הוא מרחב וקטורי

נניח ש־ U תת־מרחב של V. אזי V/U, בצירוף פעולות החיבור והכפל בסקלר כפי שהוגדרו לעיל, מרחב וקטורי.

הוכחה הבעיה הפוטנציאלית בהגדרות החיבור והכפל בסקלר לעיל מעל V/U היא שלהזזה של U אין ייצוג יחידי. ליתר דיוק, נניח ש־ $v_1, v_2, w_1, w_2 \in V$ הם כך ש־

$$v_1 + U = v_2 + U \quad \text{וגם} \quad w_1 + U = w_2 + U.$$

כדי להראות שהגדרת החיבור מעל V/U כפי שהוגדרה לעיל הגיונית, עלינו להראות ש־ $(v_1 + w_1) + U = (v_2 + w_2) + U$.

מ־ 3.101, נקבל ש־

$$v_1 - v_2 \in U \quad \text{וגם} \quad w_1 - w_2 \in U.$$

מאחר ש־ U תת־מרחב של V ולפיכך סגור תחת פעולת החיבור, מכאן נובע ש־ $(v_1 - v_2) + (w_1 - w_2) \in U$. לפיכך $(v_1 + w_1) - (v_2 + w_2) \in U$. שימוש חוזר ב־ 3.101 מראה ש־

$$(v_1 + w_1) + U = (v_2 + w_2) + U,$$

כפי שרצינו. לפיכך, הגדרת החיבור מעל V/U הגיונית.

באופן דומה, נניח ש־ $\lambda \in \mathbf{F}$. אנו עדיין תחת ההנחה ש־ $v_1 + U = v_2 + U$. מאחר ש־ U תת־מרחב של V ולפיכך סגור תחת פעולת כפל בסקלר, נקבל ש־ $\lambda(v_1 - v_2) \in U$. לפיכך $\lambda v_1 - \lambda v_2 \in U$. לכן 3.101 גורר ש־

$$(\lambda v_1) + U = (\lambda v_2) + U.$$

לפיכך ההגדרה של פעולת הכפל בסקלר מעל V/U הגיונית.

כעת משהוגדרו פעולות החיבור והכפל בסקלר מעל V/U, הווידוא שפעולות אלו אכן הופכות את V/U למרחב וקטורי הוא פשוט, ומושאר לקוראים. שימו לב שאיבר הזהות לחיבור של V/U הוא $0 + U$ (השווה ל־ U) והאיבר הנגדי לחיבור של $v + U$ הוא $(-v) + U$. ■

הרעיון הבא יוביל אותנו לאופן חישוב הממד של V/U.

3.104 הגדרה: העתקת המנה, π

נניח ש־ U תת־מרחב של V. העתקת המנה $\pi : V \to V/U$ היא ההעתקה הליניארית המוגדרת על ידי

$$\pi(v) = v + U$$

לכל $v \in V$.

על הקוראים לוודא כי π היא אכן העתקה ליניארית. אף ש־ π תלויה ב־ U כמו גם ב־ V, מרחבים אלו אינם נכללים בסימון, מאחר שזהותם ברורה מתוך ההקשר.

3.105 *dimension of quotient space*

Suppose V is finite-dimensional and U is a subspace of V. Then

$$\dim V/U = \dim V - \dim U.$$

Proof Let π denote the quotient map from V to V/U. If $v \in V$, then $v+U = 0+U$ if and only if $v \in U$ (by 3.101), which implies that $\operatorname{null} \pi = U$. The definition of π implies $\operatorname{range} \pi = V/U$. The fundamental theorem of linear maps (3.21) now implies $\dim V = \dim U + \dim V/U$, which gives the desired result. ∎

Each linear map T on V induces a linear map $\widetilde{T}$ on $V/(\operatorname{null} T)$, as defined below.

3.106 notation: $\widetilde{T}$

Suppose $T \in \mathcal{L}(V, W)$. Define $\widetilde{T} \colon V/(\operatorname{null} T) \to W$ by

$$\widetilde{T}(v + \operatorname{null} T) = Tv.$$

To show that the definition of $\widetilde{T}$ makes sense, suppose $u, v \in V$ are such that $u + \operatorname{null} T = v + \operatorname{null} T$. By 3.101, we have $u - v \in \operatorname{null} T$. Thus $T(u - v) = 0$. Hence $Tu = Tv$. Thus the definition of $\widetilde{T}$ indeed makes sense. The routine verification that $\widetilde{T}$ is a linear map from $V/(\operatorname{null} T)$ to W is left to the reader.

The next result shows that we can think of $\widetilde{T}$ as a modified version of T, with a domain that produces a one-to-one map.

3.107 *null space and range of* $\widetilde{T}$

Suppose $T \in \mathcal{L}(V, W)$. Then

(a) $\widetilde{T} \circ \pi = T$, where π is the quotient map of V onto $V/(\operatorname{null} T)$;

(b) $\widetilde{T}$ is injective;

(c) $\operatorname{range} \widetilde{T} = \operatorname{range} T$;

(d) $V/(\operatorname{null} T)$ and $\operatorname{range} T$ are isomorphic vector spaces.

Proof

(a) If $v \in V$, then $(\widetilde{T} \circ \pi)(v) = \widetilde{T}(\pi(v)) = \widetilde{T}(v + \operatorname{null} T) = Tv$, as desired.

(b) Suppose $v \in V$ and $\widetilde{T}(v + \operatorname{null} T) = 0$. Then $Tv = 0$. Thus $v \in \operatorname{null} T$. Hence 3.101 implies that $v + \operatorname{null} T = 0 + \operatorname{null} T$. This implies that $\operatorname{null} \widetilde{T} = \{0 + \operatorname{null} T\}$. Hence $\widetilde{T}$ is injective, as desired.

(c) The definition of $\widetilde{T}$ shows that $\operatorname{range} \widetilde{T} = \operatorname{range} T$.

(d) Now (b) and (c) imply that if we think of $\widetilde{T}$ as mapping into $\operatorname{range} T$, then $\widetilde{T}$ is an isomorphism from $V/(\operatorname{null} T)$ onto $\operatorname{range} T$. ∎

3.105 ממד מרחב מנה

נניח ש־ V מממד סופי ו־ U תת־מרחב של V. אזי
$$\dim V/U = \dim V - \dim U.$$

הוכחה תהא π העתקת המנה מ־ V ל־ V/U. אם $v \in V$, אזי $v + U = 0 + U$ אם־ורק־אם $v \in U$ (לפי 3.101), ומכך נסיק כי $\operatorname{null} \pi = U$. מהגדרת π נובע ש־ $\operatorname{range} \pi = V/U$. לפי המשפט היסודי של ההעתקות הליניאריות (3.21) נסיק כעת כי $\dim V = \dim U + \dim V/U$, ובכך קיבלנו את התוצאה המבוקשת. ∎

כל העתקה ליניארית T מעל V משרה העתקה ליניארית $\widetilde{T}$ מעל $V/(\operatorname{null} T)$, לפי ההגדרה להלן.

3.106 סימון: $\widetilde{T}$

נניח ש־ $T \in \mathcal{L}(V, W)$. נגדיר $\widetilde{T}\colon V/(\operatorname{null} T) \to W$ על ידי
$$\widetilde{T}(v + \operatorname{null} T) = Tv.$$

כדי להראות שהגדרת $\widetilde{T}$ הגיונית, נניח כי $u, v \in V$ כך ש־ $u + \operatorname{null} T = v + \operatorname{null} T$. לפי 3.101, נסיק ש־ $u - v \in \operatorname{null} T$. לפיכך $T(u - v) = 0$. לכן $Tu = Tv$. לפיכך הגדרת $\widetilde{T}$ בהחלט הגיונית. הווידוא השגרתי ש־ $\widetilde{T}$ העתקה ליניארית מ־ $V/(\operatorname{null} T)$ אל W מושאר לקוראים.

במסקנה הבאה נראה כי ניתן לחשוב על $\widetilde{T}$ כגרסה מותאמת של T, הפועלת בתחום המפיק העתקה חד־חד־ערכית.

3.107 מרחב האפס והטווח של $\widetilde{T}$

נניח ש־ $T \in \mathcal{L}(V, W)$. אזי

א. $T = \widetilde{T} \circ \pi$, כאשר π העתקת המנה מ־ V על $V/(\operatorname{null} T)$

ב. $\widetilde{T}$ חד־חד־ערכית

ג. $\operatorname{range} \widetilde{T} = \operatorname{range} T$

ד. $V/(\operatorname{null} T)$ ו־ $\operatorname{range} T$ מרחבים וקטוריים איזומורפיים.

הוכחה

א. אם $v \in V$, אזי $\widetilde{T}(\pi(v)) = \widetilde{T}(v + \operatorname{null} T) = Tv$, כפי שרצינו.

ב. נניח ש־ $v \in V$ וגם $\widetilde{T}(v + \operatorname{null} T) = 0$. אזי $Tv = 0$. לפיכך $v \in \operatorname{null} T$. לכן 3.101 גורר ש־ $v + \operatorname{null} T = 0 + \operatorname{null} T$. מכאן נובע ש־ $\operatorname{null} \widetilde{T} = \{0 + \operatorname{null} T\}$. לכן $\widetilde{T}$ חד־חד־ערכית, כפי שרצינו.

ג. מהגדרת $\widetilde{T}$ נובע ש־ $\operatorname{range} \widetilde{T} = \operatorname{range} T$.

ד. כעת נוכל להסיק מ־ (ב) ו־ (ג) שאם נחשוב על $\widetilde{T}$ כהעתקה אל $\operatorname{range} T$, נקבל ש־ $\widetilde{T}$ היא איזומורפיזם מ־ $V/(\operatorname{null} T)$ על $\operatorname{range} T$. ∎

Exercises 3E

1 Suppose T is a function from V to W. The *graph* of T is the subset of $V \times W$ defined by

$$\text{graph of } T = \{(v, Tv) \in V \times W : v \in V\}.$$

Prove that T is a linear map if and only if the graph of T is a subspace of $V \times W$.

Formally, a function T from V to W is a subset T of $V \times W$ such that for each $v \in V$, there exists exactly one element $(v, w) \in T$. In other words, formally a function is what is called above its graph. We do not usually think of functions in this formal manner. However, if we do become formal, then this exercise could be rephrased as follows: Prove that a function T from V to W is a linear map if and only if T is a subspace of $V \times W$.

2 Suppose that $V_1, \dots, V_m$ are vector spaces such that $V_1 \times \cdots \times V_m$ is finite-dimensional. Prove that V_k is finite-dimensional for each $k = 1, \dots, m$.

3 Suppose $V_1, \dots, V_m$ are vector spaces. Prove that $\mathcal{L}(V_1 \times \cdots \times V_m, W)$ and $\mathcal{L}(V_1, W) \times \cdots \times \mathcal{L}(V_m, W)$ are isomorphic vector spaces.

There is no assumption in the exercise above or in the two following exercises that the vector spaces are finite-dimensional.

4 Suppose $W_1, \dots, W_m$ are vector spaces. Prove that $\mathcal{L}(V, W_1 \times \cdots \times W_m)$ and $\mathcal{L}(V, W_1) \times \cdots \times \mathcal{L}(V, W_m)$ are isomorphic vector spaces.

5 For m a positive integer, define V^m by

$$V^m = \underbrace{V \times \cdots \times V}_{m \text{ times}}.$$

Prove that V^m and $\mathcal{L}(\mathbf{F}^m, V)$ are isomorphic vector spaces.

6 Suppose that v, x are vectors in V and that U, W are subspaces of V such that $v + U = x + W$. Prove that $U = W$.

7 Let $U = \{(x, y, z) \in \mathbf{R}^3 : 2x + 3y + 5z = 0\}$. Suppose $A \subseteq \mathbf{R}^3$. Prove that A is a translate of U if and only if there exists $c \in \mathbf{R}$ such that

$$A = \{(x, y, z) \in \mathbf{R}^3 : 2x + 3y + 5z = c\}.$$

8 (a) Suppose $T \in \mathcal{L}(V, W)$ and $c \in W$. Prove that $\{x \in V : Tx = c\}$ is either the empty set or is a translate of null T.

(b) Explain why the set of solutions to a system of linear equations such as 3.27 is either the empty set or is a translate of some subspace of $\mathbf{F}^n$.

9 Prove that a nonempty subset A of V is a translate of some subspace of V if and only if $\lambda v + (1 - \lambda)w \in A$ for all $v, w \in A$ and all $\lambda \in \mathbf{F}$.

10 Suppose $A_1 = v + U_1$ and $A_2 = w + U_2$ for some $v, w \in V$ and some subspaces U_1, U_2 of V. Prove that the intersection $A_1 \cap A_2$ is either a translate of some subspace of V or is the empty set.

תרגילים 5ג

1 נניח ש־ T פונקציה מ־ V אל W. הגרף של T הוא התת־קבוצה של $V \times W$ המוגדרת על ידי

$$\text{הגרף של } T = \{(v, Tv) \in V \times W : v \in V\}.$$

הוכיחו ש־ T העתקה ליניארית אם־ורק־אם הגרף של T הוא תת־מרחב של $V \times W$.

באופן פורמלי, פונקציה T מ־ V אל W היא תת־קבוצה T של $V \times W$ כך שלכל $v \in V$, קיים איבר יחיד $(v, w) \in T$. בפילים אחרות, ההגדרה הפורמלית של פונקציה היא מה שהגדרנו לעיל כגרף שלה. בדרך כלל איננו חושבים על פונקציות בהיבט הפורמלי. בכל זאת, לאור ההגדרה הפורמלית, נוכל לנסח תרגיל זה באופן הבא: הוכיחו שפונקציה T מ־ V אל W היא העתקה ליניארית אם־ורק־אם T תת־מרחב של $V \times W$.

2 נניח ש־ $V_1, \dots, V_m$ מרחבים וקטוריים כך ש־ $V_1 \times \cdots \times V_m$ מממד סופי. הוכיחו ש־ V_k מממד סופי לכל $k = 1, \dots, m$.

3 נניח ש־ $V_1, \dots, V_m$ מרחבים וקטוריים. הוכיחו ש־ $\mathcal{L}(V_1 \times \cdots \times V_m, W)$ ו־ $\mathcal{L}(V_1, W) \times \cdots \times \mathcal{L}(V_m, W)$ מרחבים וקטוריים איזומורפיים.

אין הנחה בתרגיל לעיל או בשני התרגילים הבאים שהמרחבים הווקטוריים מממד סופי.

4 נניח ש־ $W_1, \dots, W_m$ מרחבים וקטוריים. הוכיחו ש־ $\mathcal{L}(V, W_1 \times \cdots \times W_m)$ ו־ $\mathcal{L}(V, W_1) \times \cdots \times \mathcal{L}(V, W_m)$ מרחבים וקטוריים איזומורפיים.

5 עבור m מספר שלם חיובי, הגדירו V^m על ידי

$$V^m = \underbrace{V \times \cdots \times V}_{m \text{ פעמים}}.$$

הוכיחו ש־ V^m ו־ $\mathcal{L}(\mathbf{F}^m, V)$ מרחבים וקטוריים איזומורפיים.

6 נניח ש־ v, x וקטורים ב־ V ו־ U, W תתי־מרחב של V כך ש־ $v + U = x + W$. הוכיחו ש־ $U = W$.

7 יהי $U = \{(x, y, z) \in \mathbf{R}^3 : 2x + 3y + 5z = 0\}$. נניח ש־ $A \subseteq \mathbf{R}^3$. הוכיחו ש־ A הזזה של U אם־ורק־אם קיים $c \in \mathbf{R}$ כך ש־

$$A = \{(x, y, z) \in \mathbf{R}^3 : 2x + 3y + 5z = c\}.$$

8 א. נניח ש־ $T \in \mathcal{L}(V, W)$ ו־ $c \in W$. הוכיחו ש־ $\{x \in V : Tx = c\}$ היא או הקבוצה הריקה או הזזה של $\operatorname{null} T$.

ב. הסבירו מדוע קבוצת הפתרונות של מערכת משוואות ליניאריות כמו זו שב־3.27 היא או הקבוצה הריקה או הזזה של תת־מרחב כלשהו של $\mathbf{F}^n$.

9 הוכיחו שתת־קבוצה שאינה ריקה A של V היא הזזה של תת־מרחב כלשהו של V אם־ורק־אם $\lambda v + (1 - \lambda)w \in A$ לכל $v, w \in A$ ולכל $\lambda \in \mathbf{F}$.

10 נניח ש־ $A_1 = v + U_1$ ו־ $A_2 = w + U_2$ עבור $v, w \in V$ כלשהם ו־ U_1, U_2 תתי־מרחב כלשהם של V. הוכיחו שהחיתוך $A_1 \cap A_2$ הוא או הזזה של תת־מרחב כלשהו של V או הקבוצה הריקה.

11 Suppose $U = \{(x_1, x_2, \dots) \in \mathbf{F}^\infty : x_k \neq 0 \text{ for only finitely many } k\}$.

(a) Show that U is a subspace of $\mathbf{F}^\infty$.

(b) Prove that $\mathbf{F}^\infty / U$ is infinite-dimensional.

12 Suppose $v_1, \dots, v_m \in V$. Let

$$A = \{\lambda_1 v_1 + \cdots + \lambda_m v_m : \lambda_1, \dots, \lambda_m \in \mathbf{F} \text{ and } \lambda_1 + \cdots + \lambda_m = 1\}.$$

(a) Prove that A is a translate of some subspace of V.

(b) Prove that if B is a translate of some subspace of V and $\{v_1, \dots, v_m\} \subseteq B$, then $A \subseteq B$.

(c) Prove that A is a translate of some subspace of V of dimension less than m.

13 Suppose U is a subspace of V such that V/U is finite-dimensional. Prove that V is isomorphic to $U \times (V/U)$.

14 Suppose U and W are subspaces of V and $V = U \oplus W$. Suppose $w_1, \dots, w_m$ is a basis of W. Prove that $w_1 + U, \dots, w_m + U$ is a basis of V/U.

15 Suppose U is a subspace of V and $v_1 + U, \dots, v_m + U$ is a basis of V/U and $u_1, \dots, u_n$ is a basis of U. Prove that $v_1, \dots, v_m, u_1, \dots, u_n$ is a basis of V.

16 Suppose $\varphi \in \mathcal{L}(V, \mathbf{F})$ and $\varphi \neq 0$. Prove that $\dim V/(\operatorname{null} \varphi) = 1$.

17 Suppose U is a subspace of V such that $\dim V/U = 1$. Prove that there exists $\varphi \in \mathcal{L}(V, \mathbf{F})$ such that $\operatorname{null} \varphi = U$.

18 Suppose that U is a subspace of V such that V/U is finite-dimensional.

(a) Show that if W is a finite-dimensional subspace of V and $V = U + W$, then $\dim W \geq \dim V/U$.

(b) Prove that there exists a finite-dimensional subspace W of V such that $\dim W = \dim V/U$ and $V = U \oplus W$.

19 Suppose $T \in \mathcal{L}(V, W)$ and U is a subspace of V. Let π denote the quotient map from V onto V/U. Prove that there exists $S \in \mathcal{L}(V/U, W)$ such that $T = S \circ \pi$ if and only if $U \subseteq \operatorname{null} T$.

11 נניח ש־ U היא קבוצת כל הסדרות $(x_1, x_2, \dots) \in \mathbf{F}^\infty$ המקיימות $x_k \neq 0$ רק עבור מספר סופי של ערכי k.

א. הראו ש־ U תת־מרחב של $\mathbf{F}^\infty$.

ב. הוכיחו ש־ $\mathbf{F}^\infty / U$ מממד אינסופי.

12 נניח ש־ $v_1, \dots, v_m \in V$. יהי

$$A = \{\lambda_1 v_1 + \cdots + \lambda_m v_m : \lambda_1, \dots, \lambda_m \in \mathbf{F} \text{ וגם } \lambda_1 + \cdots + \lambda_m = 1\}.$$

א. הוכיחו ש־ A הזזה של תת־מרחב כלשהו של V.

ב. הוכיחו שאם B הזזה של תת־מרחב כלשהו של V ו־ $\{v_1, \dots, v_m\} \subseteq B$, אזי $A \subseteq B$.

ג. הוכיחו ש־ A הזזה של תת־מרחב כלשהו של V מממד קטן מ־ m.

13 נניח ש־ U תת־מרחב של V כך ש־ V/U מממד סופי. הוכיחו ש־ V איזומורפי ל־ $U \times (V/U)$.

14 נניח ש־ U ו־ W תתי־מרחב של V ו־ $V = U \oplus W$. נניח ש־ $w_1, \dots, w_m$ בסיס של W. הוכיחו ש־ $w_1 + U, \dots, w_m + U$ בסיס של V/U.

15 נניח ש־ U תת־מרחב של V ו־ $v_1 + U, \dots, v_m + U$ בסיס של V/U ו־ $u_1, \dots, u_n$ בסיס של U. הוכיחו ש־ $v_1, \dots, v_m, u_1, \dots, u_n$ בסיס של V.

16 נניח ש־ $\varphi \in \mathcal{L}(V, \mathbf{F})$ ו־ $\varphi \neq 0$. הוכיחו ש־ $\dim V/(\operatorname{null} \varphi) = 1$.

17 נניח ש־ U תת־מרחב של V כך ש־ $\dim V/U = 1$. הוכיחו שקיים $\varphi \in \mathcal{L}(V, \mathbf{F})$ כך ש־ $\operatorname{null} \varphi = U$.

18 נניח ש־ U תת־מרחב של V כך ש־ V/U מממד סופי.

א. הראו שאם W תת־מרחב מממד סופי של V ו־ $V = U + W$, אזי $\dim W \geq \dim V/U$.

ב. הוכיחו שקיים תת־מרחב מממד סופי W של V כך ש־ $\dim W = \dim V/U$ וגם $V = U \oplus W$.

19 נניח ש־ $T \in \mathcal{L}(V, W)$ ו־ U תת־מרחב של V. תהי π העתקת המנה מ־ V על V/U. הוכיחו שקיים $S \in \mathcal{L}(V/U, W)$ כך ש־ $T = S \circ \pi$ אם־ורק־אם $U \subseteq \operatorname{null} T$.

אלגברה ליניארית בדרך הישר, מהדורה רביעית, מחבר: שלדון אקסלר
Linear Algebra Done Right, fourth edition, by Sheldon Axler

3F Duality

Dual Space and Dual Map

Linear maps into the scalar field $\mathbf{F}$ play a special role in linear algebra, and thus they get a special name.

3.108 definition: *linear functional*

A *linear functional* on V is a linear map from V to $\mathbf{F}$. In other words, a linear functional is an element of $\mathcal{L}(V, \mathbf{F})$.

3.109 example: *linear functionals*

- Define $\varphi\colon \mathbf{R}^3 \to \mathbf{R}$ by $\varphi(x, y, z) = 4x - 5y + 2z$. Then φ is a linear functional on $\mathbf{R}^3$.
- Fix $(c_1, \dots, c_n) \in \mathbf{F}^n$. Define $\varphi\colon \mathbf{F}^n \to \mathbf{F}$ by $\varphi(x_1, \dots, x_n) = c_1 x_1 + \cdots + c_n x_n$. Then φ is a linear functional on $\mathbf{F}^n$.
- Define $\varphi\colon \mathcal{P}(\mathbf{R}) \to \mathbf{R}$ by
$$\varphi(p) = 3p''(5) + 7p(4).$$
Then φ is a linear functional on $\mathcal{P}(\mathbf{R})$.
- Define $\varphi\colon \mathcal{P}(\mathbf{R}) \to \mathbf{R}$ by
$$\varphi(p) = \int_0^1 p$$
for each $p \in \mathcal{P}(\mathbf{R})$. Then φ is a linear functional on $\mathcal{P}(\mathbf{R})$.

The vector space $\mathcal{L}(V, \mathbf{F})$ also gets a special name and special notation.

3.110 definition: *dual space, V'*

The *dual space* of V, denoted by V', is the vector space of all linear functionals on V. In other words, $V' = \mathcal{L}(V, \mathbf{F})$.

3.111 $\dim V' = \dim V$

Suppose V is finite-dimensional. Then V' is also finite-dimensional and
$$\dim V' = \dim V.$$

Proof By 3.72 we have
$$\dim V' = \dim \mathcal{L}(V, \mathbf{F}) = (\dim V)(\dim \mathbf{F}) = \dim V,$$
as desired. ∎

6ג דואליות

מרחב דואלי והעתקה דואלית

להעתקות ליניאריות אל שדה הסקלרים $\mathbf{F}$ תפקיד מיוחד באלגברה ליניארית, ולפיכך מוענק להן שם מיוחד.

3.108 הגדרה: פונקציונל ליניארי

פונקציונל ליניארי מעל V הוא העתקה ליניארית מ־ V אל $\mathbf{F}$. במילים אחרות, פונקציונל ליניארי הוא איבר של $\mathcal{L}(V, \mathbf{F})$.

3.109 דוגמה: פונקציונלים ליניאריים

- נגדיר $\varphi : \mathbf{R}^3 \to \mathbf{R}$ על ידי $\varphi(x, y, z) = 4x - 5y + 2z$. אזי φ הוא פונקציונל ליניארי מעל $\mathbf{R}^3$.
- נקבע $(c_1, \dots, c_n) \in \mathbf{F}^n$. נגדיר $\varphi : \mathbf{F}^n \to \mathbf{F}$ על ידי $\varphi(x_1, \dots, x_n) = c_1 x_1 + \cdots + c_n x_n$. אזי φ הוא פונקציונל ליניארי מעל $\mathbf{F}^n$.
- נגדיר $\varphi : \mathcal{P}(\mathbf{R}) \to \mathbf{R}$ על ידי
$$\varphi(p) = 3p''(5) + 7p(4).$$
אזי φ הוא פונקציונל ליניארי מעל $\mathcal{P}(\mathbf{R})$.
- נגדיר $\varphi : \mathcal{P}(\mathbf{R}) \to \mathbf{R}$ על ידי
$$\varphi(p) = \int_0^1 p$$
לכל $p \in \mathcal{P}(\mathbf{R})$. אזי φ הוא פונקציונל ליניארי מעל $\mathcal{P}(\mathbf{R})$.

המרחב הווקטורי $\mathcal{L}(V, \mathbf{F})$ גם הוא מקבל שם וסימון מיוחדים.

3.110 הגדרה: מרחב דואלי, V'

המרחב הדואלי של V, המסומן ב־ V', הוא המרחב הווקטורי של כל הפונקציונלים הליניאריים מעל V. במילים אחרות, $V' = \mathcal{L}(V, \mathbf{F})$.

3.111 $\dim V' = \dim V$

נניח ש־ V מממד סופי. אזי V' גם הוא מממד סופי ו־
$$\dim V' = \dim V.$$

הוכחה לפי 3.72 מתקיים

$$\dim V' = \dim \mathcal{L}(V, \mathbf{F}) = (\dim V)(\dim \mathbf{F}) = \dim V,$$

כפי שרצינו.

In the following definition, the linear map lemma (3.4) implies that each φ_j is well defined.

3.112 definition: *dual basis*

If $v_1, \dots, v_n$ is a basis of V, then the *dual basis* of $v_1, \dots, v_n$ is the list $\varphi_1, \dots, \varphi_n$ of elements of V', where each φ_j is the linear functional on V such that

$$\varphi_j(v_k) = \begin{cases} 1 & \text{if } k = j, \\ 0 & \text{if } k \neq j. \end{cases}$$

3.113 example: *the dual basis of the standard basis of* $\mathbf{F}^n$

Suppose n is a positive integer. For $1 \le j \le n$, define φ_j to be the linear functional on $\mathbf{F}^n$ that selects the jth coordinate of a vector in $\mathbf{F}^n$. Thus

$$\varphi_j(x_1, \dots, x_n) = x_j$$

for each $(x_1, \dots, x_n) \in \mathbf{F}^n$.

Let $e_1, \dots, e_n$ be the standard basis of $\mathbf{F}^n$. Then

$$\varphi_j(e_k) = \begin{cases} 1 & \text{if } k = j, \\ 0 & \text{if } k \neq j. \end{cases}$$

Thus $\varphi_1, \dots, \varphi_n$ is the dual basis of the standard basis $e_1, \dots, e_n$ of $\mathbf{F}^n$.

The next result shows that the dual basis of a basis of V consists of the linear functionals on V that give the coefficients for expressing a vector in V as a linear combination of the basis vectors.

3.114 *dual basis gives coefficients for linear combination*

Suppose $v_1, \dots, v_n$ is a basis of V and $\varphi_1, \dots, \varphi_n$ is the dual basis. Then

$$v = \varphi_1(v)v_1 + \cdots + \varphi_n(v)v_n$$

for each $v \in V$.

Proof Suppose $v \in V$. Then there exist $c_1, \dots, c_n \in \mathbf{F}$ such that

3.115 $$v = c_1v_1 + \cdots + c_nv_n.$$

If $j \in \{1, \dots, n\}$, then applying φ_j to both sides of the equation above gives

$$\varphi_j(v) = c_j.$$

Substituting the values for $c_1, \dots, c_n$ given by the equation above into 3.115 shows that $v = \varphi_1(v)v_1 + \cdots + \varphi_n(v)v_n$.

מלמת ההעתקה הליניארית (3.4) נובע כי בהגדרה להלן, כל φ_j מוגדר היטב.

3.112 הגדרה: בסיס דואלי

אם $v_1, \dots, v_n$ בסיס של V, אזי הבסיס הדואלי של $v_1, \dots, v_n$ הוא הרשימה $\varphi_1, \dots, \varphi_n$ של איברי V', כאשר כל φ_j הוא הפונקציונל הליניארי מעל V כך ש־

$$\varphi_j(v_k) = \begin{cases} 1 & \text{if } k = j, \\ 0 & \text{if } k \neq j. \end{cases}$$

3.113 דוגמה: הבסיס הדואלי של הבסיס הסטנדרטי של $\mathbf{F}^n$

נניח ש־ n מספר שלם חיובי. עבור $1 \le j \le n$, נגדיר את φ_j כפונקציונל הליניארי מעל $\mathbf{F}^n$ הבוחר את הקואורדינטה ה־ j^{th} של וקטור ב־ $\mathbf{F}^n$. לפיכך

$$\varphi_j(x_1, \dots, x_n) = x_j$$

לכל $(x_1, \dots, x_n) \in \mathbf{F}^n$.

יהיו $e_1, \dots, e_n$ הבסיס הסטנדרטי של $\mathbf{F}^n$. אזי

$$\varphi_j(e_k) = \begin{cases} 1 & \text{if } k = j, \\ 0 & \text{if } k \neq j. \end{cases}$$

לפיכך $\varphi_1, \dots, \varphi_n$ הוא הבסיס הדואלי של הבסיס הסטנדרטי $e_1, \dots, e_n$ של $\mathbf{F}^n$.

המסקנה הבאה מראה כי בסיס דואלי של בסיס של V מורכב מהפונקציונלים הליניאריים מעל V המספקים את המקדמים לביטוי וקטור ב־ V כצירוף ליניארי של וקטורי הבסיס.

3.114 בסיס דואלי מספק מקדמים עבור צירוף ליניארי

נניח ש־ $v_1, \dots, v_n$ בסיס של V ו־ $\varphi_1, \dots, \varphi_n$ הוא הבסיס הדואלי. אזי

$$v = \varphi_1(v)v_1 + \cdots + \varphi_n(v)v_n$$

לכל $v \in V$.

הוכחה נניח ש־ $v \in V$. אזי קיימים $c_1, \dots, c_n \in \mathbf{F}$ כך ש־

$$v = c_1v_1 + \cdots + c_nv_n. \tag{3.115}$$

אם $j \in \{1, \dots, n\}$, אזי הפעלת φ_j על שני אגפי המשוואה לעיל נותנת

$$\varphi_j(v) = c_j.$$

נציב את הערכים של $c_1, \dots, c_n$ שקיבלנו מהמשוואה לעיל בתוך 3.115 ונקבל ש־ $v = \varphi_1(v)v_1 + \cdots + \varphi_n(v)v_n$. ■

The next result shows that the dual basis is indeed a basis of the dual space. Thus the terminology "dual basis" is justified.

3.116 *dual basis is a basis of the dual space*

Suppose V is finite-dimensional. Then the dual basis of a basis of V is a basis of V'.

Proof Suppose $v_1, \dots, v_n$ is a basis of V. Let $\varphi_1, \dots, \varphi_n$ denote the dual basis.

To show that $\varphi_1, \dots, \varphi_n$ is a linearly independent list of elements of V', suppose $a_1, \dots, a_n \in \mathbf{F}$ are such that

3.117 $$a_1\varphi_1 + \cdots + a_n\varphi_n = 0.$$

Now

$$(a_1\varphi_1 + \cdots + a_n\varphi_n)(v_k) = a_k$$

for each $k = 1, \dots, n$. Thus 3.117 shows that $a_1 = \cdots = a_n = 0$. Hence $\varphi_1, \dots, \varphi_n$ is linearly independent.

Because $\varphi_1, \dots, \varphi_n$ is a linearly independent list in V' whose length equals $\dim V'$ (by 3.111), we can conclude that $\varphi_1, \dots, \varphi_n$ is a basis of V' (see 2.38). ∎

In the definition below, note that if T is a linear map from V to W then T' is a linear map from W' to V'.

3.118 definition: *dual map, T'*

Suppose $T \in \mathcal{L}(V, W)$. The *dual map* of T is the linear map $T' \in \mathcal{L}(W', V')$ defined for each $\varphi \in W'$ by

$$T'(\varphi) = \varphi \circ T.$$

If $T \in \mathcal{L}(V, W)$ and $\varphi \in W'$, then $T'(\varphi)$ is defined above to be the composition of the linear maps φ and T. Thus $T'(\varphi)$ is indeed a linear map from V to $\mathbf{F}$; in other words, $T'(\varphi) \in V'$.

The following two bullet points show that T' is a linear map from W' to V'.

- If $\varphi, \psi \in W'$, then

$$T'(\varphi + \psi) = (\varphi + \psi) \circ T = \varphi \circ T + \psi \circ T = T'(\varphi) + T'(\psi).$$

- If $\lambda \in \mathbf{F}$ and $\varphi \in W'$, then

$$T'(\lambda\varphi) = (\lambda\varphi) \circ T = \lambda(\varphi \circ T) = \lambda T'(\varphi).$$

The prime notation appears with two unrelated meanings in the next example: D' denotes the dual of the linear map D, and p' denotes the derivative of a polynomial p.

המסקנה הבאה מראה כי הבסיס הדואלי הוא אכן בסיס למרחב הדואלי. לפיכך המונח "בסיס דואלי" מוצדק.

3.116 בסיס דואלי הוא בסיס של המרחב הדואלי

נניח ש־ V מממד סופי. אזי הבסיס הדואלי של בסיס של V הוא בסיס של V'.

הוכחה נניח ש־ $v_1, \dots, v_n$ בסיס של V. נסמן ב־ $\varphi_1, \dots, \varphi_n$ את הבסיס הדואלי.

כדי להראות ש־ $\varphi_1, \dots, \varphi_n$ היא רשימת איברים של V' בלתי תלויה לינארית, נניח ש־ $a_1, \dots, a_n \in \mathbf{F}$ כך ש־

$$a_1\varphi_1 + \cdots + a_n\varphi_n = 0. \tag{3.117}$$

כעת

$$(a_1\varphi_1 + \cdots + a_n\varphi_n)(v_k) = a_k$$

לכל $k = 1, \dots, n$. לפיכך 3.117 מראה לנו ש־ $a_1 = \cdots = a_n = 0$. לכן $\varphi_1, \dots, \varphi_n$ בלתי תלויה לינארית.

מאחר ש־ $\varphi_1, \dots, \varphi_n$ רשימה בלתי תלויה לינארית ב־ V' שאורכה שווה ל־ $\dim V'$ (לפי 3.111), אנו מסיקים כי $\varphi_1, \dots, \varphi_n$ בסיס של V' (ראו 2.38). ∎

נשים לב כי בהגדרה הבאה, אם T העתקה לינארית מ־ V אל W אזי T' העתקה לינארית מ־ W' אל V'.

3.118 הגדרה: העתקה דואלית, T'

נניח ש־ $T \in \mathcal{L}(V, W)$. ההעתקה הדואלית של T היא ההעתקה הלינארית $T' \in \mathcal{L}(W', V')$ המוגדרת לכל $\varphi \in W'$ על ידי

$$T'(\varphi) = \varphi \circ T.$$

אם $T \in \mathcal{L}(V, W)$ ו־ $\varphi \in W'$, אזי $T'(\varphi)$ מוגדרת לעיל כהרכבה של ההעתקות הלינאריות φ ו־ T. לפיכך $T'(\varphi)$ היא אכן העתקה לינארית מ־ V אל $\mathbf{F}$; במילים אחרות, $T'(\varphi) \in V'$.

שתי הנקודות הבאות מראות כי T' העתקה לינארית מ־ W' אל V'.

- אם $\varphi, \psi \in W'$, אזי
$$T'(\varphi + \psi) = (\varphi + \psi) \circ T = \varphi \circ T + \psi \circ T = T'(\varphi) + T'(\psi).$$
- אם $\lambda \in \mathbf{F}$ ו־ $\varphi \in W'$, אזי
$$T'(\lambda\varphi) = (\lambda\varphi) \circ T = \lambda(\varphi \circ T) = \lambda T'(\varphi).$$

הסימון באמצעות תג מופיע בדוגמה הבאה בשתי משמעויות בלתי קשורות: D' מסמן את ההעתקה הדואלית של ההעתקה הלינארית D, ואילו p' מסמן את הנגזרת של הפולינום p.

3.119 example: *dual map of the differentiation linear map*

Define $D: \mathcal{P}(\mathbf{R}) \to \mathcal{P}(\mathbf{R})$ by $Dp = p'$.

- Suppose φ is the linear functional on $\mathcal{P}(\mathbf{R})$ defined by $\varphi(p) = p(3)$. Then $D'(\varphi)$ is the linear functional on $\mathcal{P}(\mathbf{R})$ given by

$$(D'(\varphi))(p) = (\varphi \circ D)(p) = \varphi(Dp) = \varphi(p') = p'(3).$$

Thus $D'(\varphi)$ is the linear functional on $\mathcal{P}(\mathbf{R})$ taking p to $p'(3)$.

- Suppose φ is the linear functional on $\mathcal{P}(\mathbf{R})$ defined by $\varphi(p) = \int_0^1 p$. Then $D'(\varphi)$ is the linear functional on $\mathcal{P}(\mathbf{R})$ given by

$$\begin{aligned}(D'(\varphi))(p) &= (\varphi \circ D)(p)\\ &= \varphi(Dp)\\ &= \varphi(p')\\ &= \int_0^1 p'\\ &= p(1) - p(0).\end{aligned}$$

Thus $D'(\varphi)$ is the linear functional on $\mathcal{P}(\mathbf{R})$ taking p to $p(1) - p(0)$.

In the next result, (a) and (b) imply that the function that takes T to T' is a linear map from $\mathcal{L}(V, W)$ to $\mathcal{L}(W', V')$.

In (c) below, note the reversal of order from ST on the left to $T'S'$ on the right.

3.120 *algebraic properties of dual maps*

Suppose $T \in \mathcal{L}(V, W)$. Then

(a) $(S + T)' = S' + T'$ for all $S \in \mathcal{L}(V, W)$;

(b) $(\lambda T)' = \lambda T'$ for all $\lambda \in \mathbf{F}$;

(c) $(ST)' = T'S'$ for all $S \in \mathcal{L}(W, U)$.

Proof The proofs of (a) and (b) are left to the reader.

To prove (c), suppose $\varphi \in U'$. Then

$$(ST)'(\varphi) = \varphi \circ (ST) = (\varphi \circ S) \circ T = T'(\varphi \circ S) = T'(S'(\varphi)) = (T'S')(\varphi),$$

where the first, third, and fourth equalities above hold because of the definition of the dual map, the second equality holds because composition of functions is associative, and the last equality follows from the definition of composition.

The equation above shows that $(ST)'(\varphi) = (T'S')(\varphi)$ for all $\varphi \in U'$. Thus $(ST)' = T'S'$. ∎

Some books use the notation V^* and T^* for duality instead of V' and T'. However, here we reserve the notation T^* for the adjoint, which will be introduced when we study linear maps on inner product spaces in Chapter 7.

3.119 דוגמה: העתקה דואלית של העתקת הגזירה

נגדיר $D: \mathcal{P}(\mathbf{R}) \to \mathcal{P}(\mathbf{R})$ על ידי $Dp = p'$.

- נניח ש־ φ הוא הפונקציונל הליניארי מעל $\mathcal{P}(\mathbf{R})$ המוגדר על ידי $\varphi(p) = p(3)$. אזי $D'(\varphi)$ הוא הפונקציונל הליניארי מעל $\mathcal{P}(\mathbf{R})$ הנתון על ידי
$$(D'(\varphi))(p) = (\varphi \circ D)(p) = \varphi(Dp) = \varphi(p') = p'(3).$$
לפיכך $D'(\varphi)$ הוא הפונקציונל הליניארי מעל $\mathcal{P}(\mathbf{R})$ המעתיק את p אל $p'(3)$.
- נניח ש־ φ הוא הפונקציונל הליניארי מעל $\mathcal{P}(\mathbf{R})$ המוגדר על ידי $\varphi(p) = \int_0^1 p$. אזי $D'(\varphi)$ הוא הפונקציונל הליניארי מעל $\mathcal{P}(\mathbf{R})$ הנתון על ידי
$$\begin{aligned}(D'(\varphi))(p) &= (\varphi \circ D)(p)\\ &= \varphi(Dp)\\ &= \varphi(p')\\ &= \int_0^1 p'\\ &= p(1) - p(0).\end{aligned}$$
לפיכך $D'(\varphi)$ הוא הפונקציונל הליניארי מעל $\mathcal{P}(\mathbf{R})$ המעתיק את p אל $p(1) - p(0)$.

במסקנה הבאה מסעיפים (א) ו־ (ב) נובע כי הפונקציה המעתיקה את T אל T' היא העתקה ליניארית מ־ $\mathcal{L}(V, W)$ אל $\mathcal{L}(W', V')$.
שימו לב להיפוך הסדר בסעיף (ג) מ־ ST באגף השמאלי ל־ $T'S'$ באגף הימני.

3.120 תכונות אלגבריות של העתקות דואליות

נניח ש־ $T \in \mathcal{L}(V, W)$. אזי

א. $(S + T)' = S' + T'$ לכל $S \in \mathcal{L}(V, W)$

ב. $(\lambda T)' = \lambda T'$ לכל $\lambda \in \mathbf{F}$

ג. $(ST)' = T'S'$ לכל $S \in \mathcal{L}(W, U)$.

הוכחה הוכחות (א) ו־ (ב) מושארות לקוראים.

להוכחת (ג), נניח ש־ $\varphi \in U'$. אזי
$$(ST)'(\varphi) = \varphi \circ (ST) = (\varphi \circ S) \circ T = T'(\varphi \circ S) = T'(S'(\varphi)) = (T'S')(\varphi),$$
כאשר השוויונות הראשון, השלישי, והרביעי לעיל מתקיימים לאור הגדרת ההעתקה הדואלית, השוויון השני מתקיים מאחר שפעולת הרכבת פונקציות היא אסוציאטיבית, והשוויון האחרון נובע מהגדרת ההרכבה.

בחלק מהספרים דואליות מסומנת ב־ V^* ו־ T^* במקום V' ו־ T'. אולם, בספר זה הסימון T^* שמור להעתקה הצמודה שאותה נכיר כשנלמד על העתקות מעל מרחבי מכפלה פנימית בפרק 7.

המשוואה לעיל מראה כי $(ST)'(\varphi) = (T'S')(\varphi)$ לכל $\varphi \in U'$. לפיכך $(ST)' = T'S'$.

Null Space and Range of Dual of Linear Map

Our goal in this subsection is to describe null T' and range T' in terms of range T and null T. To do this, we will need the next definition.

3.121 definition: *annihilator*, U^0

For $U \subseteq V$, the *annihilator* of U, denoted by U^0, is defined by

$$U^0 = \{\varphi \in V' : \varphi(u) = 0 \text{ for all } u \in U\}.$$

3.122 example: *element of an annihilator*

Suppose U is the subspace of $\mathcal{P}(\mathbf{R})$ consisting of polynomial multiples of x^2. If φ is the linear functional on $\mathcal{P}(\mathbf{R})$ defined by $\varphi(p) = p'(0)$, then $\varphi \in U^0$.

For $U \subseteq V$, the annihilator U^0 is a subset of the dual space V'. Thus U^0 depends on the vector space containing U, so a notation such as U^0_V would be more precise. However, the containing vector space will always be clear from the context, so we will use the simpler notation U^0.

3.123 example: *the annihilator of a two-dimensional subspace of* $\mathbf{R}^5$

Let e_1, e_2, e_3, e_4, e_5 denote the standard basis of $\mathbf{R}^5$; let $\varphi_1, \varphi_2, \varphi_3, \varphi_4, \varphi_5 \in (\mathbf{R}^5)'$ denote the dual basis of e_1, e_2, e_3, e_4, e_5. Suppose

$$U = \operatorname{span}(e_1, e_2) = \{(x_1, x_2, 0, 0, 0) \in \mathbf{R}^5 : x_1, x_2 \in \mathbf{R}\}.$$

We want to show that $U^0 = \operatorname{span}(\varphi_3, \varphi_4, \varphi_5)$.

Recall (see 3.113) that φ_j is the linear functional on $\mathbf{R}^5$ that selects the jth coordinate: $\varphi_j(x_1, x_2, x_3, x_4, x_5) = x_j$.

First suppose $\varphi \in \operatorname{span}(\varphi_3, \varphi_4, \varphi_5)$. Then there exist $c_3, c_4, c_5 \in \mathbf{R}$ such that $\varphi = c_3\varphi_3 + c_4\varphi_4 + c_5\varphi_5$. If $(x_1, x_2, 0, 0, 0) \in U$, then

$$\varphi(x_1, x_2, 0, 0, 0) = (c_3\varphi_3 + c_4\varphi_4 + c_5\varphi_5)(x_1, x_2, 0, 0, 0) = 0.$$

Thus $\varphi \in U^0$. Hence we have shown that $\operatorname{span}(\varphi_3, \varphi_4, \varphi_5) \subseteq U^0$.

To show the inclusion in the other direction, suppose that $\varphi \in U^0$. Because the dual basis is a basis of $(\mathbf{R}^5)'$, there exist $c_1, c_2, c_3, c_4, c_5 \in \mathbf{R}$ such that $\varphi = c_1\varphi_1 + c_2\varphi_2 + c_3\varphi_3 + c_4\varphi_4 + c_5\varphi_5$. Because $e_1 \in U$ and $\varphi \in U^0$, we have

$$0 = \varphi(e_1) = (c_1\varphi_1 + c_2\varphi_2 + c_3\varphi_3 + c_4\varphi_4 + c_5\varphi_5)(e_1) = c_1.$$

Similarly, $e_2 \in U$ and thus $c_2 = 0$. Hence $\varphi = c_3\varphi_3 + c_4\varphi_4 + c_5\varphi_5$. Thus $\varphi \in \operatorname{span}(\varphi_3, \varphi_4, \varphi_5)$, which shows that $U^0 \subseteq \operatorname{span}(\varphi_3, \varphi_4, \varphi_5)$.

Thus $U^0 = \operatorname{span}(\varphi_3, \varphi_4, \varphi_5)$.

מרחב האפס והטווח של העתקה דואלית

מטרתנו בתת־יחידה זו היא לתאר את $\operatorname{null} T'$ ו־ $\operatorname{range} T'$ במונחים של $\operatorname{range} T$ ו־ $\operatorname{null} T$. לצורך כך נזדקק להגדרה הבאה.

3.121 הגדרה: מְאַפֵּס, U^0

עבור $U \subseteq V$, הַמְאַפֵּס של U, המסומן ב־ U^0, מוגדר על ידי
$$U^0 = \{\varphi \in V' : \varphi(u) = 0 \text{ לכל } u \in U\}.$$

3.122 דוגמה: איבר במאפס

נניח ש־ U תת־מרחב של $\mathcal{P}(\mathbf{R})$ המורכב מכפולות פולינומיות של x^2. אם φ הוא הפונקציונל הליניארי מעל $\mathcal{P}(\mathbf{R})$ המוגדר על ידי $\varphi(p) = p'(0)$, אזי $\varphi \in U^0$.

עבור $U \subseteq V$, הַמְאַפֵּס U^0 הוא תת־קבוצה של המרחב הדואלי V'. לפיכך U^0 תלוי במרחב הווקטורי המכיל את U, ולכן סימון כגון U^0_V יהיה מדויק יותר. אולם, מאחר שהמרחב הווקטורי המכיל תמיד יהיה ברור מהקשר, אנו נשתמש בסימון הפשוט יותר U^0.

3.123 דוגמה: המאפס של תת־מרחב דו־ממדי של $\mathbf{R}^5$

נסמן ב־ e_1, e_2, e_3, e_4, e_5 את הבסיס הסטנדרטי של $\mathbf{R}^5$; יהיו $\varphi_1, \varphi_2, \varphi_3, \varphi_4, \varphi_5 \in (\mathbf{R}^5)'$ הבסיס הדואלי של e_1, e_2, e_3, e_4, e_5. נניח ש־
$$U = \operatorname{span}(e_1, e_2) = \{(x_1, x_2, 0, 0, 0) \in \mathbf{R}^5 : x_1, x_2 \in \mathbf{R}\}.$$
ברצוננו להראות כי $U^0 = \operatorname{span}(\varphi_3, \varphi_4, \varphi_5)$.

נזכיר (ראו 3.113) כי φ_j הוא הפונקציונל הליניארי מעל $\mathbf{R}^5$ הבוחר את הקואורדינטה ה־ j^{th} באופן הבא: $\varphi_j(x_1, x_2, x_3, x_4, x_5) = x_j$.

תחילה נניח ש־ $\varphi \in \operatorname{span}(\varphi_3, \varphi_4, \varphi_5)$. אזי קיימים $c_3, c_4, c_5 \in \mathbf{R}$ כך ש־ $\varphi = c_3\varphi_3 + c_4\varphi_4 + c_5\varphi_5$. אם $(x_1, x_2, 0, 0, 0) \in U$, אזי
$$\varphi(x_1, x_2, 0, 0, 0) = (c_3\varphi_3 + c_4\varphi_4 + c_5\varphi_5)(x_1, x_2, 0, 0, 0) = 0.$$
לפיכך $\varphi \in U^0$. לכן הראינו כי $\operatorname{span}(\varphi_3, \varphi_4, \varphi_5) \subseteq U^0$.

כעת נראה את ההכלה בכיוון ההפוך.

נניח כי $\varphi \in U^0$. מאחר שהבסיס הדואלי הוא בסיס של $(\mathbf{R}^5)'$, קיימים סקלרים $c_1, c_2, c_3, c_4, c_5 \in \mathbf{R}$ כך ש־ $\varphi = c_1\varphi_1 + c_2\varphi_2 + c_3\varphi_3 + c_4\varphi_4 + c_5\varphi_5$. מאחר ש־ $e_1 \in U$ ו־ $\varphi \in U^0$, נקבל
$$0 = \varphi(e_1) = (c_1\varphi_1 + c_2\varphi_2 + c_3\varphi_3 + c_4\varphi_4 + c_5\varphi_5)(e_1) = c_1.$$
באופן דומה, $e_2 \in U$ ולפיכך $c_2 = 0$. לכן $\varphi = c_3\varphi_3 + c_4\varphi_4 + c_5\varphi_5$. לפיכך $\varphi \in \operatorname{span}(\varphi_3, \varphi_4, \varphi_5)$, מה שמראה כי $U^0 \subseteq \operatorname{span}(\varphi_3, \varphi_4, \varphi_5)$.

לפיכך $U^0 = \operatorname{span}(\varphi_3, \varphi_4, \varphi_5)$.

3.124 *the annihilator is a subspace*

Suppose $U \subseteq V$. Then U^0 is a subspace of V'.

Proof Note that $0 \in U^0$ (here 0 is the zero linear functional on V) because the zero linear functional applied to every vector in U equals $0 \in \mathbf{F}$.

Suppose $\varphi, \psi \in U^0$. Thus $\varphi, \psi \in V'$ and $\varphi(u) = \psi(u) = 0$ for every $u \in U$. If $u \in U$, then

$$(\varphi + \psi)(u) = \varphi(u) + \psi(u) = 0 + 0 = 0.$$

Thus $\varphi + \psi \in U^0$.

Similarly, U^0 is closed under scalar multiplication. Thus 1.34 implies that U^0 is a subspace of V'. ∎

The next result shows that $\dim U^0$ is the difference of $\dim V$ and $\dim U$. For example, this shows that if U is a two-dimensional subspace of $\mathbf{R}^5$, then U^0 is a three-dimensional subspace of $(\mathbf{R}^5)'$, as in Example 3.123.

The next result can be proved following the pattern of Example 3.123: choose a basis $u_1, \dots, u_m$ of U, extend to a basis $u_1, \dots, u_m, \dots, u_n$ of V, let $\varphi_1, \dots, \varphi_m, \dots, \varphi_n$ be the dual basis of V', and then show that $\varphi_{m+1}, \dots, \varphi_n$ is a basis of U^0, which implies the desired result. You should construct the proof just outlined, even though a slicker proof is presented here.

3.125 *dimension of the annihilator*

Suppose V is finite-dimensional and U is a subspace of V. Then

$$\dim U^0 = \dim V - \dim U.$$

Proof Let $i \in \mathcal{L}(U, V)$ be the inclusion map defined by $i(u) = u$ for each $u \in U$. Thus i' is a linear map from V' to U'. The fundamental theorem of linear maps (3.21) applied to i' shows that

$$\dim \operatorname{range} i' + \dim \operatorname{null} i' = \dim V'.$$

However, $\operatorname{null} i' = U^0$ (as can be seen by thinking about the definitions) and $\dim V' = \dim V$ (by 3.111), so we can rewrite the equation above as

3.126 $$\dim \operatorname{range} i' + \dim U^0 = \dim V.$$

If $\varphi \in U'$, then φ can be extended to a linear functional ψ on V (see, for example, Exercise 13 in Section 3A). The definition of i' shows that $i'(\psi) = \varphi$. Thus $\varphi \in \operatorname{range} i'$, which implies that $\operatorname{range} i' = U'$. Hence

$$\dim \operatorname{range} i' = \dim U' = \dim U,$$

and then 3.126 becomes the equation $\dim U + \dim U^0 = \dim V$, as desired. ∎

3.124 המאפס הוא תת־מרחב

נניח ש־ $U \subseteq V$. אזי U^0 תת־מרחב של V'.

הוכחה נשים לב כי $0 \in U^0$ (כאן ה־ 0 הוא הפונקציונל הליניארי האפס מעל V) מאחר שהפעלת פונקציונל האפס על כל וקטור ב־ U תוצאתה $0 \in \mathbf{F}$.

נניח ש־ $\varphi, \psi \in U^0$. לפיכך $\varphi, \psi \in V'$ ו־ $\varphi(u) = \psi(u) = 0$ לכל $u \in U$. אם $u \in U$, אזי

$$(\varphi + \psi)(u) = \varphi(u) + \psi(u) = 0 + 0 = 0.$$

לפיכך $\varphi + \psi \in U^0$.

באופן דומה, U^0 סגור תחת פעולת כפל בסקלר. לפיכך 1.34 גורר ש־ U^0 תת־מרחב של V'. ■

במסקנה הבאה נראה ש־ $\dim U^0$ הוא ההפרש בין $\dim V$ ל־ $\dim U$. לדוגמה, המסקנה מראה לנו כי אם U תת־מרחב דו־ממדי של $\mathbf{R}^5$, אזי U^0 תת־מרחב תלת־ממדי של $(\mathbf{R}^5)'$, כפי שראינו בדוגמה 3.123.

את המסקנה הבאה אפשר להוכיח אם נלך בעקבות ההוכחה בדוגמה 3.123: נבחר בסיס $u_1, \dots, u_m$ של U, נשלים לבסיס $u_1, \dots, u_m, \dots, u_n$ של V, יהי $\varphi_1, \dots, \varphi_m, \dots, \varphi_n$ הבסיס הדואלי של V', ובנקודה זו נוכל להראות ש־ $\varphi_{m+1}, \dots, \varphi_n$ בסיס של U^0, מה שיוביל אל התוצאה הרצויה. עליכם לבנות את ההוכחה במתווה שתואר, זאת אף שאנו מציגים להלן הוכחה אלגנטית יותר.

3.125 ממד המאפס

נניח ש־ V מממד סופי ו־ U תת־מרחב של V. אזי

$$\dim U^0 = \dim V - \dim U.$$

הוכחה תהי $i \in \mathcal{L}(U, V)$ העתקת ההכלה המוגדרת על ידי $i(u) = u$ לכל $u \in U$. לפיכך i' העתקה ליניארית מ־ V' אל U'. נפעיל את המשפט היסודי של ההעתקות הליניאריות (3.21) על i' ונקבל

$$\dim \operatorname{range} i' + \dim \operatorname{null} i' = \dim V'.$$

אולם $\operatorname{null} i' = U^0$ (כפי שניתן לראות כשבוחנים את ההגדרות) ו־ $\dim V' = \dim V$ (לפי 3.111), כך שנוכל לכתוב מחדש את המשוואה לעיל באופן הבא

$$\dim \operatorname{range} i' + \dim U^0 = \dim V. \tag{3.126}$$

אם $\varphi \in U'$, אזי נוכל להרחיב את הגדרת φ לפונקציונל ליניארי ψ מעל V (ראו, לדוגמה, תרגיל 13 ביחידה ג1). מהגדרת i' נובע כי $i'(\psi) = \varphi$. לפיכך $\varphi \in \operatorname{range} i'$, מה שמוביל למסקנה כי $\operatorname{range} i' = U'$. לכן

$$\dim \operatorname{range} i' = \dim U' = \dim U,$$

בשלב זה 3.126 הופכת למשוואה $\dim U + \dim U^0 = \dim V$, כפי שרצינו. ■

The next result can be a useful tool to show that a subspace is as big as possible—see (a)—or to show that a subspace is as small as possible—see (b).

3.127 *condition for the annihilator to equal* $\{0\}$ *or the whole space*

Suppose V is finite-dimensional and U is a subspace of V. Then

(a) $U^0 = \{0\} \iff U = V$;

(b) $U^0 = V' \iff U = \{0\}$.

Proof To prove (a), we have

$$\begin{aligned} U^0 = \{0\} &\iff \dim U^0 = 0 \\ &\iff \dim U = \dim V \\ &\iff U = V, \end{aligned}$$

where the second equivalence follows from 3.125 and the third equivalence follows from 2.39.

Similarly, to prove (b) we have

$$\begin{aligned} U^0 = V' &\iff \dim U^0 = \dim V' \\ &\iff \dim U^0 = \dim V \\ &\iff \dim U = 0 \\ &\iff U = \{0\}, \end{aligned}$$

where one direction of the first equivalence follows from 2.39, the second equivalence follows from 3.111, and the third equivalence follows from 3.125. ∎

The proof of (a) in the next result does not use the hypothesis that V and W are finite-dimensional.

3.128 *the null space of* T'

Suppose V and W are finite-dimensional and $T \in \mathcal{L}(V, W)$. Then

(a) $\operatorname{null} T' = (\operatorname{range} T)^0$;

(b) $\dim \operatorname{null} T' = \dim \operatorname{null} T + \dim W - \dim V$.

Proof

(a) First suppose $\varphi \in \operatorname{null} T'$. Thus $0 = T'(\varphi) = \varphi \circ T$. Hence

$$0 = (\varphi \circ T)(v) = \varphi(Tv) \quad \text{for every } v \in V.$$

Thus $\varphi \in (\operatorname{range} T)^0$. This implies that $\operatorname{null} T' \subseteq (\operatorname{range} T)^0$.

To prove the inclusion in the opposite direction, now suppose $\varphi \in (\operatorname{range} T)^0$. Thus $\varphi(Tv) = 0$ for every vector $v \in V$. Hence $0 = \varphi \circ T = T'(\varphi)$. In other words, $\varphi \in \operatorname{null} T'$, which shows that $(\operatorname{range} T)^0 \subseteq \operatorname{null} T'$, completing the proof of (a).

המסקנה הבאה יכולה לשמש ככלי שבאמצעותו ניתן להראות שגודלו של תת־מרחב מקסימלי—ראו (א)—או להראות שגודלו של תת־מרחב מינימלי—ראו (ב).

3.127 תנאי לכך שהמאפס שווה ל־ $\{0\}$ או למרחב כולו

נניח ש־ V מממד סופי ו־ U תת־מרחב של V. אזי

א. $U^0 = \{0\}$ אם־ורק־אם $U = V$

ב. $U^0 = V'$ אם־ורק־אם $U = \{0\}$.

הוכחה להוכחת (א), מתקיים

$$\begin{aligned} U^0 = \{0\} &\iff \dim U^0 = 0 \\ &\iff \dim U = \dim V \\ &\iff U = V, \end{aligned}$$

כאשר השקילות השנייה נובעת מ־ 3.125 והשקילות השלישית נובעת מ־ 2.39.

באופן דומה, להוכחת (ב) מתקיים

$$\begin{aligned} U^0 = V' &\iff \dim U^0 = \dim V' \\ &\iff \dim U^0 = \dim V \\ &\iff \dim U = 0 \\ &\iff U = \{0\}, \end{aligned}$$

כאשר כיוון אחד של השקילות הראשונה נובע מ־ 2.39, השקילות השנייה נובעת מ־ 3.111, והשקילות השלישית נובעת מ־ 3.125. ■

ההוכחה של (א) במסקנה הבאה אינה משתמשת בהנחה ש־ V ו־ W מממד סופי.

3.128 מרחב האפס של T'

נניח ש־ V ו־ W מממד סופי ו־ $T \in \mathcal{L}(V, W)$. אזי

א. $\operatorname{null} T' = (\operatorname{range} T)^0$

ב. $\dim \operatorname{null} T' = \dim \operatorname{null} T + \dim W - \dim V$.

הוכחה

א. ראשית נניח כי $\varphi \in \operatorname{null} T'$. לפיכך $0 = T'(\varphi) = \varphi \circ T$. לכן
$0 = (\varphi \circ T)(v) = \varphi(Tv)$ לכל $v \in V$.
לפיכך $\varphi \in (\operatorname{range} T)^0$. מכאן נובע ש־ $\operatorname{null} T' \subseteq (\operatorname{range} T)^0$.

להוכחת ההכלה בכיוון ההפוך, נניח כעת ש־ $\varphi \in (\operatorname{range} T)^0$. לפיכך $\varphi(Tv) = 0$ לכל וקטור $v \in V$. לכן $0 = \varphi \circ T = T'(\varphi)$. במילים אחרות, $\varphi \in \operatorname{null} T'$, מה שמראה לנו כי $(\operatorname{range} T)^0 \subseteq \operatorname{null} T'$, ובזאת השלמנו את הוכחת (א).

(b) We have

$$\begin{aligned}\dim \operatorname{null} T' &= \dim(\operatorname{range} T)^0 \\ &= \dim W - \dim \operatorname{range} T \\ &= \dim W - (\dim V - \dim \operatorname{null} T) \\ &= \dim \operatorname{null} T + \dim W - \dim V,\end{aligned}$$

where the first equality comes from (a), the second equality comes from 3.125, and the third equality comes from the fundamental theorem of linear maps (3.21). ∎

The next result can be useful because sometimes it is easier to verify that T' is injective than to show directly that T is surjective.

3.129 *T surjective is equivalent to T' injective*

Suppose V and W are finite-dimensional and $T \in \mathcal{L}(V, W)$. Then

$$T \text{ is surjective} \iff T' \text{ is injective.}$$

Proof We have

$$\begin{aligned}T \in \mathcal{L}(V, W) \text{ is surjective} &\iff \operatorname{range} T = W \\ &\iff (\operatorname{range} T)^0 = \{0\} \\ &\iff \operatorname{null} T' = \{0\} \\ &\iff T' \text{ is injective,}\end{aligned}$$

where the second equivalence comes from 3.127(a) and the third equivalence comes from 3.128(a). ∎

3.130 *the range of T'*

Suppose V and W are finite-dimensional and $T \in \mathcal{L}(V, W)$. Then

(a) $\dim \operatorname{range} T' = \dim \operatorname{range} T$;

(b) $\operatorname{range} T' = (\operatorname{null} T)^0$.

Proof

(a) We have

$$\begin{aligned}\dim \operatorname{range} T' &= \dim W' - \dim \operatorname{null} T' \\ &= \dim W - \dim(\operatorname{range} T)^0 \\ &= \dim \operatorname{range} T,\end{aligned}$$

where the first equality comes from 3.21, the second equality comes from 3.111 and 3.128(a), and the third equality comes from 3.125.

ב. מתקיים

$$\begin{aligned}\dim \operatorname{null} T' &= \dim(\operatorname{range} T)^0\\ &= \dim W - \dim \operatorname{range} T\\ &= \dim W - (\dim V - \dim \operatorname{null} T)\\ &= \dim \operatorname{null} T + \dim W - \dim V,\end{aligned}$$

כאשר השוויון הראשון נובע מ־ (א), השוויון השני נובע מ־ 3.125, והשוויון השלישי נובע מהמשפט היסודי של ההעתקות הליניאריות (3.21). ■

המסקנה הבאה יכולה להיות שימושית שכן לעיתים קל יותר להראות ש־ T' חד־חד־ערכית מאשר להראות ישירות ש־ T על.

3.129 T על שקול ל־ T' חד־חד־ערכית

נניח ש־ V ו־ W מממד סופי ו־ $T \in \mathcal{L}(V,W)$. אזי

T על $\iff$ T' חד־חד־ערכית.

הוכחה מתקיים

$$\begin{aligned}\operatorname{range} T = W &\iff T \text{ על } T \in \mathcal{L}(V,W)\\ (\operatorname{range} T)^0 = \{0\} &\iff\\ \operatorname{null} T' = \{0\} &\iff\\ T' \text{ חד־חד־ערכית}, &\iff\end{aligned}$$

כאשר השקילות השנייה נובעת מ־ 3.127(א) והשקילות השלישית נובעת מ־ 3.128(א). ■

3.130 הטווח של T'

נניח ש־ V ו־ W מממד סופי ו־ $T \in \mathcal{L}(V,W)$. אזי

א. $\dim \operatorname{range} T' = \dim \operatorname{range} T$

ב. $\operatorname{range} T' = (\operatorname{null} T)^0$.

הוכחה

א. מתקיים

$$\begin{aligned}\dim \operatorname{range} T' &= \dim W' - \dim \operatorname{null} T'\\ &= \dim W - \dim(\operatorname{range} T)^0\\ &= \dim \operatorname{range} T,\end{aligned}$$

כאשר השוויון הראשון נובע מ־ 3.21, השוויון השני נובע מ־ 3.111 ו־ 3.128(א), והשוויון השלישי נובע מ־ 3.125.

(b) First suppose $\varphi \in \operatorname{range} T'$. Thus there exists $\psi \in W'$ such that $\varphi = T'(\psi)$. If $v \in \operatorname{null} T$, then

$$\varphi(v) = (T'(\psi))v = (\psi \circ T)(v) = \psi(Tv) = \psi(0) = 0.$$

Hence $\varphi \in (\operatorname{null} T)^0$. This implies that $\operatorname{range} T' \subseteq (\operatorname{null} T)^0$.

We will complete the proof by showing that $\operatorname{range} T'$ and $(\operatorname{null} T)^0$ have the same dimension. To do this, note that

$$\begin{aligned} \dim \operatorname{range} T' &= \dim \operatorname{range} T \\ &= \dim V - \dim \operatorname{null} T \\ &= \dim (\operatorname{null} T)^0, \end{aligned}$$

where the first equality comes from (a), the second equality comes from 3.21, and the third equality comes from 3.125. ∎

The next result should be compared to 3.129.

3.131 *T injective is equivalent to T' surjective*

Suppose V and W are finite-dimensional and $T \in \mathcal{L}(V, W)$. Then

$$T \text{ is injective} \iff T' \text{ is surjective.}$$

Proof We have

$$\begin{aligned} T \text{ is injective} &\iff \operatorname{null} T = \{0\} \\ &\iff (\operatorname{null} T)^0 = V' \\ &\iff \operatorname{range} T' = V', \end{aligned}$$

where the second equivalence follows from 3.127(b) and the third equivalence follows from 3.130(b). ∎

Matrix of Dual of Linear Map

The setting for the next result is the assumption that we have a basis $v_1, \ldots, v_n$ of V, along with its dual basis $\varphi_1, \ldots, \varphi_n$ of V'. We also have a basis $w_1, \ldots, w_m$ of W, along with its dual basis $\psi_1, \ldots, \psi_m$ of W'. Thus $\mathcal{M}(T)$ is computed with respect to the bases just mentioned of V and W, and $\mathcal{M}(T')$ is computed with respect to the dual bases just mentioned of W' and V'. Using these bases gives the following pretty result.

3.132 *matrix of T' is transpose of matrix of T*

Suppose V and W are finite-dimensional and $T \in \mathcal{L}(V, W)$. Then

$$\mathcal{M}(T') = (\mathcal{M}(T))^t.$$

ב. נניח תחילה ש־ $\varphi \in \operatorname{range} T'$. לפיכך קיים $\psi \in W'$ כך ש־ $\varphi = T'(\psi)$. אם $v \in \operatorname{null} T$, אזי

$$\varphi(v) = (T'(\psi))v = (\psi \circ T)(v) = \psi(Tv) = \psi(0) = 0.$$

לכן $\varphi \in (\operatorname{null} T)^0$. מכאן נובע ש־ $\operatorname{range} T' \subseteq (\operatorname{null} T)^0$.

נשלים את ההוכחה בכך שנראה כי הממדים של $\operatorname{range} T'$ ו־ $(\operatorname{null} T)^0$ זהים. לצורך כך, נציין כי

$$\begin{aligned}\dim \operatorname{range} T' &= \dim \operatorname{range} T \\ &= \dim V - \dim \operatorname{null} T \\ &= \dim (\operatorname{null} T)^0,\end{aligned}$$

כאשר השוויון הראשון נובע מ־ (א), השוויון השני נובע מ־ 3.21, והשוויון השלישי נובע מ־ 3.125. ■

את המסקנה הבאה יש להשוות ל־ 3.129.

3.131 T חד־חד־ערכית שקול ל־ T' על

נניח ש־ V ו־ W מממד סופי ו־ $T \in \mathcal{L}(V, W)$. אזי

T חד־חד־ערכית $\iff$ T' על.

הוכחה מתקיים

$$\begin{aligned}T \text{ חד־חד־ערכית} &\iff \operatorname{null} T = \{0\} \\ &\iff (\operatorname{null} T)^0 = V' \\ &\iff \operatorname{range} T' = V',\end{aligned}$$

כאשר השקילות השנייה נובעת מ־ 3.127(ב) והשקילות השלישית נובעת מ־ 3.130(ב). ■

מטריצה של העתקה דואלית

המסגרת עבור המסקנה הבאה היא ההנחה שיש לנו בסיס $v_1, \ldots, v_n$ של V, יחד עם הבסיס הדואלי שלו $\varphi_1, \ldots, \varphi_n$ של V'. בנוסף יש לנו בסיס $w_1, \ldots, w_m$ של W, יחד עם הבסיס הדואלי שלו $\psi_1, \ldots, \psi_m$ של W'. לפיכך $\mathcal{M}(T)$ מחושבת ביחס לבסיסים של V ו־ W שאותם ציינו זה עתה, ואילו $\mathcal{M}(T')$ מחושבת ביחס לבסיסים הדואליים שאותם ציינו זה עתה של W' ו־ V'. שימוש בבסיסים אלו יוביל אל המסקנה היפה הבאה.

3.132 מטריצת T' היא המטריצה המשוחלפת של T

נניח ש־ V ו־ W מממד סופי ו־ $T \in \mathcal{L}(V, W)$. אזי

$$\mathcal{M}(T') = (\mathcal{M}(T))^t.$$

Proof Let $A = \mathcal{M}(T)$ and $C = \mathcal{M}(T')$. Suppose $1 \le j \le m$ and $1 \le k \le n$. From the definition of $\mathcal{M}(T')$ we have

$$T'(\psi_j) = \sum_{r=1}^{n} C_{r,j}\varphi_r.$$

The left side of the equation above equals $\psi_j \circ T$. Thus applying both sides of the equation above to v_k gives

$$\begin{aligned}(\psi_j \circ T)(v_k) &= \sum_{r=1}^{n} C_{r,j}\varphi_r(v_k)\\ &= C_{k,j}.\end{aligned}$$

We also have

$$\begin{aligned}(\psi_j \circ T)(v_k) &= \psi_j(Tv_k)\\ &= \psi_j\Big(\sum_{r=1}^{m} A_{r,k}w_r\Big)\\ &= \sum_{r=1}^{m} A_{r,k}\psi_j(w_r)\\ &= A_{j,k}.\end{aligned}$$

Comparing the last line of the last two sets of equations, we have $C_{k,j} = A_{j,k}$. Thus $C = A^t$. In other words, $\mathcal{M}(T') = (\mathcal{M}(T))^t$, as desired. ∎

Now we use duality to give an alternative proof that the column rank of a matrix equals the row rank of the matrix. This result was previously proved using different tools—see 3.57.

3.133 *column rank equals row rank*

Suppose $A \in \mathbf{F}^{m,n}$. Then the column rank of A equals the row rank of A.

Proof Define $T\colon \mathbf{F}^{n,1} \to \mathbf{F}^{m,1}$ by $Tx = Ax$. Thus $\mathcal{M}(T) = A$, where $\mathcal{M}(T)$ is computed with respect to the standard bases of $\mathbf{F}^{n,1}$ and $\mathbf{F}^{m,1}$. Now

$$\begin{aligned}\text{column rank of } A &= \text{column rank of } \mathcal{M}(T)\\ &= \dim \text{range } T\\ &= \dim \text{range } T'\\ &= \text{column rank of } \mathcal{M}(T')\\ &= \text{column rank of } A^t\\ &= \text{row rank of } A,\end{aligned}$$

where the second equality comes from 3.78, the third equality comes from 3.130(a), the fourth equality comes from 3.78, the fifth equality comes from 3.132, and the last equality follows from the definitions of row and column rank. ∎

See Exercise 8 in Section 7A for another alternative proof of the result above.

הוכחה תהיינה $A = \mathcal{M}(T)$ ו־ $C = \mathcal{M}(T')$. נניח ש־ $1 \le j \le m$ ו־ $1 \le k \le n$. מהגדרת $\mathcal{M}(T')$ נקבל

$$T'(\psi_j) = \sum_{r=1}^{n} C_{r,j}\varphi_r.$$

האגף השמאלי של המשוואה לעיל שווה ל־ $\psi_j \circ T$. לפיכך הפעלת שני אגפי המשוואה לעיל על v_k נותנת

$$\begin{aligned}(\psi_j \circ T)(v_k) &= \sum_{r=1}^{n} C_{r,j}\varphi_r(v_k)\\ &= C_{k,j}.\end{aligned}$$

בנוסף, מתקיים

$$\begin{aligned}(\psi_j \circ T)(v_k) &= \psi_j(Tv_k)\\ &= \psi_j\Big(\sum_{r=1}^{m} A_{r,k}w_r\Big)\\ &= \sum_{r=1}^{m} A_{r,k}\psi_j(w_r)\\ &= A_{j,k}.\end{aligned}$$

בהשוואת השורה האחרונה של כל אחת משתי קבוצות המשוואות, אנו מקבלים $C_{k,j} = A_{j,k}$. לפיכך $C = A^{\mathrm{t}}$. במילים אחרות, $\mathcal{M}(T') = (\mathcal{M}(T))^{\mathrm{t}}$, כפי שרצינו. ■

כעת אנו משתמשים בדואליות כדי לספק הוכחה חלופית לכך שדרגת העמודות של מטריצה שווה לדרגת השורות של המטריצה. את מסקנה זו הוכחנו לעיל באמצעים אחרים—ראו 3.57.

3.133 דרגת העמודות שווה לדרגת השורות

נניח ש־ $A \in \mathbf{F}^{m,n}$. אזי דרגת העמודות של A שווה לדרגת השורות של A.

הוכחה נגדיר $T : \mathbf{F}^{n,1} \to \mathbf{F}^{m,1}$ על ידי $Tx = Ax$. לפיכך $\mathcal{M}(T) = A$, כאשר $\mathcal{M}(T)$ מחושבת ביחס לבסיסים הסטנדרטיים של $\mathbf{F}^{n,1}$ ו־ $\mathbf{F}^{m,1}$. נקבל כעת

$$\begin{aligned}\text{דרגת העמודות של } A &= \text{דרגת העמודות של } \mathcal{M}(T)\\ &= \dim \operatorname{range} T\\ &= \dim \operatorname{range} T'\\ &= \text{דרגת העמודות של } \mathcal{M}(T')\\ &= \text{דרגת העמודות של } A^{\mathrm{t}}\\ &= \text{דרגת השורות של } A,\end{aligned}$$

כאשר השוויון השני נובע מ־ 3.78, השוויון השלישי נובע מ־ 3.130(א), השוויון הרביעי נובע מ־ 3.78, השוויון החמישי נובע מ־ 3.132, והשוויון האחרון נובע מההגדרות של דרגות השורות והעמודות. ■

ראו תרגיל 8 ביחידה 1ז להוכחה חלופית נוספת למסקנה לעיל.

Exercises 3F

1 Explain why each linear functional is surjective or is the zero map.

2 Give three distinct examples of linear functionals on $\mathbf{R}^{[0,1]}$.

3 Suppose V is finite-dimensional and $v \in V$ with $v \neq 0$. Prove that there exists $\varphi \in V'$ such that $\varphi(v) = 1$.

4 Suppose V is finite-dimensional and U is a subspace of V such that $U \neq V$. Prove that there exists $\varphi \in V'$ such that $\varphi(u) = 0$ for every $u \in U$ but $\varphi \neq 0$.

5 Suppose $T \in \mathcal{L}(V, W)$ and $w_1, \dots, w_m$ is a basis of range T. Hence for each $v \in V$, there exist unique numbers $\varphi_1(v), \dots, \varphi_m(v)$ such that

$$Tv = \varphi_1(v)w_1 + \cdots + \varphi_m(v)w_m,$$

thus defining functions $\varphi_1, \dots, \varphi_m$ from V to $\mathbf{F}$. Show that each of the functions $\varphi_1, \dots, \varphi_m$ is a linear functional on V.

6 Suppose $\varphi, \beta \in V'$. Prove that $\operatorname{null} \varphi \subseteq \operatorname{null} \beta$ if and only if there exists $c \in \mathbf{F}$ such that $\beta = c\varphi$.

7 Suppose that $V_1, \dots, V_m$ are vector spaces. Prove that $(V_1 \times \cdots \times V_m)'$ and ${V_1}' \times \cdots \times {V_m}'$ are isomorphic vector spaces.

8 Suppose $v_1, \dots, v_n$ is a basis of V and $\varphi_1, \dots, \varphi_n$ is the dual basis of V'. Define $\Gamma \colon V \to \mathbf{F}^n$ and $\Lambda \colon \mathbf{F}^n \to V$ by

$$\Gamma(v) = (\varphi_1(v), \dots, \varphi_n(v)) \quad \text{and} \quad \Lambda(a_1, \dots, a_n) = a_1v_1 + \cdots + a_nv_n.$$

Explain why Γ and Λ are inverses of each other.

9 Suppose m is a positive integer. Show that the dual basis of the basis $1, x, \dots, x^m$ of $\mathcal{P}_m(\mathbf{R})$ is $\varphi_0, \varphi_1, \dots, \varphi_m$, where

$$\varphi_k(p) = \frac{p^{(k)}(0)}{k!}.$$

Here $p^{(k)}$ denotes the kth derivative of p, with the understanding that the 0th derivative of p is p.

10 Suppose m is a positive integer.

(a) Show that $1, x - 5, \dots, (x - 5)^m$ is a basis of $\mathcal{P}_m(\mathbf{R})$.

(b) What is the dual basis of the basis in (a)?

11 Suppose $v_1, \dots, v_n$ is a basis of V and $\varphi_1, \dots, \varphi_n$ is the corresponding dual basis of V'. Suppose $\psi \in V'$. Prove that

$$\psi = \psi(v_1)\varphi_1 + \cdots + \psi(v_n)\varphi_n.$$

תרגילים 6ג

1 הסבירו מדוע כל פונקציונל ליניארי הוא או עָל או העתקת האפס.

2 הביאו שלוש דוגמאות שונות של פונקציונלים ליניאריים מעל $\mathbf{R}^{[0,1]}$.

3 נניח ש־ V מממד סופי ו־ $v \in V$ עם $v \neq 0$. הוכיחו שקיים $\varphi \in V'$ כך ש־ $\varphi(v) = 1$.

4 נניח ש־ V מממד סופי ו־ U תת־מרחב של V כך ש־ $U \neq V$. הוכיחו שקיים $\varphi \in V'$ כך ש־ $\varphi(u) = 0$ לכל $u \in U$ אבל $\varphi \neq 0$.

5 נניח ש־ $T \in \mathcal{L}(V, W)$ ו־ $w_1, \dots, w_m$ בסיס של $\operatorname{range} T$. לכן לכל $v \in V$, קיימים מספרים יחידים $\varphi_1(v), \dots, \varphi_m(v)$ כך ש־

$$Tv = \varphi_1(v)w_1 + \cdots + \varphi_m(v)w_m,$$

ובכך מוגדרות פונקציות $\varphi_1, \dots, \varphi_m$ מ־ V ל־ $\mathbf{F}$. הראו שכל אחת מהפונקציות $\varphi_1, \dots, \varphi_m$ היא פונקציונל ליניארי מעל V.

6 נניח ש־ $\varphi, \beta \in V'$. הוכיחו ש־ $\operatorname{null} \varphi \subseteq \operatorname{null} \beta$ אם־ורק־אם קיים $c \in \mathbf{F}$ כך ש־ $\beta = c\varphi$.

7 נניח ש־ $V_1, \dots, V_m$ מרחבים וקטוריים. הוכיחו ש־ $(V_1 \times \cdots \times V_m)'$ ו־ $V_1' \times \cdots \times V_m'$ מרחבים וקטוריים איזומורפיים.

8 נניח ש־ $v_1, \dots, v_n$ בסיס של V ו־ $\varphi_1, \dots, \varphi_n$ הבסיס הדואלי של V'. הגדירו $\Gamma \colon V \to \mathbf{F}^n$ ו־ $\Lambda \colon \mathbf{F}^n \to V$ על ידי

$$\Lambda(a_1, \dots, a_n) = a_1v_1 + \cdots + a_nv_n \quad \text{וגם} \quad \Gamma(v) = (\varphi_1(v), \dots, \varphi_n(v))$$

הסבירו מדוע Γ ו־ Λ הופכיות זו של זו.

9 נניח ש־ m מספר שלם חיובי. הראו שהבסיס הדואלי של הבסיס $1, x, \dots, x^m$ של $\mathcal{P}_m(\mathbf{R})$ הוא $\varphi_0, \varphi_1, \dots, \varphi_m$, כאשר

$$\varphi_k(p) = \frac{p^{(k)}(0)}{k!}.$$

כאן $p^{(k)}$ מסמן את הנגזרת ה־ k^{th} של p, תוך הבנה כי הנגזרת ה־ 0th של p היא p.

10 נניח ש־ m מספר שלם חיובי.

א. הראו ש־ $1, x-5, \dots, (x-5)^m$ בסיס של $\mathcal{P}_m(\mathbf{R})$.

ב. מהו הבסיס הדואלי של הבסיס ב־ (א)?

11 נניח ש־ $v_1, \dots, v_n$ בסיס של V ו־ $\varphi_1, \dots, \varphi_n$ הבסיס הדואלי המתאים של V'. נניח ש־ $\psi \in V'$. הוכיחו ש־

$$\psi = \psi(v_1)\varphi_1 + \cdots + \psi(v_n)\varphi_n.$$

12 Suppose $S, T \in \mathcal{L}(V, W)$.

(a) Prove that $(S + T)' = S' + T'$.

(b) Prove that $(\lambda T)' = \lambda T'$ for all $\lambda \in \mathbf{F}$.

This exercise asks you to verify (a) and (b) in 3.120.

13 Show that the dual map of the identity operator on V is the identity operator on V'.

14 Define $T : \mathbf{R}^3 \to \mathbf{R}^2$ by

$$T(x, y, z) = (4x + 5y + 6z, 7x + 8y + 9z).$$

Suppose φ_1, φ_2 denotes the dual basis of the standard basis of $\mathbf{R}^2$ and ψ_1, ψ_2, ψ_3 denotes the dual basis of the standard basis of $\mathbf{R}^3$.

(a) Describe the linear functionals $T'(\varphi_1)$ and $T'(\varphi_2)$.

(b) Write $T'(\varphi_1)$ and $T'(\varphi_2)$ as linear combinations of ψ_1, ψ_2, ψ_3.

15 Define $T : \mathcal{P}(\mathbf{R}) \to \mathcal{P}(\mathbf{R})$ by

$$(Tp)(x) = x^2 p(x) + p''(x)$$

for each $x \in \mathbf{R}$.

(a) Suppose $\varphi \in \mathcal{P}(\mathbf{R})'$ is defined by $\varphi(p) = p'(4)$. Describe the linear functional $T'(\varphi)$ on $\mathcal{P}(\mathbf{R})$.

(b) Suppose $\varphi \in \mathcal{P}(\mathbf{R})'$ is defined by $\varphi(p) = \int_0^1 p$. Evaluate $(T'(\varphi))(x^3)$.

16 Suppose W is finite-dimensional and $T \in \mathcal{L}(V, W)$. Prove that

$$T' = 0 \iff T = 0.$$

17 Suppose V and W are finite-dimensional and $T \in \mathcal{L}(V, W)$. Prove that T is invertible if and only if $T' \in \mathcal{L}(W', V')$ is invertible.

18 Suppose V and W are finite-dimensional. Prove that the map that takes $T \in \mathcal{L}(V, W)$ to $T' \in \mathcal{L}(W', V')$ is an isomorphism of $\mathcal{L}(V, W)$ onto $\mathcal{L}(W', V')$.

19 Suppose $U \subseteq V$. Explain why

$$U^0 = \{\varphi \in V' : U \subseteq \operatorname{null} \varphi\}.$$

20 Suppose V is finite-dimensional and U is a subspace of V. Show that

$$U = \{v \in V : \varphi(v) = 0 \text{ for every } \varphi \in U^0\}.$$

21 Suppose V is finite-dimensional and U and W are subspaces of V.

(a) Prove that $W^0 \subseteq U^0$ if and only if $U \subseteq W$.

(b) Prove that $W^0 = U^0$ if and only if $U = W$.

12 נניח ש־ $S, T \in \mathcal{L}(V, W)$.

א. הוכיחו ש־ $(S+T)' = S' + T'$.

ב. הוכיחו ש־ $(\lambda T)' = \lambda T'$ לכל $\lambda \in \mathbf{F}$.

תרגיל זה מבקש מכם לוודא את (א) ו־ (ב) ב־ 3.120.

13 הראו שההעתקה הדואלית של אופרטור הזהות מעל V היא אופרטור הזהות מעל V'.

14 נגדיר $T: \mathbf{R}^3 \to \mathbf{R}^2$ על ידי

$$T(x, y, z) = (4x + 5y + 6z, 7x + 8y + 9z).$$

נניח ש־ φ_1, φ_2 מסמנים את הבסיס הדואלי של הבסיס הסטנדרטי של $\mathbf{R}^2$ ו־ ψ_1, ψ_2, ψ_3 מסמנים את הבסיס הדואלי של הבסיס הסטנדרטי של $\mathbf{R}^3$.

א. תארו את הפונקציונלים הליניאריים $T'(\varphi_1)$ ו־ $T'(\varphi_2)$.

ב. כתבו את $T'(\varphi_1)$ ו־ $T'(\varphi_2)$ כצירוף ליניארי של ψ_1, ψ_2, ψ_3.

15 נגדיר $T: \mathcal{P}(\mathbf{R}) \to \mathcal{P}(\mathbf{R})$ על ידי

$$(Tp)(x) = x^2 p(x) + p''(x)$$

לכל $x \in \mathbf{R}$.

א. נניח ש־ $\varphi \in \mathcal{P}(\mathbf{R})'$ מוגדר על ידי $\varphi(p) = p'(4)$. תארו את הפונקציונל הליניארי $T'(\varphi)$ מעל $\mathcal{P}(\mathbf{R})$.

ב. נניח ש־ $\varphi \in \mathcal{P}(\mathbf{R})'$ מוגדר על ידי $\varphi(p) = \int_0^1 p$. חשבו $(T'(\varphi))(x^3)$.

16 נניח ש־ W מממד סופי ו־ $T \in \mathcal{L}(V, W)$. הוכיחו ש־

$$T' = 0 \iff T = 0.$$

17 נניח ש־ V ו־ W מממד סופי ו־ $T \in \mathcal{L}(V, W)$. הוכיחו ש־ T הפיכה אם־ורק־אם $T' \in \mathcal{L}(W', V')$ הפיכה.

18 נניח ש־ V ו־ W מממד סופי. הוכיחו שההעתקה המעתיקה $T \in \mathcal{L}(V, W)$ אל $T' \in \mathcal{L}(W', V')$ היא איזומורפיזם של $\mathcal{L}(V, W)$ על $\mathcal{L}(W', V')$.

19 נניח ש־ $U \subseteq V$. הסבירו מדוע

$$U^0 = \{\varphi \in V' : U \subseteq \operatorname{null} \varphi\}.$$

20 נניח ש־ V מממד סופי ו־ U תת־מרחב של V. הראו ש־

$$U = \{v \in V : \varphi(v) = 0 \text{ לכל } \varphi \in U^0\}.$$

21 נניח ש־ V מממד סופי ו־ U ו־ W תתי־מרחב של V.

א. הוכיחו ש־ $W^0 \subseteq U^0$ אם־ורק־אם $U \subseteq W$.

ב. הוכיחו ש־ $W^0 = U^0$ אם־ורק־אם $U = W$.

22 Suppose V is finite-dimensional and U and W are subspaces of V.

(a) Show that $(U + W)^0 = U^0 \cap W^0$.
(b) Show that $(U \cap W)^0 = U^0 + W^0$.

23 Suppose V is finite-dimensional and $\varphi_1, \dots, \varphi_m \in V'$. Prove that the following three sets are equal to each other.

(a) $\operatorname{span}(\varphi_1, \dots, \varphi_m)$
(b) $((\operatorname{null} \varphi_1) \cap \cdots \cap (\operatorname{null} \varphi_m))^0$
(c) $\{\varphi \in V' : (\operatorname{null} \varphi_1) \cap \cdots \cap (\operatorname{null} \varphi_m) \subseteq \operatorname{null} \varphi\}$

24 Suppose V is finite-dimensional and $v_1, \dots, v_m \in V$. Define a linear map $\Gamma : V' \to \mathbf{F}^m$ by $\Gamma(\varphi) = (\varphi(v_1), \dots, \varphi(v_m))$.

(a) Prove that $v_1, \dots, v_m$ spans V if and only if Γ is injective.
(b) Prove that $v_1, \dots, v_m$ is linearly independent if and only if Γ is surjective.

25 Suppose V is finite-dimensional and $\varphi_1, \dots, \varphi_m \in V'$. Define a linear map $\Gamma : V \to \mathbf{F}^m$ by $\Gamma(v) = (\varphi_1(v), \dots, \varphi_m(v))$.

(a) Prove that $\varphi_1, \dots, \varphi_m$ spans V' if and only if Γ is injective.
(b) Prove that $\varphi_1, \dots, \varphi_m$ is linearly independent if and only if Γ is surjective.

26 Suppose V is finite-dimensional and Ω is a subspace of V'. Prove that

$$\Omega = \{v \in V : \varphi(v) = 0 \text{ for every } \varphi \in \Omega\}^0.$$

27 Suppose $T \in \mathcal{L}(\mathcal{P}_5(\mathbf{R}))$ and $\operatorname{null} T' = \operatorname{span}(\varphi)$, where φ is the linear functional on $\mathcal{P}_5(\mathbf{R})$ defined by $\varphi(p) = p(8)$. Prove that

$$\operatorname{range} T = \{p \in \mathcal{P}_5(\mathbf{R}) : p(8) = 0\}.$$

28 Suppose V is finite-dimensional and $\varphi_1, \dots, \varphi_m$ is a linearly independent list in V'. Prove that

$$\dim((\operatorname{null} \varphi_1) \cap \cdots \cap (\operatorname{null} \varphi_m)) = (\dim V) - m.$$

29 Suppose V and W are finite-dimensional and $T \in \mathcal{L}(V, W)$.

(a) Prove that if $\varphi \in W'$ and $\operatorname{null} T' = \operatorname{span}(\varphi)$, then $\operatorname{range} T = \operatorname{null} \varphi$.
(b) Prove that if $\psi \in V'$ and $\operatorname{range} T' = \operatorname{span}(\psi)$, then $\operatorname{null} T = \operatorname{null} \psi$.

30 Suppose V is finite-dimensional and $\varphi_1, \dots, \varphi_n$ is a basis of V'. Show that there exists a basis of V whose dual basis is $\varphi_1, \dots, \varphi_n$.

31 Suppose U is a subspace of V. Let $i : U \to V$ be the inclusion map defined by $i(u) = u$. Thus $i' \in \mathcal{L}(V', U')$.

(a) Show that $\operatorname{null} i' = U^0$.
(b) Prove that if V is finite-dimensional, then $\operatorname{range} i' = U'$.
(c) Prove that if V is finite-dimensional, then $\tilde{i'}$ is an isomorphism from V'/U^0 onto U'.

The isomorphism in (c) is natural in that it does not depend on a choice of basis in either vector space.

22 נניח ש־ V מממד סופי ו־ U ו־ W תתי־מרחב של V.

א. הראו ש־ $(U+W)^0 = U^0 \cap W^0$.

ב. הראו ש־ $(U \cap W)^0 = U^0 + W^0$.

23 נניח ש־ V מממד סופי ו־ $\varphi_1, \ldots, \varphi_m \in V'$. הוכיחו ששלוש הקבוצות הבאות שוות זו לזו.

א. $\operatorname{span}(\varphi_1, \ldots, \varphi_m)$

ב. $((\operatorname{null}\varphi_1) \cap \cdots \cap (\operatorname{null}\varphi_m))^0$

ג. $\{\varphi \in V' : (\operatorname{null}\varphi_1) \cap \cdots \cap (\operatorname{null}\varphi_m) \subseteq \operatorname{null}\varphi\}$

24 נניח ש־ V מממד סופי ו־ $v_1, \ldots, v_m \in V$. הגדירו העתקה ליניארית $\Gamma : V' \to \mathbf{F}^m$ על ידי $\Gamma(\varphi) = (\varphi(v_1), \ldots, \varphi(v_m))$.

א. הוכיחו ש־ $v_1, \ldots, v_m$ פורסת את V אם־ורק־אם Γ חד־חד־ערכית.

ב. הוכיחו ש־ $v_1, \ldots, v_m$ בלתי תלויה ליניארית אם־ורק־אם Γ על.

25 נניח ש־ V מממד סופי ו־ $\varphi_1, \ldots, \varphi_m \in V'$. הגדירו העתקה ליניארית $\Gamma : V \to \mathbf{F}^m$ על ידי $\Gamma(v) = (\varphi_1(v), \ldots, \varphi_m(v))$.

א. הוכיחו ש־ $\varphi_1, \ldots, \varphi_m$ פורסת את V' אם־ורק־אם Γ חד־חד־ערכית.

ב. הוכיחו ש־ $\varphi_1, \ldots, \varphi_m$ בלתי תלויה ליניארית אם־ורק־אם Γ על.

26 נניח ש־ V מממד סופי ו־ Ω תת־מרחב של V'. הוכיחו ש־

$$\Omega = \{v \in V : \varphi(v) = 0 \text{ לכל } \varphi \in \Omega\}^0.$$

27 נניח ש־ $T \in \mathcal{L}(\mathcal{P}_5(\mathbf{R}))$ ו־ $\operatorname{null} T' = \operatorname{span}(\varphi)$, כאשר φ הפונקציונל הליניארי מעל $\mathcal{P}_5(\mathbf{R})$ המוגדר על ידי $\varphi(p) = p(8)$. הוכיחו ש־

$$\operatorname{range} T = \{p \in \mathcal{P}_5(\mathbf{R}) : p(8) = 0\}.$$

28 נניח ש־ V מממד סופי ו־ $\varphi_1, \ldots, \varphi_m$ רשימה בלתי תלויה ליניארית ב־ V'. הוכיחו ש־

$$\dim((\operatorname{null}\varphi_1) \cap \cdots \cap (\operatorname{null}\varphi_m)) = (\dim V) - m.$$

29 נניח ש־ V ו־ W מממד סופי ו־ $T \in \mathcal{L}(V, W)$.

א. הוכיחו שאם $\varphi \in W'$ ו־ $\operatorname{null} T' = \operatorname{span}(\varphi)$, אזי $\operatorname{range} T = \operatorname{null}\varphi$.

ב. הוכיחו שאם $\psi \in V'$ ו־ $\operatorname{range} T' = \operatorname{span}(\psi)$, אזי $\operatorname{null} T = \operatorname{null}\psi$.

30 נניח ש־ V מממד סופי ו־ $\varphi_1, \ldots, \varphi_n$ בסיס של V'. הראו שקיים בסיס של V שבסיסו הדואלי הוא $\varphi_1, \ldots, \varphi_n$.

31 נניח ש־ U תת־מרחב של V. תהי $i : U \to V$ העתקת ההכלה המוגדרת על ידי $i(u) = u$. לפיכך $i' \in \mathcal{L}(V', U')$.

א. הראו ש־ $\operatorname{null} i' = U^0$.

ב. הוכיחו שאם V מממד סופי, אזי $\operatorname{range} i' = U'$.

ג. הוכיחו שאם V מממד סופי, אזי $\tilde{i}$ איזומורפיזם מ־ V'/U^0 על U'.

האיזומורפיזם ב־ (ג) הוא טבעי במובן שאינו תלוי בבחירת בסיס של אף אחד מהמרחבים הווקטוריים.

אלגברה ליניארית בדרך הישר, מהדורה רביעית, מחבר: שלדון אקסלר
Linear Algebra Done Right, fourth edition, by Sheldon Axler

32 The *double dual space* of V, denoted by V'', is defined to be the dual space of V'. In other words, $V'' = (V')'$. Define $\Lambda : V \to V''$ by

$$(\Lambda v)(\varphi) = \varphi(v)$$

for each $v \in V$ and each $\varphi \in V'$.

(a) Show that Λ is a linear map from V to V''.

(b) Show that if $T \in \mathcal{L}(V)$, then $T'' \circ \Lambda = \Lambda \circ T$, where $T'' = (T')'$.

(c) Show that if V is finite-dimensional, then Λ is an isomorphism from V onto V''.

Suppose V is finite-dimensional. Then V and V' are isomorphic, but finding an isomorphism from V onto V' generally requires choosing a basis of V. In contrast, the isomorphism Λ from V onto V'' does not require a choice of basis and thus is considered more natural.

33 Suppose U is a subspace of V. Let $\pi : V \to V/U$ be the usual quotient map. Thus $\pi' \in \mathcal{L}((V/U)', V')$.

(a) Show that π' is injective.

(b) Show that range $\pi' = U^0$.

(c) Conclude that π' is an isomorphism from $(V/U)'$ onto U^0.

The isomorphism in (c) is natural in that it does not depend on a choice of basis in either vector space. In fact, there is no assumption here that any of these vector spaces are finite-dimensional.

32 המרחב הזואלי הכפול של V, המסומן ב־ V'', מוגדר כמרחב הדואלי של V'. במילים אחרות, $V'' = (V')'$. הגדירו $\Lambda : V \to V''$ באופן הבא

$$(\Lambda v)(\varphi) = \varphi(v)$$

לכל $v \in V$ ולכל $\varphi \in V'$.

א. הראו ש־ Λ העתקה ליניארית מ־ V אל V''.

ב. הראו שאם $T \in \mathcal{L}(V)$, אזי $T'' \circ \Lambda = \Lambda \circ T$, כאשר $T'' = (T')'$.

ג. הראו שאם V מממד סופי, אזי Λ איזומורפיזם מ־ V על V''.

ATOM-03-748

נניח ש־ V מממד סופי. אזי V ו־ V' איזומורפיים, אך מציאת איזומורפיזם מ־ V על V' בדרך כלל דורשת בחירת בסיס של V. בניגוד לכך, האיזומורפיזם Λ מ־ V על V'' אינו דורש בחירת בסיס ולפיכך נחשב לטבעי יותר.

33 נניח ש־ U תת־מרחב של V. תהי $\pi : V \to V/U$ העתקת המנה הרגילה. לפיכך $\pi' \in \mathcal{L}((V/U)', V')$.

א. הראו ש־ π' חד־חד־ערכית.

ב. הראו ש־ $\operatorname{range} \pi' = U^0$.

ג. הסיקו כי π' איזומורפיזם מ־ $(V/U)'$ על U^0.

האיזומורפיזם ב־ (ג) הוא טבעי במובן שאינו תלוי בבחירת בסיס של אף אחד מהמרחבים הווקטוריים. למעשה, אין כאן הנחה שאחד מהמרחבים הווקטוריים הללו מממד סופי.

Chapter 4

Polynomials

This chapter contains material on polynomials that we will use to investigate linear maps from a vector space to itself. Many results in this chapter will already be familiar to you from other courses; they are included here for completeness.

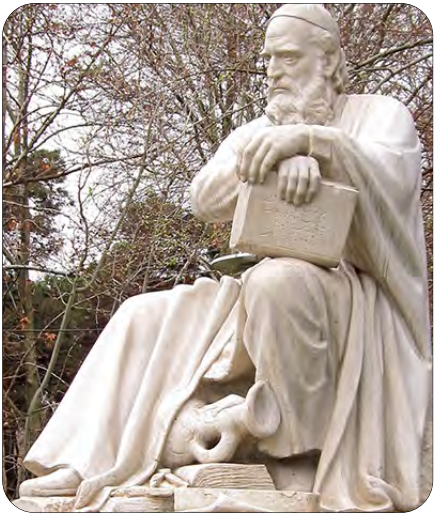
Because this chapter is not about linear algebra, your instructor may go through it rapidly. You may not be asked to scrutinize all the proofs. Make sure, however, that you at least read and understand the statements of all results in this chapter—they will be used in later chapters.

This chapter begins with a brief discussion of algebraic properties of the complex numbers. Then we prove that a nonconstant polynomial cannot have more zeros than its degree. We also give a linear-algebra-based proof of the division algorithm for polynomials, which is worth reading even if you are already familiar with a proof that does not use linear algebra.

As we will see, the fundamental theorem of algebra leads to a factorization of every polynomial into degree-one factors if the scalar field is **C** or to factors of degree at most two if the scalar field is **R**.

standing assumption for this chapter

- **F** denotes **R** or **C**.



Alireza Javaheri CC BY

Statue of mathematician and poet Omar Khayyam (1048–1131), whose algebra book written in 1070 contained the first serious study of cubic polynomials.

פרק ד

פולינומים

פרק זה מכיל חומר על פולינומים שנזדקק לו כשנעסוק בהעתקות ליניאריות ממרחב וקטורי לעצמו. אף שרבות מהמסקנות בפרק זה כבר מוכרות לכם מקורסים אחרים, הן הוכללו כאן למען השלמות.

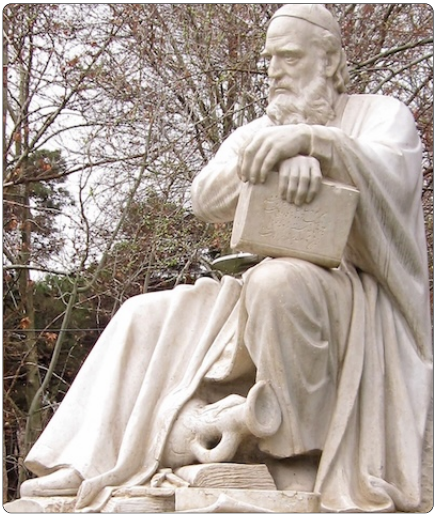
מאחר שפרק זה אינו עוסק באלגברה ליניארית, מנחה הקורס עשוי לעבור עליו בזריזות, וייתכן שלא תידרשו לנתח לעומק את כל ההוכחות. עם זאת, ודאו שלכל הפחות קראתם והבנתם את ניסוחי כל המסקנות בפרק—הן תהיינה בשימוש בפרקים הבאים.

הפרק נפתח בסקירה קצרה של המאפיינים האלגבריים של המספרים המרוכבים. לאחר מכן אנו מוכיחים כי מספר השורשים של פולינום שאינו קבוע הוא לכל היותר מעלתו. בנוסף אנו מספקים הוכחה המבוססת על אלגברה ליניארית לאלגוריתם החלוקה של פולינומים, שכדאי לקרוא אותה גם אם כבר מוכרת לכם הוכחה שאינה משתמשת באלגברה ליניארית.

כפי שנראה, המשפט היסודי של האלגברה יוביל אותנו לפירוק של כל פולינום לגורמים ממעלה ראשונה אם שדה הסקלרים הוא $\mathbf{C}$, או לגורמים ממעלה שנייה לכל היותר אם שדה הסקלרים הוא $\mathbf{R}$.

הנחות המוצא לפרק זה

- $\mathbf{F}$ מייצג $\mathbf{R}$ או $\mathbf{C}$.



Alireza Javaheri CC BY

פסלו של המתמטיקאי והמשורר עומר ח'יאם (1048–1131), שבספר האלגברה שכתב בשנת 1070 מופיע לראשונה חקר מעמיק של פולינומים ממעלה שלישית.

Before discussing polynomials with complex or real coefficients, we need to learn a bit more about the complex numbers.

4.1 definition: *real part,* $\operatorname{Re} z$*, imaginary part,* $\operatorname{Im} z$

Suppose $z = a + bi$, where a and b are real numbers.

- The *real part* of z, denoted by $\operatorname{Re} z$, is defined by $\operatorname{Re} z = a$.
- The *imaginary part* of z, denoted by $\operatorname{Im} z$, is defined by $\operatorname{Im} z = b$.

Thus for every complex number z, we have

$$z = \operatorname{Re} z + (\operatorname{Im} z)i.$$

4.2 definition: *complex conjugate,* $\overline{z}$*, absolute value,* $|z|$

Suppose $z \in \mathbf{C}$.

- The *complex conjugate* of $z \in \mathbf{C}$, denoted by $\overline{z}$, is defined by
$$\overline{z} = \operatorname{Re} z - (\operatorname{Im} z)i.$$
- The *absolute value* of a complex number z, denoted by $|z|$, is defined by
$$|z| = \sqrt{(\operatorname{Re} z)^2 + (\operatorname{Im} z)^2}.$$

4.3 example: *real and imaginary part, complex conjugate, absolute value*

Suppose $z = 3 + 2i$. Then

- $\operatorname{Re} z = 3$ and $\operatorname{Im} z = 2$;
- $\overline{z} = 3 - 2i$;
- $|z| = \sqrt{3^2 + 2^2} = \sqrt{13}$.

Identifying a complex number $z \in \mathbf{C}$ with the ordered pair $(\operatorname{Re} z, \operatorname{Im} z) \in \mathbf{R}^2$ identifies $\mathbf{C}$ with $\mathbf{R}^2$. Note that $\mathbf{C}$ is a one-dimensional complex vector space, but we can also think of $\mathbf{C}$ (identified with $\mathbf{R}^2$) as a two-dimensional real vector space.

The absolute value of each complex number is a nonnegative number. Specifically, if $z \in \mathbf{C}$, then $|z|$ equals the distance from the origin in $\mathbf{R}^2$ to the point $(\operatorname{Re} z, \operatorname{Im} z) \in \mathbf{R}^2$.

The real and imaginary parts, complex conjugate, and absolute value have the properties listed in the following multipart result.

You should verify that $z = \overline{z}$ if and only if z is a real number.

בטרם נעסוק בפולינומים עם מקדמים מרוכבים או ממשיים, עלינו ללמוד מעט יותר על המספרים המרוכבים.

4.1 הגדרה: חלק ממשי, $\operatorname{Re} z$, חלק מדומה, $\operatorname{Im} z$

נניח ש־ $z = a + bi$, כאשר a ו־ b מספרים ממשיים.

- החלק הממשי של z, המסומן ב־ $\operatorname{Re} z$, מוגדר על ידי $\operatorname{Re} z = a$.
- החלק המדומה של z, המסומן ב־ $\operatorname{Im} z$, מוגדר על ידי $\operatorname{Im} z = b$.

לפיכך, עבור כל מספר מרוכב z, נקבל

$$z = \operatorname{Re} z + (\operatorname{Im} z)i.$$

4.2 הגדרה: צמוד מרוכב, $\bar{z}$, ערך מוחלט, $|z|$

נניח ש־ $z \in \mathbf{C}$.

- הצמוד המרוכב של $z \in \mathbf{C}$, המסומן ב־ $\bar{z}$, מוגדר על ידי
$$\bar{z} = \operatorname{Re} z - (\operatorname{Im} z)i.$$
- הערך המוחלט של מספר מרוכב z, המסומן ב־ $|z|$, מוגדר על ידי
$$|z| = \sqrt{(\operatorname{Re} z)^2 + (\operatorname{Im} z)^2}.$$

4.3 דוגמה: חלק ממשי ומדומה, צמוד מרוכב, ערך מוחלט

נניח ש־ $z = 3 + 2i$. אזי

- $\operatorname{Re} z = 3$ ו־ $\operatorname{Im} z = 2$
- $\bar{z} = 3 - 2i$
- $|z| = \sqrt{3^2 + 2^2} = \sqrt{13}$.

זיהוי מספר מרוכב $z \in \mathbf{C}$ עם הזוג הסדור $(\operatorname{Re} z, \operatorname{Im} z) \in \mathbf{R}^2$ מזהה את $\mathbf{C}$ עם $\mathbf{R}^2$. נשים לב כי $\mathbf{C}$ הוא מרחב וקטורי מרוכב חד־ממדי, אך עדיין נוכל לחשוב על $\mathbf{C}$ (המזוהה עם $\mathbf{R}^2$) כמרחב וקטורי ממשי דו־ממדי.

הערך המוחלט של כל מספר מרוכב הוא מספר אי־שלילי. ליתר דיוק, אם $z \in \mathbf{C}$, אזי $|z|$ שווה למרחק מהראשית ב־ $\mathbf{R}^2$ לנקודה $(\operatorname{Re} z, \operatorname{Im} z) \in \mathbf{R}^2$.

החלקים הממשי והמדומה, הצמוד המרוכב והערך המוחלט מקיימים את התכונות במסקנה מרובת הסעיפים להלן.

עליכם לווודא ש־ $z = \bar{z}$ אם־ורק־אם z מספר ממשי.

4.4 *properties of complex numbers*

Suppose $w, z \in \mathbf{C}$. Then the following equalities and inequalities hold.

sum of z and $\overline{z}$
$z + \overline{z} = 2 \operatorname{Re} z$.

difference of z and $\overline{z}$
$z - \overline{z} = 2(\operatorname{Im} z)i$.

product of z and $\overline{z}$
$z\overline{z} = |z|^2$.

additivity and multiplicativity of complex conjugate
$\overline{w + z} = \overline{w} + \overline{z}$ and $\overline{wz} = \overline{w}\,\overline{z}$.

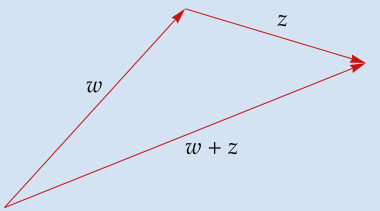
double complex conjugate
$\overline{\overline{z}} = z$.

real and imaginary parts are bounded by $|z|$
$|\operatorname{Re} z| \le |z|$ and $|\operatorname{Im} z| \le |z|$.

absolute value of the complex conjugate
$|\overline{z}| = |z|$.

multiplicativity of absolute value
$|wz| = |w|\,|z|$.

triangle inequality
$|w + z| \le |w| + |z|$.



Geometric interpretation of triangle inequality: The length of each side of a triangle is less than or equal to the sum of the lengths of the two other sides.

Proof Except for the last item above, the routine verifications of the assertions above are left to the reader. To verify the triangle inequality, we have

$$\begin{aligned}
|w + z|^2 &= (w + z)(\overline{w} + \overline{z}) \\
&= w\overline{w} + z\overline{z} + w\overline{z} + z\overline{w} \\
&= |w|^2 + |z|^2 + w\overline{z} + \overline{w\overline{z}} \\
&= |w|^2 + |z|^2 + 2\operatorname{Re}(w\overline{z}) \\
&\le |w|^2 + |z|^2 + 2|w\overline{z}| \\
&= |w|^2 + |z|^2 + 2|w|\,|z| \\
&= (|w| + |z|)^2.
\end{aligned}$$

Taking square roots now gives the desired inequality $|w + z| \le |w| + |z|$. ∎

See Exercise 2 for the reverse triangle inequality.

4.4 תכונות מספרים מרוכבים

נניח ש־ $w, z \in \mathbf{C}$. אזי מתקיימים השוויונות והאי־שוויונות הבאים.

סכום z ו־ $\overline{z}$
$$z + \overline{z} = 2 \operatorname{Re} z.$$

הפרש z ו־ $\overline{z}$
$$z - \overline{z} = 2(\operatorname{Im} z)i.$$

מכפלת z ו־ $\overline{z}$
$$z\overline{z} = |z|^2.$$

אדיטיביות וכפליות של צמוד מרוכב
$$\overline{w + z} = \overline{w} + \overline{z} \quad \text{ו־} \quad \overline{wz} = \overline{w}\,\overline{z}.$$

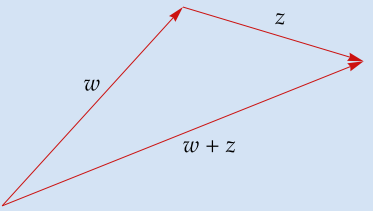
הצמדה כפולה של מרוכב
$$\overline{\overline{z}} = z.$$

החלקים הממשי והמדומה תחומים על ידי $|z|$
$$|\operatorname{Re} z| \le |z| \quad \text{ו־} \quad |\operatorname{Im} z| \le |z|.$$

ערך מוחלט של צמוד מרוכב
$$|\overline{z}| = |z|.$$

כפליות ערך מוחלט
$$|wz| = |w|\,|z|.$$

אי־שוויון המשולש
$$|w + z| \le |w| + |z|.$$



הוכחה מלבד הסעיף האחרון, ההוכחות השגרתיות לטענות לעיל מושארות לקוראים. נוכיח כעת את אי־שוויון המשולש. מתקיים

$$\begin{aligned}
|w + z|^2 &= (w + z)(\overline{w} + \overline{z}) \\
&= w\overline{w} + z\overline{z} + w\overline{z} + z\overline{w} \\
&= |w|^2 + |z|^2 + w\overline{z} + \overline{w\overline{z}} \\
&= |w|^2 + |z|^2 + 2 \operatorname{Re}(w\overline{z}) \\
&\le |w|^2 + |z|^2 + 2|w\overline{z}| \\
&= |w|^2 + |z|^2 + 2|w|\,|z| \\
&= (|w| + |z|)^2.
\end{aligned}$$

הוצאת שורש כעת נותנת את אי־שוויון הרצוי $|w + z| \le |w| + |z|$. ■

פרשנות גאומטרית לאי־שוויון המשולש: אורך כל צלע במשולש הוא קטן או שווה לסכום אורכן של שתי הצלעות האחרות.

ראו תרגיל 2 לאי־שוויון המשולש ההפוך.

Zeros of Polynomials

Recall that a function $p: \mathbf{F} \to \mathbf{F}$ is called a polynomial of degree m if there exist $a_0, \dots, a_m \in \mathbf{F}$ with $a_m \neq 0$ such that

$$p(z) = a_0 + a_1 z + \cdots + a_m z^m$$

for all $z \in \mathbf{F}$. A polynomial could have more than one degree if the representation of p in the form above were not unique. Our first task is to show that this cannot happen.

The solutions to the equation $p(z) = 0$ play a crucial role in the study of a polynomial $p \in \mathcal{P}(\mathbf{F})$. Thus these solutions have a special name.

4.5 definition: *zero of a polynomial*

A number $\lambda \in \mathbf{F}$ is called a *zero* (or *root*) of a polynomial $p \in \mathcal{P}(\mathbf{F})$ if

$$p(\lambda) = 0.$$

The next result is the key tool that we will use to show that the degree of a polynomial is unique.

4.6 *each zero of a polynomial corresponds to a degree-one factor*

Suppose m is a positive integer and $p \in \mathcal{P}(\mathbf{F})$ is a polynomial of degree m. Suppose $\lambda \in \mathbf{F}$. Then $p(\lambda) = 0$ if and only if there exists a polynomial $q \in \mathcal{P}(\mathbf{F})$ of degree $m - 1$ such that

$$p(z) = (z - \lambda)q(z)$$

for every $z \in \mathbf{F}$.

Proof First suppose $p(\lambda) = 0$. Let $a_0, a_1, \dots, a_m \in \mathbf{F}$ be such that

$$p(z) = a_0 + a_1 z + \cdots + a_m z^m$$

for all $z \in \mathbf{F}$. Then

4.7 $$p(z) = p(z) - p(\lambda) = a_1(z - \lambda) + \cdots + a_m(z^m - \lambda^m)$$

for all $z \in \mathbf{F}$. For each $k \in \{1, \dots, m\}$, the equation

$$z^k - \lambda^k = (z - \lambda) \sum_{j=1}^{k} \lambda^{j-1} z^{k-j}$$

shows that $z^k - \lambda^k$ equals $z - \lambda$ times some polynomial of degree $k - 1$. Thus 4.7 shows that p equals $z - \lambda$ times some polynomial of degree $m - 1$, as desired.

To prove the implication in the other direction, now suppose that there is a polynomial $q \in \mathcal{P}(\mathbf{F})$ such that $p(z) = (z - \lambda)q(z)$ for every $z \in \mathbf{F}$. Then $p(\lambda) = (\lambda - \lambda)q(\lambda) = 0$, as desired.

שורשי פולינומים

נזכיר כי פונקציה $p: \mathbf{F} \to \mathbf{F}$ נקראת פולינום ממעלה m אם קיימים $a_0, \dots, a_m \in \mathbf{F}$ ו־ $a_m \neq 0$ כך ש־

$$p(z) = a_0 + a_1 z + \cdots + a_m z^m$$

לכל $z \in \mathbf{F}$. אם ייצוגו של p באופן שתואר לעיל לא היה יחיד, הייתה יכולה להיות לפולינום יותר ממעלה אחת. משימתנו הראשונה היא לשלול אפשרות זו.

לפתרונות המשוואה $p(z) = 0$ תפקיד מרכזי בחקר הפולינום $p \in \mathcal{P}(\mathbf{F})$. לפיכך לפתרונות אלו הוענק שם מיוחד.

4.5 הגדרה: שורש פולינום

מספר $\lambda \in \mathbf{F}$ נקרא שורש של פולינום $p \in \mathcal{P}(\mathbf{F})$ אם

$$p(\lambda) = 0.$$

המסקנה הבאה משמשת ככלי המרכזי בבואנו להוכיח כי מעלת הפולינום יחידה.

4.6 כל שורש פולינום מתאים לגורם ממעלה ראשונה

נניח ש־ m מספר שלם חיובי ו־ $p \in \mathcal{P}(\mathbf{F})$ פולינום ממעלה m. נניח ש־ $\lambda \in \mathbf{F}$. אזי $p(\lambda) = 0$ אם־ורק־אם קיים פולינום $q \in \mathcal{P}(\mathbf{F})$ ממעלה $m - 1$ כך ש־

$$p(z) = (z - \lambda) q(z)$$

לכל $z \in \mathbf{F}$.

הוכחה ראשית נניח ש־ $p(\lambda) = 0$. יהיו $a_0, a_1, \dots, a_m \in \mathbf{F}$ כך ש־

$$p(z) = a_0 + a_1 z + \cdots + a_m z^m$$

לכל $z \in \mathbf{F}$. אזי

$$p(z) = p(z) - p(\lambda) = a_1(z - \lambda) + \cdots + a_m(z^m - \lambda^m) \tag{4.7}$$

לכל $z \in \mathbf{F}$. עבור כל $k \in \{1, \dots, m\}$, המשוואה

$$z^k - \lambda^k = (z - \lambda) \sum_{j=1}^{k} \lambda^{j-1} z^{k-j}$$

מראה ש־ $z^k - \lambda^k$ שווה ל־ $z - \lambda$ כפול פולינום כלשהו ממעלה $k - 1$. לפיכך לפי 4.7, p שווה ל־ $z - \lambda$ כפול פולינום כלשהו ממעלה $m - 1$, כפי שרצינו.

להוכחה בכיוון ההפוך, נניח כעת שקיים פולינום $q \in \mathcal{P}(\mathbf{F})$ כך ש־ $p(z) = (z - \lambda) q(z)$ לכל $z \in \mathbf{F}$. אזי $p(\lambda) = (\lambda - \lambda) q(\lambda) = 0$, כפי שרצינו. ∎

Now we can prove that polynomials do not have too many zeros.

4.8 *degree m implies at most m zeros*

Suppose m is a positive integer and $p \in \mathcal{P}(\mathbf{F})$ is a polynomial of degree m. Then p has at most m zeros in $\mathbf{F}$.

Proof We will use induction on m. The desired result holds if $m = 1$ because if $a_1 \neq 0$ then the polynomial $a_0 + a_1 z$ has only one zero (which equals $-a_0/a_1$). Thus assume that $m > 1$ and the desired result holds for $m - 1$.

If p has no zeros in $\mathbf{F}$, then the desired result holds and we are done. Thus suppose p has a zero $\lambda \in \mathbf{F}$. By 4.6, there is a polynomial $q \in \mathcal{P}(\mathbf{F})$ of degree $m - 1$ such that

$$p(z) = (z - \lambda)q(z)$$

for every $z \in \mathbf{F}$. Our induction hypothesis implies that q has at most $m - 1$ zeros in $\mathbf{F}$. The equation above shows that the zeros of p in $\mathbf{F}$ are exactly the zeros of q in $\mathbf{F}$ along with λ. Thus p has at most m zeros in $\mathbf{F}$. ∎

The result above implies that the coefficients of a polynomial are uniquely determined (because if a polynomial had two different sets of coefficients, then subtracting the two representations of the polynomial would give a polynomial with some nonzero coefficients but infinitely many zeros). In particular, the degree of a polynomial is uniquely defined.

Recall that the degree of the 0 polynomial is defined to be $-\infty$. When necessary, use the expected arithmetic with $-\infty$. For example, $-\infty < m$ and $-\infty + m = -\infty$ for every integer m.

The 0 polynomial is declared to have degree $-\infty$ so that exceptions are not needed for various reasonable results such as $\deg(pq) = \deg p + \deg q$.

Division Algorithm for Polynomials

If p and s are nonnegative integers, with $s \neq 0$, then there exist nonnegative integers q and r such that

$$p = sq + r$$

and $r < s$. Think of dividing p by s, getting quotient q with remainder r. Our next result gives an analogous result for polynomials. Thus the next result is often called the division algorithm for polynomials, although as stated here it is not really an algorithm, just a useful result.

The division algorithm for polynomials could be proved without using any linear algebra. However, as is appropriate for a linear algebra textbook, the proof given here uses linear algebra techniques and makes nice use of a basis of $\mathcal{P}_n(\mathbf{F})$, which is the $(n + 1)$-dimensional vector space of polynomials with coefficients in $\mathbf{F}$ and of degree at most n.

Think of the division algorithm for polynomials as giving a remainder polynomial r when the polynomial p is divided by the polynomial s.

כעת באפשרותנו להוכיח שלפולינומים מספר שורשים מוגבל.

4.8 מעלה m משמעותה m שורשים לכל היותר

נניח ש־ m מספר שלם חיובי ו־ $p \in \mathcal{P}(\mathbf{F})$ פולינום ממעלה m. אזי ל־ p יש לכל היותר m שורשים ב־ $\mathbf{F}$.

הוכחה נוכיח באמצעות אינדוקציה על m. אם $m = 1$ המסקנה מתקיימת, שהרי אם $a_1 \neq 0$ אזי לפולינום $a_0 + a_1 z$ שורש אחד בלבד (השווה ל־ $-a_0/a_1$). לפיכך נניח ש־ $m > 1$ והמסקנה מתקיימת עבור $m - 1$.

אם ל־ p אין שורשים ב־ $\mathbf{F}$, אזי המסקנה מתקיימת וההוכחה הסתיימה. לפיכך נניח שיש ל־ p שורש $\lambda \in \mathbf{F}$. לפי 4.6, קיים פולינום $q \in \mathcal{P}(\mathbf{F})$ ממעלה $m - 1$ כך שמתקיים

$$p(z) = (z - \lambda)q(z)$$

לכל $z \in \mathbf{F}$. לפי הנחת האינדוקציה, ל־ q יש לכל היותר $m-1$ שורשים ב־ $\mathbf{F}$. המשוואה לעיל מראה ששורשי p ב־ $\mathbf{F}$ הם בדיוק שורשי q ב־ $\mathbf{F}$ בצירוף λ. לפיכך ל־ p יש לכל היותר m שורשים ב־ $\mathbf{F}$. ∎

המסקנה לעיל גוררת שמקדמי הפולינום נקבעים באופן יחיד (שהרי אם לפולינום היו שתי קבוצות שונות של מקדמים, תוצאת חיסור ייצוג אחד מהאחר הייתה פולינום שחלק ממקדמיו שונים מאפס אך בעל אינסוף שורשים). בפרט, מעלת הפולינום מוגדרת באופן יחיד.

נזכיר כי מעלת פולינום האפס מוגדרת כ־ $-\infty$. בעת הצורך נבצע את החישוב הצפוי עם $-\infty$. לדוגמה, $-\infty < m$ ו־ $-\infty + m = -\infty$ לכל מספר שלם m.

לפולינום האפס הוגדרה המעלה $-\infty$ כך שלא ייזדרשו החרגות עבור מסקנות סבירות שונות כגון $\deg(pq) = \deg p + \deg q$.

אלגוריתם חלוקת פולינומים

אם p ו־ s מספרים שלמים אי־שליליים, ו־ $s \neq 0$, אזי קיימים מספרים שלמים אי־שליליים q ו־ r כך ש־

$$p = sq + r$$

וגם $r < s$. חשבו על תהליך שבו מחלקים p ב־ s, ומקבלים מנה q עם שארית r. המסקנה הבאה נותנת תוצאה אנלוגית לפולינומים. לפיכך, לעיתים קרובות, היא נקראת האלגוריתם לחלוקת פולינומים, אף שבצורה שבה נוסחה כאן, אין זה אלגוריתם של ממש כי אם תוצאה שימושית.

חשבו על האלגוריתם לחלוקת פולינומים כתהליך שבו מופק פולינום שארית r כתוצאה מחילוק הפולינום p בפולינום s.

את האלגוריתם לחלוקת פולינומים ניתן להוכיח ללא שימוש באלגברה ליניארית כלל. אולם, כיאה לספר לימוד אלגברה ליניארית, נציג כאן הוכחה שמשתמשת בטכניקות אלגברה ליניארית ואף עושה שימוש נאה בבסיס של $\mathcal{P}_n(\mathbf{F})$, שהוא המרחב הווקטורי ה־ $(n+1)$־ממדי של פולינומים עם מקדמים מ־ $\mathbf{F}$ ממעלה n לכל היותר.

4.9 *division algorithm for polynomials*

Suppose that $p, s \in \mathcal{P}(\mathbf{F})$, with $s \neq 0$. Then there exist unique polynomials $q, r \in \mathcal{P}(\mathbf{F})$ such that

$$p = sq + r$$

and $\deg r < \deg s$.

Proof Let $n = \deg p$ and let $m = \deg s$. If $n < m$, then take $q = 0$ and $r = p$ to get the desired equation $p = sq + r$ with $\deg r < \deg s$. Thus we now assume that $n \geq m$.

The list

4.10 $$1, z, \dots, z^{m-1}, s, zs, \dots, z^{n-m}s$$

is linearly independent in $\mathcal{P}_n(\mathbf{F})$ because each polynomial in this list has a different degree. Also, the list 4.10 has length $n + 1$, which equals $\dim \mathcal{P}_n(\mathbf{F})$. Hence 4.10 is a basis of $\mathcal{P}_n(\mathbf{F})$ [by 2.38].

Because $p \in \mathcal{P}_n(\mathbf{F})$ and 4.10 is a basis of $\mathcal{P}_n(\mathbf{F})$, there exist unique constants $a_0, a_1, \dots, a_{m-1} \in \mathbf{F}$ and $b_0, b_1, \dots, b_{n-m} \in \mathbf{F}$ such that

4.11 $$\begin{aligned} p &= a_0 + a_1 z + \cdots + a_{m-1} z^{m-1} + b_0 s + b_1 z s + \cdots + b_{n-m} z^{n-m} s \\ &= \underbrace{a_0 + a_1 z + \cdots + a_{m-1} z^{m-1}}_{r} + s(\underbrace{b_0 + b_1 z + \cdots + b_{n-m} z^{n-m}}_{q}). \end{aligned}$$

With r and q as defined above, we see that p can be written as $p = sq + r$ with $\deg r < \deg s$, as desired.

The uniqueness of $q, r \in \mathcal{P}(\mathbf{F})$ satisfying these conditions follows from the uniqueness of the constants $a_0, a_1, \dots, a_{m-1} \in \mathbf{F}$ and $b_0, b_1, \dots, b_{n-m} \in \mathbf{F}$ satisfying 4.11. ∎

Factorization of Polynomials over **C**

We have been handling polynomials with complex coefficients and polynomials with real coefficients simultaneously, letting $\mathbf{F}$ denote $\mathbf{R}$ or $\mathbf{C}$. Now we will see differences between these two cases. First we treat polynomials with complex coefficients. Then we will use those results to prove corresponding results for polynomials with real coefficients.

The fundamental theorem of algebra is an existence theorem. Its proof does not lead to a method for finding zeros. The quadratic formula gives the zeros explicitly for polynomials of degree 2. Similar but more complicated formulas exist for polynomials of degree 3 and 4. No such formulas exist for polynomials of degree 5 and above.

Our proof of the fundamental theorem of algebra implicitly uses the result that a continuous real-valued function on a closed disk in $\mathbf{R}^2$ attains a minimum value. A web search can lead you to several

4.9 אלגוריתם חלוקת פולינומים

נניח ש־ $p, s \in \mathcal{P}(\mathbf{F})$, ו־ $s \neq 0$. אזי קיימים פולינומים יחידים $q, r \in \mathcal{P}(\mathbf{F})$ כך ש־

$$p = sq + r$$

כאשר $\deg r < \deg s$.

הוכחה נגדיר $n = \deg p$ ו־ $m = \deg s$. אם $n < m$, אזי ניקח $q = 0$ ו־ $r = p$ ונקבל את המשוואה $p = sq + r$ יחד עם $\deg r < \deg s$. לפיכך נניח כעת ש־ $n \geq m$.

הרשימה

$$1, z, \dots, z^{m-1}, s, zs, \dots, z^{n-m}s \tag{4.10}$$

בלתי תלויה ליניארית ב־ $\mathcal{P}_n(\mathbf{F})$ שכן לכל פולינום ברשימה זו מעלה שונה. בנוסף, אורך הרשימה 4.10 הוא $n + 1$, מספר השווה ל־ $\dim \mathcal{P}_n(\mathbf{F})$. לכן 4.10 בסיס של $\mathcal{P}_n(\mathbf{F})$ [לפי 2.38].

מאחר ש־ $p \in \mathcal{P}_n(\mathbf{F})$ ו־ 4.10 בסיס של $\mathcal{P}_n(\mathbf{F})$, קיימים קבועים יחידים

$$a_0, a_1, \dots, a_{m-1} \in \mathbf{F} \quad \text{ו־} \quad b_0, b_1, \dots, b_{n-m} \in \mathbf{F}$$

כך ש־

$$\begin{aligned} p &= a_0 + a_1 z + \cdots + a_{m-1} z^{m-1} + b_0 s + b_1 z s + \cdots + b_{n-m} z^{n-m} s \\ &= \underbrace{a_0 + a_1 z + \cdots + a_{m-1} z^{m-1}}_{r} + s(\underbrace{b_0 + b_1 z + \cdots + b_{n-m} z^{n-m}}_{q}). \end{aligned} \tag{4.11}$$

בהינתן r ו־ q כפי שהוגדרו לעיל, אנו רואים כי ניתן לבטא את p כ־ $p = sq + r$ כאשר $\deg r < \deg s$, כפי שרצינו.

יחידותם של $q, r \in \mathcal{P}(\mathbf{F})$ המקיימים תנאים אלו נובעת מיחידותם של הקבועים $a_0, a_1, \dots, a_{m-1} \in \mathbf{F}$ ו־ $b_0, b_1, \dots, b_{n-m} \in \mathbf{F}$ המקיימים את 4.11. ■

פירוק לגורמים של פולינומים מעל C

המשפט היסודי של האלגברה הוא משפט על קיומם של שורשים. הוכחתו אינה מובילה לשיטה למציאתם. נוסחת השורשים לפתרון משוואה ריבועית מוצאת במפורש את השורשים עבור פולינומים ממעלה 2. נוסחאות דומות אך מורכבות יותר קיימות עבור פולינומים ממעלה 3 ו־ 4. לא קיימות נוסחאות כאלו עבור פולינומים ממעלה 5 ומעבר.

עסקנו עד כה בפולינומים עם מקדמים מרוכבים ובפולינומים עם מקדמים ממשיים במקביל, כש־ $\mathbf{F}$ מייצג $\mathbf{R}$ או $\mathbf{C}$. כעת נראה הבדלים בין שני מקרים אלו. תחילה נעסוק בפולינומים עם מקדמים מרוכבים. לאחר מכן, נשתמש במסקנות אלו להוכיח מסקנות תואמות עבור פולינומים עם מקדמים ממשיים.

ההוכחה שנביא למשפט היסודי של האלגברה עושה שימוש מובלע במשפט הקובע כי פונקציה ממשית רציפה המוגדרת מעל עיגול סגור ב־ $\mathbf{R}^2$ מקבלת ערך מינימום. חיפוש ברשת עשוי להוביל אתכם למספר הוכחות נוספות למשפט היסודי של האלגברה.

other proofs of the fundamental theorem of algebra. The proof using Liouville's theorem is particularly nice if you are comfortable with analytic functions. All proofs of the fundamental theorem of algebra need to use some analysis, because the result is not true if $\mathbf{C}$ is replaced, for example, with the set of numbers of the form $c + di$ where c, d are rational numbers.

4.12 *fundamental theorem of algebra, first version*

Every nonconstant polynomial with complex coefficients has a zero in $\mathbf{C}$.

Proof De Moivre's theorem, which you can prove using induction on k and the addition formulas for cosine and sine, states that if k is a positive integer and $\theta \in \mathbf{R}$, then

$$(\cos\theta + i\sin\theta)^k = \cos k\theta + i\sin k\theta.$$

Suppose $w \in \mathbf{C}$ and k is a positive integer. Using polar coordinates, we know that there exist $r \geq 0$ and $\theta \in \mathbf{R}$ such that

$$r(\cos\theta + i\sin\theta) = w.$$

De Moivre's theorem implies that

$$\left(r^{1/k}\left(\cos\tfrac{\theta}{k} + i\sin\tfrac{\theta}{k}\right)\right)^k = w.$$

Thus every complex number has a kth root, a fact that we will soon use.

Suppose p is a nonconstant polynomial with complex coefficients and highest-order nonzero term $c_m z^m$. Then $|p(z)| \to \infty$ as $|z| \to \infty$ (because $|p(z)|/|z^m| \to |c_m|$ as $|z| \to \infty$). Thus the continuous function $z \mapsto |p(z)|$ has a global minimum at some point $\zeta \in \mathbf{C}$. To show that $p(\zeta) = 0$, suppose that $p(\zeta) \neq 0$.

Define a new polynomial q by

$$q(z) = \frac{p(z+\zeta)}{p(\zeta)}.$$

The function $z \mapsto |q(z)|$ has a global minimum value of 1 at $z = 0$. Write

$$q(z) = 1 + a_k z^k + \cdots + a_m z^m,$$

where k is the smallest positive integer such that the coefficient of z^k is nonzero; in other words, $a_k \neq 0$.

Let $\beta \in \mathbf{C}$ be such that $\beta^k = -\dfrac{1}{a_k}$. There is a constant $c > 1$ such that if $t \in (0,1)$, then

$$\begin{aligned} |q(t\beta)| &\leq |1 + a_k t^k \beta^k| + t^{k+1}c \\ &= 1 - t^k(1 - tc). \end{aligned}$$

Thus taking t to be $1/(2c)$ in the inequality above, we have $|q(t\beta)| < 1$, which contradicts the assumption that the global minimum of $z \mapsto |q(z)|$ is 1. This contradiction implies that $p(\zeta) = 0$, showing that p has a zero, as desired. ∎

ההוכחה העושה שימוש במשפט לִיוּבִיל יפה במיוחד אם אתם חשים בנוח עם פונקציות אנליטיות. כל ההוכחות של המשפט היסודי של האלגברה חייבות לעשות שימוש מסוים באנליזה, מאחר שהתוצאה אינה נכונה אם $\mathbf{C}$ מוחלף, לדוגמה, בקבוצת המספרים מהצורה $c+di$ כאשר c,d רציונליים.

4.12 המשפט היסודי של האלגברה, גרסה ראשונה

לכל פולינום שאינו קבוע עם מקדמים מרוכבים יש שורש ב־ $\mathbf{C}$.

הוכחה משפט דה־מואבר, הניתן להוכחה באינדוקציה על k ובנוסחאות החיבור עבור הפונקציות סינוס וקוסינוס, קובע כי אם k מספר שלם חיובי, ו־ $\theta \in \mathbf{R}$, אזי

$$(\cos\theta + i\sin\theta)^k = \cos k\theta + i\sin k\theta.$$

נניח ש־ $w \in \mathbf{C}$ ו־ k מספר שלם חיובי. בשימוש בקואורדינטות קוטביות, אנו יודעים כי קיימים $r \geq 0$ ו־ $\theta \in \mathbf{R}$ כך ש־

$$r(\cos\theta + i\sin\theta) = w.$$

ממשפט דה־מואבר נובעת המשוואה

$$\Big(r^{1/k}\big(\cos\tfrac{\theta}{k} + i\sin\tfrac{\theta}{k}\big)\Big)^k = w.$$

לפיכך לכל מספר מרוכב יש שורש מסדר k, עובדה שבה נעשה שימוש בקרוב.

נניח ש־ p פולינום שאינו קבוע עם מקדמים מרוכבים, שאיברו השונה מאפס מהסדר הגבוה ביותר הוא $c_m z^m$. אזי, $|p(z)| \to \infty$ כאשר $|z| \to \infty$ (מאחר ש־ $|p(z)|/|z^m| \to |c_m|$ כאשר $|z| \to \infty$). לפיכך, לפונקציה הרציפה $z \mapsto |p(z)|$ יש מינימום גלובלי בנקודה כלשהי $\zeta \in \mathbf{C}$. כדי להראות ש־ $p(\zeta) = 0$, נניח ש־ $p(\zeta) \neq 0$.

נגדיר פולינום חדש q על ידי

$$q(z) = \frac{p(z+\zeta)}{p(\zeta)}.$$

הערך בנקודת המינימום הגלובלי של הפונקציה $z \mapsto |q(z)|$ הוא 1 עבור $z = 0$. נכתוב

$$q(z) = 1 + a_k z^k + \cdots + a_m z^m,$$

כאשר k המספר השלם החיובי הקטן ביותר כך שהמקדם של z^k שונה מאפס; במילים אחרות, $a_k \neq 0$.

יהי $\beta \in \mathbf{C}$ כך ש־ $\beta^k = -\dfrac{1}{a_k}$. קיים קבוע $c > 1$ כך שאם $t \in (0,1)$, אזי

$$\begin{aligned} |q(t\beta)| &\leq |1 + a_k t^k \beta^k| + t^{k+1} c \\ &= 1 - t^k(1 - tc). \end{aligned}$$

לפיכך אם נציב $t = 1/(2c)$ באי־השוויון לעיל, נקבל $|q(t\beta)| < 1$, בסתירה להנחה שהמינימום הגלובלי של $z \mapsto |q(z)|$ הוא 1. מסתירה זו נובע ש־ $p(\zeta) = 0$, ובכך הראינו כי ל־ p יש שורש, כפי שרצינו.

Computers can use clever numerical methods to find good approximations to the zeros of any polynomial, even when exact zeros cannot be found. For example, no one will ever give an exact formula for a zero of the polynomial p defined by

$$p(x) = x^5 - 5x^4 - 6x^3 + 17x^2 + 4x - 7.$$

However, a computer can find that the zeros of p are approximately the five numbers $-1.87,\ -0.74,\ 0.62,\ 1.47,\ 5.51$.

The first version of the fundamental theorem of algebra leads to the following factorization result for polynomials with complex coefficients. Note that in this factorization, the zeros of p are the numbers $\lambda_1, \ldots, \lambda_m$, which are the only values of z for which the right side of the equation in the next result equals 0.

4.13 *fundamental theorem of algebra, second version*

If $p \in \mathcal{P}(\mathbf{C})$ is a nonconstant polynomial, then p has a unique factorization (except for the order of the factors) of the form

$$p(z) = c(z - \lambda_1) \cdots (z - \lambda_m),$$

where $c, \lambda_1, \ldots, \lambda_m \in \mathbf{C}$.

Proof Let $p \in \mathcal{P}(\mathbf{C})$ and let $m = \deg p$. We will use induction on m. If $m = 1$, then the desired factorization exists and is unique. So assume that $m > 1$ and that the desired factorization exists and is unique for all polynomials of degree $m - 1$.

First we will show that the desired factorization of p exists. By the first version of the fundamental theorem of algebra (4.12), p has a zero $\lambda \in \mathbf{C}$. By 4.6, there is a polynomial q of degree $m - 1$ such that

$$p(z) = (z - \lambda)q(z)$$

for all $z \in \mathbf{C}$. Our induction hypothesis implies that q has the desired factorization, which when plugged into the equation above gives the desired factorization of p.

Now we turn to the question of uniqueness. The number c is uniquely determined as the coefficient of z^m in p. So we only need to show that except for the order, there is only one way to choose $\lambda_1, \ldots, \lambda_m$. If

$$(z - \lambda_1) \cdots (z - \lambda_m) = (z - \tau_1) \cdots (z - \tau_m)$$

for all $z \in \mathbf{C}$, then because the left side of the equation above equals 0 when $z = \lambda_1$, one of the τ's on the right side equals λ_1. Relabeling, we can assume that $\tau_1 = \lambda_1$. Now if $z \neq \lambda_1$, we can divide both sides of the equation above by $z - \lambda_1$, getting

$$(z - \lambda_2) \cdots (z - \lambda_m) = (z - \tau_2) \cdots (z - \tau_m)$$

for all $z \in \mathbf{C}$ except possibly $z = \lambda_1$. Actually the equation above holds for all $z \in \mathbf{C}$, because otherwise by subtracting the right side from the left side we would get a nonzero polynomial that has infinitely many zeros. The equation above and our induction hypothesis imply that except for the order, the λ's are the same as the τ's, completing the proof of uniqueness. ∎

בשימוש במחשב תוך הפעלת שיטות נומריות מתוחכמות נוכל למצוא קירובים טובים לשורשים של כל פולינום, גם כאשר לא ניתן למצוא את השורשים המדויקים. לדוגמה, איש לא יוכל לספק נוסחה מדויקת לשורש של הפולינום p המוגדר על ידי

$$p(x) = x^5 - 5x^4 - 6x^3 + 17x^2 + 4x - 7.$$

אולם, באמצעות מחשב נוכל למצוא שהשורשים של p הם בקירוב חמשת המספרים $-1.87,\ -0.74,\ 0.62,\ 1.47,\ 5.51$.

המשפט היסודי של האלגברה בגרסתו הראשונה מוביל לפירוק לגורמים הבא עבור פולינומים עם מקדמים מרוכבים. שימו לב כי בפירוק לגורמים זה, שורשי p הם המספרים $\lambda_1, \dots, \lambda_m$, שהם ערכי z היחידים שעבורם האגף הימני של המשוואה במסקנה להלן שווה 0.

4.13 המשפט היסודי של האלגברה, גרסה שנייה

אם $p \in \mathcal{P}(\mathbf{C})$ פולינום שאינו קבוע, אזי ל־ p פירוק לגורמים יחיד (עד כדי סדר הופעתם) מהצורה

$$p(z) = c(z - \lambda_1) \cdots (z - \lambda_m),$$

כאשר $c, \lambda_1, \dots, \lambda_m \in \mathbf{C}$.

הוכחה יהי $p \in \mathcal{P}(\mathbf{C})$ ויהי $m = \deg p$. נוכיח באמצעות אינדוקציה על m. אם $m = 1$, אזי הפירוק לגורמים קיים והוא יחיד כפי שרצינו. לכן נניח ש־ $m > 1$ והפירוק לגורמים הרצוי קיים באופן יחיד לכל הפולינומים ממעלה $m - 1$.

תחילה נראה את קיומם של הגורמים של p. לפי המשפט היסודי של האלגברה בגרסתו הראשונה (4.12), לפולינום p יש שורש $\lambda \in \mathbf{C}$. לפי 4.6, קיים פולינום q ממעלה $m - 1$ כך ש־

$$p(z) = (z - \lambda)q(z)$$

לכל $z \in \mathbf{C}$. הנחת האינדוקציה שלנו גוררת שעבור q הפירוק לגורמים שרצינו קיים, וכאשר נציב אותו במשוואה לעיל, נקבל את הפירוק לגורמים שרצינו עבור p.

כעת נפנה לשאלת היחידות. המספר c נקבע באופן יחיד כמקדם של z^m ב־ p. לכן כל שעלינו להראות הוא שמלבד סדר הופעת הגורמים, יש אפשרות אחת בלבד לבחירת $\lambda_1, \dots, \lambda_m$. אם

$$(z - \lambda_1) \cdots (z - \lambda_m) = (z - \tau_1) \cdots (z - \tau_m)$$

לכל $z \in \mathbf{C}$, אזי, מאחר שהאגף השמאלי של המשוואה לעיל שווה 0 כאשר $z = \lambda_1$, אחד ה־ τ־ים באגף הימני של המשוואה שווה ל־ λ_1. בסימון מחדש, נוכל להניח כי $\tau_1 = \lambda_1$. כעת אם $z \neq \lambda_1$, נוכל לחלק את שני אגפי המשוואה לעיל ב־ $z - \lambda_1$, ונקבל

$$(z - \lambda_2) \cdots (z - \lambda_m) = (z - \tau_2) \cdots (z - \tau_m)$$

לכל $z \in \mathbf{C}$ מלבד אולי עבור $z = \lambda_1$. למעשה המשוואה לעיל מתקיימת לכל $z \in \mathbf{C}$, שכן אחרת, אם נחסר את האגף הימני מהאגף השמאלי, נקבל פולינום השונה מאפס בעל אינסוף שורשים. מהמשוואה לעיל ומהנחת האינדוקציה נובעת המסקנה שמלבד סדר הופעת הגורמים, ערכי ה־ λ זהים לערכי ה־ τ, ובכך הושלמה הוכחת היחידות. ■

Factorization of Polynomials over **R**

A polynomial with real coefficients may have no real zeros. For example, the polynomial $1 + x^2$ has no real zeros.

To obtain a factorization theorem over **R**, we will use our factorization theorem over **C**. We begin with the next result.

The failure of the fundamental theorem of algebra for **R** *accounts for the differences between linear algebra on real and complex vector spaces, as we will see in later chapters.*

4.14 *polynomials with real coefficients have nonreal zeros in pairs*

Suppose $p \in \mathcal{P}(\mathbf{C})$ is a polynomial with real coefficients. If $\lambda \in \mathbf{C}$ is a zero of p, then so is $\overline{\lambda}$.

Proof Let

$$p(z) = a_0 + a_1 z + \cdots + a_m z^m,$$

where $a_0, \ldots, a_m$ are real numbers. Suppose $\lambda \in \mathbf{C}$ is a zero of p. Then

$$a_0 + a_1\lambda + \cdots + a_m\lambda^m = 0.$$

Take the complex conjugate of both sides of this equation, obtaining

$$a_0 + a_1\overline{\lambda} + \cdots + a_m\overline{\lambda}^m = 0,$$

where we have used basic properties of the complex conjugate (see 4.4). The equation above shows that $\overline{\lambda}$ is a zero of p. ∎

We want a factorization theorem for polynomials with real coefficients. We begin with the following result.

Think about the quadratic formula in connection with the result below.

4.15 *factorization of a quadratic polynomial*

Suppose $b, c \in \mathbf{R}$. Then there is a polynomial factorization of the form

$$x^2 + bx + c = (x - \lambda_1)(x - \lambda_2)$$

with $\lambda_1, \lambda_2 \in \mathbf{R}$ if and only if $b^2 \geq 4c$.

Proof Notice that

$$x^2 + bx + c = \left(x + \frac{b}{2}\right)^2 + \left(c - \frac{b^2}{4}\right).$$

First suppose $b^2 < 4c$. Then the right side of the equation above is positive for every $x \in \mathbf{R}$. Hence the polynomial $x^2 + bx + c$ has no real zeros and thus cannot be factored in the form $(x - \lambda_1)(x - \lambda_2)$ with $\lambda_1, \lambda_2 \in \mathbf{R}$.

The equation above is the basis of the technique called completing the square.

פירוק לגורמים של פולינומים מעל $\mathbf{R}$

היעדר תוקפו של המשפט היסודי של האלגברה עבור $\mathbf{R}$ מסביר את ההבדלים בין אלגברה ליניארית מעל מרחבים וקטוריים ממשיים ומרוכבים, כפי שנראה בפרקי ההמשך.

לפולינום עם מקדמים ממשיים, ייתכן שלא יהיו שורשים ממשיים. לדוגמה, לפולינום $1 + x^2$ אין שורשים ממשיים.

להשגת משפט פירוק לגורמים מעל $\mathbf{R}$, נשתמש במשפט שלנו לפירוק לגורמים מעל $\mathbf{C}$. נתחיל במסקנה להלן.

4.14 השורשים הלא־ממשיים של פולינומים עם מקדמים ממשיים באים בזוגות

נניח ש־ $p \in \mathcal{P}(\mathbf{C})$ הוא פולינום עם מקדמים ממשיים. אם $\lambda \in \mathbf{C}$ שורש של p, אזי גם $\overline{\lambda}$ שורש של p.

הוכחה יהי

$$p(z) = a_0 + a_1 z + \cdots + a_m z^m,$$

כאשר $a_0, \ldots, a_m$ מספרים ממשיים. נניח ש־ $\lambda \in \mathbf{C}$ שורש של p. אזי

$$a_0 + a_1\lambda + \cdots + a_m\lambda^m = 0.$$

ניקח את הצמוד המרוכב של שני אגפי המשוואה, וכתוצאה מכך נקבל

$$a_0 + a_1\overline{\lambda} + \cdots + a_m\overline{\lambda}^m = 0,$$

כאשר השתמשנו בתכונות הבסיסיות של הצמוד המרוכב (ראו 4.4). המשוואה לעיל מראה ש־ $\overline{\lambda}$ שורש של p. ■

חשבו על נוסחת השורשים לפתרון משוואה ריבועית בקשר למסקנה הבאה.

אנו מעוניינים במשפט לפירוק לגורמים עבור פולינומים עם מקדמים ממשיים. נפתח אם כן במסקנה הבאה.

4.15 פירוק לגורמים של פולינום ממעלה שנייה

נניח ש־ $b, c \in \mathbf{R}$. אזי קיים פירוק לגורמים של הפולינום מהצורה

$$x^2 + bx + c = (x - \lambda_1)(x - \lambda_2)$$

כאשר $\lambda_1, \lambda_2 \in \mathbf{R}$ אם־ורק־אם $b^2 \geq 4c$.

הוכחה נשים לב כי

$$x^2 + bx + c = \left(x + \frac{b}{2}\right)^2 + \left(c - \frac{b^2}{4}\right).$$

המשוואה לעיל היא הבסיס לטכניקה הנקראת השלמה לריבוע.

נניח תחילה ש־ $b^2 < 4c$. אזי האגף הימני של המשוואה חיובי לכל $x \in \mathbf{R}$. לכן לפולינום $x^2 + bx + c$ אין שורשים ממשיים, ולפיכך לא נוכל לפרק אותו לגורמים מהצורה $(x - \lambda_1)(x - \lambda_2)$ כאשר $\lambda_1, \lambda_2 \in \mathbf{R}$.

Conversely, now suppose $b^2 \ge 4c$. Then there is a real number d such that $d^2 = \frac{b^2}{4} - c$. From the displayed equation above, we have

$$\begin{aligned} x^2 + bx + c &= \left(x + \frac{b}{2}\right)^2 - d^2 \\ &= \left(x + \frac{b}{2} + d\right)\left(x + \frac{b}{2} - d\right), \end{aligned}$$

which gives the desired factorization. ∎

The next result gives a factorization of a polynomial over $\mathbf{R}$. The idea of the proof is to use the second version of the fundamental theorem of algebra (4.13), which gives a factorization of p as a polynomial with complex coefficients. Complex but nonreal zeros of p come in pairs; see 4.14. Thus if the factorization of p as an element of $\mathcal{P}(\mathbf{C})$ includes terms of the form $(x - \lambda)$ with λ a nonreal complex number, then $(x - \overline{\lambda})$ is also a term in the factorization. Multiplying together these two terms, we get

$$x^2 - 2(\operatorname{Re}\lambda)x + |\lambda|^2,$$

which is a quadratic term of the required form.

The idea sketched in the paragraph above almost provides a proof of the existence of our desired factorization. However, we need to be careful about one point. Suppose λ is a nonreal complex number and $(x - \lambda)$ is a term in the factorization of p as an element of $\mathcal{P}(\mathbf{C})$. We are guaranteed by 4.14 that $(x - \overline{\lambda})$ also appears as a term in the factorization, but 4.14 does not state that these two factors appear the same number of times, as needed to make the idea above work. However, the proof works around this point.

In the next result, either m or M may equal 0. The numbers $\lambda_1, \dots, \lambda_m$ are precisely the real zeros of p, for these are the only real values of x for which the right side of the equation in the next result equals 0.

4.16 *factorization of a polynomial over* $\mathbf{R}$

Suppose $p \in \mathcal{P}(\mathbf{R})$ is a nonconstant polynomial. Then p has a unique factorization (except for the order of the factors) of the form

$$p(x) = c(x - \lambda_1) \cdots (x - \lambda_m)(x^2 + b_1 x + c_1) \cdots (x^2 + b_M x + c_M),$$

where $c, \lambda_1, \dots, \lambda_m, b_1, \dots, b_M, c_1, \dots, c_M \in \mathbf{R}$, with $b_k^2 < 4c_k$ for each k.

Proof First we will prove that the desired factorization exists, and after that we will prove the uniqueness.

Think of p as an element of $\mathcal{P}(\mathbf{C})$. If all (complex) zeros of p are real, then we have the desired factorization by 4.13. Thus suppose p has a zero $\lambda \in \mathbf{C}$ with $\lambda \notin \mathbf{R}$. By 4.14, $\overline{\lambda}$ is a zero of p. Thus we can write

בכיוון ההפוך, כעת נניח ש־ $b^2 \geq 4c$. אזי, קיים מספר ממשי d כך ש־ $d^2 = \frac{b^2}{4} - c$. מהמשוואה שהצגנו לעיל, נקבל

$$\begin{aligned} x^2 + bx + c &= \left(x + \frac{b}{2}\right)^2 - d^2 \\ &= \left(x + \frac{b}{2} + d\right)\left(x + \frac{b}{2} - d\right), \end{aligned}$$

ובזאת קיבלנו את הפירוק לגורמים שרצינו. ■

מהמסקנה הבאה נקבל את הפירוק לגורמים מעל $\mathbf{R}$. הרעיון בהוכחה הוא להשתמש במשפט היסודי של האלגברה בגרסתו השנייה (4.13), שממנו מתקבל הפירוק לגורמים של p כפולינום עם מקדמים מרוכבים. השורשים המרוכבים הלא־ממשיים של p באים בזוגות; ראו 4.14. לפיכך, אם הפירוק לגורמים של p כאיבר ב־ $\mathcal{P}(\mathbf{C})$ מכיל גורם מהצורה $(x - \lambda)$ כאשר λ מספר מרוכב שאינו ממשי, אזי $(x - \overline{\lambda})$ גם הוא גורם בפירוק לגורמים. בהכפלת שני גורמים אלו זה בזה, נקבל

$$(x^2 - 2(\operatorname{Re}\lambda)x + |\lambda|^2),$$

שהוא ביטוי ריבועי בצורה הנדרשת.

הרעיון ששרטטנו בפסקה לעיל כמעט מספק את ההוכחה לקיומו של הפירוק לגורמים המבוקש. אך עלינו להיזהר בנקודה אחת. נניח ש־ λ הוא מספר מרוכב שאינו ממשי ו־ $(x - \lambda)$ הוא גורם בפירוק לגורמים של p כאיבר של $\mathcal{P}(\mathbf{C})$. 4.14 מבטיחה ש־ $(x - \overline{\lambda})$ מופיע גם הוא כגורם בפירוק לגורמים, אך 4.14 אינה קובעת שמספר המופעים של שני גורמים אלו זהה, מה שנדרש כדי שהרעיון לעיל יעבוד. אולם, בפועל ההוכחה עוקפת נקודה זו.

במסקנה הבאה ייתכן ש־ m או M שווה 0. המספרים $\lambda_1, \ldots, \lambda_m$ הם במדויק השורשים הממשיים של p, מאחר שהם הערכים הממשיים היחידים של x שעבורם האגף הימני של המשוואה במסקנה הבאה שווה 0.

4.16 פירוק לגורמים של פולינום מעל $\mathbf{R}$

נניח ש־ $p \in \mathcal{P}(\mathbf{R})$ הוא פולינום שאינו קבוע. אזי ל־ p יש פירוק לגורמים יחיד (עד כדי סדר הופעתם) מהצורה

$$p(x) = c(x - \lambda_1) \cdots (x - \lambda_m)(x^2 + b_1 x + c_1) \cdots (x^2 + b_M x + c_M),$$

כאשר $c, \lambda_1, \ldots, \lambda_m, b_1, \ldots, b_M, c_1, \ldots, c_M \in \mathbf{R}$ ומתקיים $b_k^2 < 4c_k$ לכל k.

הוכחה תחילה נוכיח שהפירוק לגורמים המבוקש אכן קיים, ולאחר מכן נוכיח שהוא יחיד.

נחשוב על p כעל איבר ב־ $\mathcal{P}(\mathbf{C})$. אם כל השורשים (המרוכבים) של p הם ממשיים, קיבלנו את הפירוק לגורמים הרצוי לפי 4.13. לפיכך, נניח כעת שלפולינום p יש שורש $\lambda \in \mathbf{C}$ כאשר $\lambda \notin \mathbf{R}$. לפי 4.14, $\overline{\lambda}$ גם הוא שורש של p. לפיכך נוכל לכתוב

$$\begin{aligned} p(x) &= (x-\lambda)(x-\overline{\lambda})q(x) \\ &= (x^2 - 2(\operatorname{Re}\lambda)x + |\lambda|^2)q(x) \end{aligned}$$

for some polynomial $q \in \mathcal{P}(\mathbf{C})$ of degree two less than the degree of p. If we can prove that q has real coefficients, then using induction on the degree of p completes the proof of the existence part of this result.

To prove that q has real coefficients, we solve the equation above for q, getting

$$q(x) = \frac{p(x)}{x^2 - 2(\operatorname{Re}\lambda)x + |\lambda|^2}$$

for all $x \in \mathbf{R}$. The equation above implies that $q(x) \in \mathbf{R}$ for all $x \in \mathbf{R}$. Writing

$$q(x) = a_0 + a_1 x + \cdots + a_{n-2}x^{n-2},$$

where $n = \deg p$ and $a_0, \ldots, a_{n-2} \in \mathbf{C}$, we thus have

$$0 = \operatorname{Im} q(x) = (\operatorname{Im} a_0) + (\operatorname{Im} a_1)x + \cdots + (\operatorname{Im} a_{n-2})x^{n-2}$$

for all $x \in \mathbf{R}$. This implies that $\operatorname{Im} a_0, \ldots, \operatorname{Im} a_{n-2}$ all equal 0 (by 4.8). Thus all coefficients of q are real, as desired. Hence the desired factorization exists.

Now we turn to the question of uniqueness of our factorization. A factor of p of the form $x^2 + b_k x + c_k$ with $b_k^2 < 4c_k$ can be uniquely written as $(x-\lambda_k)(x-\overline{\lambda_k})$ with $\lambda_k \in \mathbf{C}$. A moment's thought shows that two different factorizations of p as an element of $\mathcal{P}(\mathbf{R})$ would lead to two different factorizations of p as an element of $\mathcal{P}(\mathbf{C})$, contradicting 4.13. ∎

Exercises 4

1 Suppose $w, z \in \mathbf{C}$. Verify the following equalities and inequalities.

(a) $z + \overline{z} = 2\operatorname{Re} z$

(b) $z - \overline{z} = 2(\operatorname{Im} z)i$

(c) $z\overline{z} = |z|^2$

(d) $\overline{w+z} = \overline{w} + \overline{z}$ and $\overline{wz} = \overline{w}\,\overline{z}$

(e) $\overline{\overline{z}} = z$

(f) $|\operatorname{Re} z| \le |z|$ and $|\operatorname{Im} z| \le |z|$

(g) $|\overline{z}| = |z|$

(h) $|wz| = |w|\,|z|$

The results above are the parts of 4.4 that were left to the reader.

2 Prove that if $w, z \in \mathbf{C}$, then $\big||w| - |z|\big| \le |w - z|$.

The inequality above is called the ***reverse triangle inequality****.*

$$\begin{aligned} p(x) &= (x-\lambda)(x-\overline{\lambda})q(x) \\ &= (x^2 - 2(\operatorname{Re}\lambda)x + |\lambda|^2)q(x) \end{aligned}$$

עבור פולינום $q \in \mathcal{P}(\mathbf{C})$ כלשהו ממעלה הפחותה בשתיים ממעלת p. אם ביכולתנו להוכיח כי ל־ q מקדמים ממשיים, אזי שימוש באינדוקציה על מעלת p ישלים את החלק של הוכחת הקיום של מסקנה זו.

כדי להוכיח שלפולינום q מקדמים ממשיים, נפתור את המשוואה לעיל עבור q ונקבל

$$q(x) = \frac{p(x)}{x^2 - 2(\operatorname{Re}\lambda)x + |\lambda|^2}$$

לכל $x \in \mathbf{R}$. מהמשוואה לעיל נובע כי $q(x) \in \mathbf{R}$ לכל $x \in \mathbf{R}$. בכתיבת

$$q(x) = a_0 + a_1 x + \cdots + a_{n-2}x^{n-2},$$

כאשר $n = \deg p$ ו־ $a_0, \ldots, a_{n-2} \in \mathbf{C}$, לפיכך מתקיים

$$0 = \operatorname{Im} q(x) = (\operatorname{Im} a_0) + (\operatorname{Im} a_1)x + \cdots + (\operatorname{Im} a_{n-2})x^{n-2}$$

לכל $x \in \mathbf{R}$. מכאן נובע ש־ $\operatorname{Im} a_0, \ldots, \operatorname{Im} a_{n-2}$ כולם שווים 0 (לפי 4.8). לפיכך כל מקדמי q ממשיים, כפי שרצינו. לכן הפירוק לגורמים המבוקש קיים.

כעת נעבור לשאלת היחידות של הפירוק לגורמים. גורם של p מהצורה $x^2 + b_k x + c_k$ כאשר $b_k^2 < 4c_k$ ניתן לבטא באופן יחיד בצורה $(x-\lambda_k)(x-\overline{\lambda_k})$ עבור $\lambda_k \in \mathbf{C}$. במעט מחשבה נראה ששני פירוקים לגורמים של p כאיבר ב־ $\mathcal{P}(\mathbf{R})$ יובילו לשני פירוקים לגורמים של p כאיבר ב־ $\mathcal{P}(\mathbf{C})$, בסתירה ל־ 4.13. ■

תרגילים ד

1 נניח ש־ $w, z \in \mathbf{C}$. ודאו את השוויונות והאי־שוויונות הבאים.

א. $z + \overline{z} = 2\operatorname{Re} z$

ב. $z - \overline{z} = 2(\operatorname{Im} z)i$

ג. $z\overline{z} = |z|^2$

ד. $\overline{w+z} = \overline{w} + \overline{z}$ ו־ $\overline{wz} = \overline{w}\,\overline{z}$

ה. $\overline{\overline{z}} = z$

ו. $|\operatorname{Re} z| \le |z|$ ו־ $|\operatorname{Im} z| \le |z|$

ז. $|\overline{z}| = |z|$

ח. $|wz| = |w|\,|z|$

התוצאות לעיל הן החלקים ב־ 4.4 שהושארו לקוראים.

2 הוכיחו שאם $w, z \in \mathbf{C}$, אזי $\big||w| - |z|\big| \le |w - z|$.

אי־שוויון זה נקרא **אי־שוויון המשולש ההפוך**.

3 Suppose V is a complex vector space and $\varphi \in V'$. Define $\sigma : V \to \mathbf{R}$ by $\sigma(v) = \operatorname{Re} \varphi(v)$ for each $v \in V$. Show that

$$\varphi(v) = \sigma(v) - i\sigma(iv)$$

for all $v \in V$.

4 Suppose m is a positive integer. Is the set

$$\{0\} \cup \{p \in \mathcal{P}(\mathbf{F}) : \deg p = m\}$$

a subspace of $\mathcal{P}(\mathbf{F})$?

5 Is the set

$$\{0\} \cup \{p \in \mathcal{P}(\mathbf{F}) : \deg p \text{ is even}\}$$

a subspace of $\mathcal{P}(\mathbf{F})$?

6 Suppose that m and n are positive integers with $m \le n$, and suppose $\lambda_1, \dots, \lambda_m \in \mathbf{F}$. Prove that there exists a polynomial $p \in \mathcal{P}(\mathbf{F})$ with $\deg p = n$ such that $0 = p(\lambda_1) = \cdots = p(\lambda_m)$ and such that p has no other zeros.

7 Suppose that m is a nonnegative integer, $z_1, \dots, z_{m+1}$ are distinct elements of $\mathbf{F}$, and $w_1, \dots, w_{m+1} \in \mathbf{F}$. Prove that there exists a unique polynomial $p \in \mathcal{P}_m(\mathbf{F})$ such that

$$p(z_k) = w_k$$

for each $k = 1, \dots, m + 1$.

This result can be proved without using linear algebra. However, try to find the clearer, shorter proof that uses some linear algebra.

8 Suppose $p \in \mathcal{P}(\mathbf{C})$ has degree m. Prove that p has m distinct zeros if and only if p and its derivative p' have no zeros in common.

9 Prove that every polynomial of odd degree with real coefficients has a real zero.

10 For $p \in \mathcal{P}(\mathbf{R})$, define $Tp : \mathbf{R} \to \mathbf{R}$ by

$$(Tp)(x) = \begin{cases} \dfrac{p(x) - p(3)}{x - 3} & \text{if } x \ne 3, \\ p'(3) & \text{if } x = 3 \end{cases}$$

for each $x \in \mathbf{R}$. Show that $Tp \in \mathcal{P}(\mathbf{R})$ for every polynomial $p \in \mathcal{P}(\mathbf{R})$ and also show that $T : \mathcal{P}(\mathbf{R}) \to \mathcal{P}(\mathbf{R})$ is a linear map.

11 Suppose $p \in \mathcal{P}(\mathbf{C})$. Define $q : \mathbf{C} \to \mathbf{C}$ by

$$q(z) = p(z)\,\overline{p(\overline{z})}.$$

Prove that q is a polynomial with real coefficients.

3 נניח ש־ V מרחב וקטורי מרוכב ו־ $\varphi \in V'$. הגדירו $\sigma : V \to \mathbf{R}$ על ידי $\sigma(v) = \operatorname{Re} \varphi(v)$ לכל $v \in V$. הראו ש־

$$\varphi(v) = \sigma(v) - i\sigma(iv)$$

לכל $v \in V$.

4 נניח ש־ m מספר שלם חיובי. האם הקבוצה

$$\{0\} \cup \{p \in \mathcal{P}(\mathbf{F}) : \deg p = m\}$$

תת־מרחב של $\mathcal{P}(\mathbf{F})$?

5 האם הקבוצה

$$\{0\} \cup \{p \in \mathcal{P}(\mathbf{F}) : \deg p \text{ זוגי}\}$$

תת־מרחב של $\mathcal{P}(\mathbf{F})$?

6 נניח ש־ m ו־ n מספרים שלמים חיוביים ו־ $m \le n$, ונניח ש־ $\lambda_1, \dots, \lambda_m \in \mathbf{F}$. הוכיחו שקיים פולינום $p \in \mathcal{P}(\mathbf{F})$ עם $\deg p = n$ כך ש־ $0 = p(\lambda_1) = \cdots = p(\lambda_m)$ וכך שאין שורשים אחרים ל־ p.

7 נניח ש־ m מספר שלם אי־שלילי, $z_1, \dots, z_{m+1}$ איברים של $\mathbf{F}$ השונים זה מזה, ו־ $w_1, \dots, w_{m+1} \in \mathbf{F}$. הוכיחו שקיים פולינום יחיד $p \in \mathcal{P}_m(\mathbf{F})$ כך ש־

$$p(z_k) = w_k$$

לכל $k = 1, \dots, m+1$.

אף שניתן להוכיח פסקנה זו ללא שימוש באלגברה ליניארית, נסו למצוא את ההוכחה הברורה והקצרה יותר העושה שימוש באלגברה ליניארית.

8 נניח ש־ $p \in \mathcal{P}(\mathbf{C})$ ממעלה m. הוכיחו כי ל־ p יש m שורשים שונים אם־ורק־אם אין ל־ p ונגזרתו p' שורשים משותפים.

9 הוכיחו כי לכל פולינום ממעלה אי־זוגית עם מקדמים ממשיים יש שורש ממשי.

10 עבור $p \in \mathcal{P}(\mathbf{R})$, נגדיר $Tp : \mathbf{R} \to \mathbf{R}$ על ידי

$$(Tp)(x) = \begin{cases} \dfrac{p(x) - p(3)}{x - 3} & \text{if } x \ne 3, \\ p'(3) & \text{if } x = 3 \end{cases}$$

לכל $x \in \mathbf{R}$. הראו ש־ $Tp \in \mathcal{P}(\mathbf{R})$ לכל פולינום $p \in \mathcal{P}(\mathbf{R})$ ובנוסף, הראו כי $T : \mathcal{P}(\mathbf{R}) \to \mathcal{P}(\mathbf{R})$ העתקה ליניארית.

11 נניח ש־ $p \in \mathcal{P}(\mathbf{C})$. נגדיר $q : \mathbf{C} \to \mathbf{C}$ על ידי

$$q(z) = p(z)\,\overline{p(\overline{z})}.$$

הוכיחו כי q פולינום עם מקדמים ממשיים.

12 Suppose m is a nonnegative integer and $p \in \mathcal{P}_m(\mathbf{C})$ is such that there are distinct real numbers $x_0, x_1, \dots, x_m$ with $p(x_k) \in \mathbf{R}$ for each $k = 0, 1, \dots, m$. Prove that all coefficients of p are real.

13 Suppose $p \in \mathcal{P}(\mathbf{F})$ with $p \neq 0$. Let $U = \{pq : q \in \mathcal{P}(\mathbf{F})\}$.

(a) Show that $\dim \mathcal{P}(\mathbf{F})/U = \deg p$.

(b) Find a basis of $\mathcal{P}(\mathbf{F})/U$.

14 Suppose $p, q \in \mathcal{P}(\mathbf{C})$ are nonconstant polynomials with no zeros in common. Let $m = \deg p$ and $n = \deg q$. Use linear algebra as outlined below in (a)–(c) to prove that there exist $r \in \mathcal{P}_{n-1}(\mathbf{C})$ and $s \in \mathcal{P}_{m-1}(\mathbf{C})$ such that

$$rp + sq = 1.$$

(a) Define $T : \mathcal{P}_{n-1}(\mathbf{C}) \times \mathcal{P}_{m-1}(\mathbf{C}) \to \mathcal{P}_{m+n-1}(\mathbf{C})$ by

$$T(r, s) = rp + sq.$$

Show that the linear map T is injective.

(b) Show that the linear map T in (a) is surjective.

(c) Use (b) to conclude that there exist $r \in \mathcal{P}_{n-1}(\mathbf{C})$ and $s \in \mathcal{P}_{m-1}(\mathbf{C})$ such that $rp + sq = 1$.

12 נניח ש־ m מספר שלם אי־שלילי ו־ $p \in \mathcal{P}_m(\mathbf{C})$ כך שקיימים מספרים ממשיים שונים $x_0, x_1, \ldots, x_m$ המקיימים $p(x_k) \in \mathbf{R}$ לכל $k = 0, 1, \ldots, m$. הוכיחו שכל מקדמי p ממשיים.

13 נניח ש־ $p \in \mathcal{P}(\mathbf{F})$ כאשר $p \neq 0$. יהי $U = \{pq : q \in \mathcal{P}(\mathbf{F})\}$.

א. הראו ש־ $\dim \mathcal{P}(\mathbf{F})/U = \deg p$.

ב. מצאו בסיס של $\mathcal{P}(\mathbf{F})/U$.

14 נניח ש־ $p, q \in \mathcal{P}(\mathbf{C})$ פולינומים שאינם קבועים, ללא שורשים משותפים. יהי $m = \deg p$ ו־ $n = \deg q$. השתמשו באלגברה ליניארית כמפורט בסעיפים (א)–(ג) להלן כדי להוכיח כי קיימים $r \in \mathcal{P}_{n-1}(\mathbf{C})$ ו־ $s \in \mathcal{P}_{m-1}(\mathbf{C})$ כך ש־

$$rp + sq = 1.$$

א. הגדירו $T : \mathcal{P}_{n-1}(\mathbf{C}) \times \mathcal{P}_{m-1}(\mathbf{C}) \to \mathcal{P}_{m+n-1}(\mathbf{C})$ על ידי

$$T(r, s) = rp + sq.$$

הראו שההעתקה הליניארית T חד־חד־ערכית.

ב. הראו שההעתקה הליניארית T בסעיף (א) היא עַל.

ג. השתמשו בסעיף (ב) כדי להסיק כי קיימים $r \in \mathcal{P}_{n-1}(\mathbf{C})$ ו־ $s \in \mathcal{P}_{m-1}(\mathbf{C})$ כך ש־ $rp + sq = 1$.

אלגברה ליניארית בדרך הישר, מהדורה רביעית, מחבר: שלדון אקסלר
Linear Algebra Done Right, fourth edition, by Sheldon Axler

Chapter 5

Eigenvalues and Eigenvectors

Linear maps from one vector space to another vector space were the objects of study in Chapter 3. Now we begin our investigation of operators, which are linear maps from a vector space to itself. Their study constitutes the most important part of linear algebra.

To learn about an operator, we might try restricting it to a smaller subspace. Asking for that restriction to be an operator will lead us to the notion of invariant subspaces. Each one-dimensional invariant subspace arises from a vector that the operator maps into a scalar multiple of the vector. This path will lead us to eigenvectors and eigenvalues.

We will then prove one of the most important results in linear algebra: every operator on a finite-dimensional nonzero complex vector space has an eigenvalue. This result will allow us to show that for each operator on a finite-dimensional complex vector space, there is a basis of the vector space with respect to which the matrix of the operator has at least almost half its entries equal to 0.

standing assumptions for this chapter

- **F** denotes **R** or **C**.
- V denotes a vector space over **F**.

Hans-Peter Postel CC BY

Statue of Leonardo of Pisa (1170–1250, approximate dates), also known as Fibonacci. Exercise 21 in Section 5D shows how linear algebra can be used to find the explicit formula for the Fibonacci sequence shown on the front cover.

פרק ה

ערכים עצמיים ווקטורים עצמיים

העתקות ליניאריות ממרחב וקטורי אחד למרחב וקטורי אחר היו האובייקטים שאותם למדנו בפרק ג. כעת אנו מתחילים לעסוק בחקר אופרטורים — שהם העתקות ליניאריות ממרחב וקטורי לעצמו. מחקרם מהווה את החלק החשוב ביותר באלגברה ליניארית.

כדי ללמוד על אופרטור, ננסה לצמצם אותו לתת־מרחב קטן יותר. הדרישה כי תחת צמצום זה הוא יישאר אופרטור תוביל אותנו למושג תתי־מרחב שמורים. כל תת־מרחב שמור חד־ממדי נוצר מווקטור שאותו האופרטור מעתיק אל כפולה סקלרית שלו. נתיב זה יוביל אותנו אל וקטורים עצמיים וערכים עצמיים.

לאחר מכן נוכיח את אחת המסקנות החשובות ביותר באלגברה ליניארית: לכל אופרטור מעל מרחב וקטורי מרוכב מממד סופי השונה מאפס, יש ערך עצמי. מסקנה זו תאפשר לנו להראות כי לכל אופרטור מעל מרחב וקטורי מרוכב מממד סופי, קיים בסיס למרחב, אשר ביחס אליו, לפחות כמעט מחצית מאיברי מטריצת האופרטור שווים 0.

הנחות המוצא לפרק זה

- $\mathbf{F}$ מסמן $\mathbf{R}$ או $\mathbf{C}$.
- V מסמן מרחב וקטורי מעל $\mathbf{F}$.

Hans-Peter Postel CC BY

פסלו של לאונרדו מפיזה (1250–1170, תאריכים משוערים), הידוע גם בשם פִיבּוֹנַאצִ'י. תרגיל 21 ביחידה 4ה מראה איך ניתן להשתמש באלגברה ליניארית למציאת הנוסחה המפורשת של סדרת פיבונאצ'י המופיעה על עטיפת הספר.

אלגברה ליניארית בדרך הישר, מהדורה רביעית, מחבר: שלדון אקסלר
Linear Algebra Done Right, fourth edition, by Sheldon Axler

5A Invariant Subspaces

Eigenvalues

5.1 definition: *operator*

A linear map from a vector space to itself is called an *operator*.

Suppose $T \in \mathcal{L}(V)$. If $m \geq 2$ and

$$V = V_1 \oplus \cdots \oplus V_m,$$

where each V_k is a nonzero subspace of V, then to understand the behavior of T we only need to understand the behavior of each $T|_{V_k}$; here $T|_{V_k}$ denotes the restriction of T to the smaller domain V_k. Dealing with $T|_{V_k}$ should be easier than dealing with T because V_k is a smaller vector space than V.

Recall that we defined the notation $\mathcal{L}(V)$ to mean $\mathcal{L}(V, V)$.

However, if we intend to apply tools useful in the study of operators (such as taking powers), then we have a problem: $T|_{V_k}$ may not map V_k into itself; in other words, $T|_{V_k}$ may not be an operator on V_k. Thus we are led to consider only decompositions of V of the form above in which T maps each V_k into itself. Hence we now give a name to subspaces of V that get mapped into themselves by T.

5.2 definition: *invariant subspace*

Suppose $T \in \mathcal{L}(V)$. A subspace U of V is called *invariant* under T if $Tu \in U$ for every $u \in U$.

Thus U is invariant under T if $T|_U$ is an operator on U.

5.3 example: *subspace invariant under differentiation operator*

Suppose that $T \in \mathcal{L}(\mathcal{P}(\mathbf{R}))$ is defined by $Tp = p'$. Then $\mathcal{P}_4(\mathbf{R})$, which is a subspace of $\mathcal{P}(\mathbf{R})$, is invariant under T because if $p \in \mathcal{P}(\mathbf{R})$ has degree at most 4, then p' also has degree at most 4.

5.4 example: *four invariant subspaces, not necessarily all different*

If $T \in \mathcal{L}(V)$, then the following subspaces of V are all invariant under T.

- $\{0\}$ The subspace $\{0\}$ is invariant under T because if $u \in \{0\}$, then $u = 0$ and hence $Tu = 0 \in \{0\}$.
- V The subspace V is invariant under T because if $u \in V$, then $Tu \in V$.
- $\operatorname{null} T$ The subspace $\operatorname{null} T$ is invariant under T because if $u \in \operatorname{null} T$, then $Tu = 0$, and hence $Tu \in \operatorname{null} T$.
- $\operatorname{range} T$ The subspace $\operatorname{range} T$ is invariant under T because if $u \in \operatorname{range} T$, then $Tu \in \operatorname{range} T$.

ה1 תתי־מרחב שמורים

ערכים עצמיים

5.1 הגדרה: אופרטור

העתקה ליניארית ממרחב וקטורי לעצמו נקראת אופרטור.

נניח ש־ $T \in \mathcal{L}(V)$. אם $m \geq 2$ ו־
$V = V_1 \oplus \cdots \oplus V_m$,
כאשר כל V_k הוא תת־מרחב השונה מאפס של V, אזי להבנת התנהגות T די שנבין את התנהגות כל $T|_{V_k}$; כאן $T|_{V_k}$ מסמן צמצום T לתחום קטן יותר V_k. העיסוק ב־ $T|_{V_k}$ אמור להיות קל יותר מהעיסוק ב־ T מאחר ש־ V_k מרחב וקטורי קטן יותר מ־ V.

זכרו שהגדרנו את הסימון $\mathcal{L}(V)$ שמשמעותו $\mathcal{L}(V, V)$.

אולם, אם במהלך חקר האופרטורים נפעיל כלים שימושיים (דוגמת העלאה בחזקה), אזי ניתקל בבעיה: ייתכן ש־ $T|_{V_k}$ לא תעתיק את V_k לעצמו; ובמילים אחרות, ייתכן ש־ $T|_{V_k}$ לא יהיה אופרטור מעל V_k. לפיכך נבחר לעסוק רק בפירוקים של V מהצורה לעיל, כאלו שבהם T מעתיקה את כל V_k לעצמו. לכן נעניק שם לתתי־מרחב של V המועתקים לעצמם על ידי T.

5.2 הגדרה: תת־מרחב שמור

נניח ש־ $T \in \mathcal{L}(V)$. תת־מרחב U של V נקרא שמור תחת T אם $Tu \in U$ לכל $u \in U$.

לפיכך U שמור תחת T אם $T|_U$ אופרטור מעל U.

5.3 דוגמה: תת־מרחב שמור תחת אופרטור גזירה

נניח ש־ $T \in \mathcal{L}(\mathcal{P}(\mathbf{R}))$ מוגדר על ידי $Tp = p'$. אזי $\mathcal{P}_4(\mathbf{R})$, שהוא תת־מרחב של $\mathcal{P}(\mathbf{R})$, שמור תחת T מאחר שאם מעלתו של $p \in \mathcal{P}(\mathbf{R})$ היא לכל היותר 4, אזי גם מעלתו של p' היא לכל היותר 4.

5.4 דוגמה: ארבעה תתי־מרחב שמורים שאינם בהכרח שונים

אם $T \in \mathcal{L}(V)$, אזי תתי־המרחב של V להלן, כולם שמורים תחת T.

$\{0\}$ התת־מרחב $\{0\}$ שמור תחת T מאחר שאם $u \in \{0\}$, אזי $u = 0$ ולכן $Tu = 0 \in \{0\}$.

V התת־מרחב V שמור תחת T מאחר שאם $u \in V$, אזי $Tu \in V$.

$\operatorname{null} T$ התת־מרחב $\operatorname{null} T$ שמור תחת T מאחר שאם $u \in \operatorname{null} T$, אזי $Tu = 0$, ולכן $Tu \in \operatorname{null} T$.

$\operatorname{range} T$ התת־מרחב $\operatorname{range} T$ שמור תחת T מאחר שאם $u \in \operatorname{range} T$, אזי $Tu \in \operatorname{range} T$.

Must an operator $T \in \mathcal{L}(V)$ have any invariant subspaces other than $\{0\}$ and V? Later we will see that this question has an affirmative answer if V is finite-dimensional and $\dim V > 1$ (for $\mathbf{F} = \mathbf{C}$) or $\dim V > 2$ (for $\mathbf{F} = \mathbf{R}$); see 5.19 and Exercise 29 in Section 5B.

The previous example noted that $\operatorname{null} T$ and $\operatorname{range} T$ are invariant under T. However, these subspaces do not necessarily provide easy answers to the question above about the existence of invariant subspaces other than $\{0\}$ and V, because $\operatorname{null} T$ may equal $\{0\}$ and $\operatorname{range} T$ may equal V (this happens when T is invertible).

We will return later to a deeper study of invariant subspaces. Now we turn to an investigation of the simplest possible nontrivial invariant subspaces—invariant subspaces of dimension one.

Take any $v \in V$ with $v \neq 0$ and let U equal the set of all scalar multiples of v:

$$U = \{\lambda v : \lambda \in \mathbf{F}\} = \operatorname{span}(v).$$

Then U is a one-dimensional subspace of V (and every one-dimensional subspace of V is of this form for an appropriate choice of v). If U is invariant under an operator $T \in \mathcal{L}(V)$, then $Tv \in U$, and hence there is a scalar $\lambda \in \mathbf{F}$ such that

$$Tv = \lambda v.$$

Conversely, if $Tv = \lambda v$ for some $\lambda \in \mathbf{F}$, then $\operatorname{span}(v)$ is a one-dimensional subspace of V invariant under T.

The equation $Tv = \lambda v$, which we have just seen is intimately connected with one-dimensional invariant subspaces, is important enough that the scalars λ and vectors v satisfying it are given special names.

5.5 definition: *eigenvalue*

Suppose $T \in \mathcal{L}(V)$. A number $\lambda \in \mathbf{F}$ is called an *eigenvalue* of T if there exists $v \in V$ such that $v \neq 0$ and $Tv = \lambda v$.

*The word **eigenvalue** is half-German, half-English. The German prefix **eigen** means "own" in the sense of characterizing an intrinsic property.*

In the definition above, we require that $v \neq 0$ because every scalar $\lambda \in \mathbf{F}$ satisfies $T0 = \lambda 0$.

The comments above show that V has a one-dimensional subspace invariant under T if and only if T has an eigenvalue.

5.6 example: *eigenvalue*

Define an operator $T \in \mathcal{L}(\mathbf{F}^3)$ by

$$T(x, y, z) = (7x + 3z, 3x + 6y + 9z, -6y)$$

for $(x, y, z) \in \mathbf{F}^3$. Then $T(3, 1, -1) = (18, 6, -6) = 6(3, 1, -1)$. Thus 6 is an eigenvalue of T.

האם בהכרח לאופרטור $T \in \mathcal{L}(V)$ יש אי אלו תתי־מרחב שמורים מלבד $\{0\}$ ו־ V? בהמשך נראה שלשאלה זו תשובה חיובית אם V מממד סופי ו־ $\dim V > 1$ (עבור $\mathbf{F} = \mathbf{C}$) או $\dim V > 2$ (עבור $\mathbf{F} = \mathbf{R}$); ראו 5.19 ותרגיל 29 ביחידה הזו.

הדוגמה הקודמת ציינה ש־ $\operatorname{null} T$ ו־ $\operatorname{range} T$ שמורים תחת T. אולם תתי־מרחב אלו אינם בהכרח מספקים בקלות תשובות לשאלה לעיל בנוגע לקיומם של תתי־מרחב שמורים השונים מ־ $\{0\}$ ו־ V, מאחר שייתכן ש־ $\operatorname{null} T$ שווה ל־ $\{0\}$ ו־ $\operatorname{range} T$ שווה ל־ V (זה קורה כאשר T הפיך).

בהמשך נחזור לחקר מעמיק יותר של תתי־מרחב שמורים. לעת עתה נפנה לחקירת תתי־המרחב השמורים הלא־טריוויאליים הפשוטים ביותר—תתי־מרחב שמורים חד־ממדיים.

ניקח $v \in V$ כלשהו עם $v \neq 0$ ויהי U שווה לקבוצת כל הכפולות בסקלר של v:

$$U = \{\lambda v : \lambda \in \mathbf{F}\} = \operatorname{span}(v).$$

אזי U תת־מרחב חד־ממדי של V (וכך כל תת־מרחב חד־ממדי של V הוא מהצורה הזו עבור בחירה מתאימה של v). אם U שמור תחת אופרטור $T \in \mathcal{L}(V)$, אזי $Tv \in U$, ולכן קיים סקלר $\lambda \in \mathbf{F}$ כך ש־

$$Tv = \lambda v.$$

ולהיפך, אם $Tv = \lambda v$ עבור $\lambda \in \mathbf{F}$ כלשהו, אזי $\operatorname{span}(v)$ הוא תת־מרחב חד־ממדי של V השמור תחת T.

המשוואה $Tv = \lambda v$, שראינו זה עתה קשורה באופן הדוק לתתי־מרחב חד־ממדיים שמורים, חשיבותה רבה כל כך שלסקלרים λ ולווקטורים v המקיימים אותה הוענקו שמות מיוחדים.

5.5 הגדרה: ערך עצמי

נניח ש־ $T \in \mathcal{L}(V)$. מספר $\lambda \in \mathbf{F}$ נקרא ערך עצמי של T אם קיים $v \in V$ כך ש־ $v \neq 0$ וגם $Tv = \lambda v$.

ערך עצמי באנגלית נקרא **eigenvalue**, מילה שמחציתה גרמנית ומחציתה אנגלית. הקידומת הגרמנית **eigen** משמעותה "עצמי" במובן של אפיון מהותי פנימי.

בהגדרה לעיל דרשנו ש־ $v \neq 0$ מאחר שכל סקלר $\lambda \in \mathbf{F}$ מקיים $T0 = \lambda 0$.

מההערות לעיל אנו למדים כי ל־ V יש תת־מרחב חד־ממדי שמור תחת T אם־ורק־אם יש ל־ T ערך עצמי.

5.6 דוגמה: ערך עצמי

נגדיר אופרטור $T \in \mathcal{L}(\mathbf{F}^3)$ על ידי

$$T(x, y, z) = (7x + 3z, 3x + 6y + 9z, -6y)$$

עבור $(x, y, z) \in \mathbf{F}^3$. אזי $T(3, 1, -1) = (18, 6, -6) = 6(3, 1, -1)$. לפיכך 6 ערך עצמי של T.

The equivalences in the next result, along with many deep results in linear algebra, are valid only in the context of finite-dimensional vector spaces.

5.7 *equivalent conditions to be an eigenvalue*

Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and $\lambda \in \mathbf{F}$. Then the following are equivalent.

(a) λ is an eigenvalue of T.

(b) $T - \lambda I$ is not injective.

(c) $T - \lambda I$ is not surjective.

(d) $T - \lambda I$ is not invertible.

Reminder: $I \in \mathcal{L}(V)$ is the identity operator. Thus $Iv = v$ for all $v \in V$.

Proof Conditions (a) and (b) are equivalent because the equation $Tv = \lambda v$ is equivalent to the equation $(T - \lambda I)v = 0$. Conditions (b), (c), and (d) are equivalent by 3.65. ∎

5.8 definition: *eigenvector*

Suppose $T \in \mathcal{L}(V)$ and $\lambda \in \mathbf{F}$ is an eigenvalue of T. A vector $v \in V$ is called an *eigenvector* of T corresponding to λ if $v \neq 0$ and $Tv = \lambda v$.

In other words, a nonzero vector $v \in V$ is an eigenvector of an operator $T \in \mathcal{L}(V)$ if and only if Tv is a scalar multiple of v. Because $Tv = \lambda v$ if and only if $(T - \lambda I)v = 0$, a vector $v \in V$ with $v \neq 0$ is an eigenvector of T corresponding to λ if and only if $v \in \operatorname{null}(T - \lambda I)$.

5.9 example: *eigenvalues and eigenvectors*

Suppose $T \in \mathcal{L}(\mathbf{F}^2)$ is defined by $T(w, z) = (-z, w)$.

(a) First consider the case $\mathbf{F} = \mathbf{R}$. Then T is a counterclockwise rotation by 90° about the origin in $\mathbf{R}^2$. An operator has an eigenvalue if and only if there exists a nonzero vector in its domain that gets sent by the operator to a scalar multiple of itself. A 90° counterclockwise rotation of a nonzero vector in $\mathbf{R}^2$ cannot equal a scalar multiple of itself. Conclusion: if $\mathbf{F} = \mathbf{R}$, then T has no eigenvalues (and thus has no eigenvectors).

(b) Now consider the case $\mathbf{F} = \mathbf{C}$. To find eigenvalues of T, we must find the scalars λ such that $T(w, z) = \lambda(w, z)$ has some solution other than $w = z = 0$. The equation $T(w, z) = \lambda(w, z)$ is equivalent to the simultaneous equations

$$-z = \lambda w, \quad w = \lambda z. \tag{5.10}$$

Substituting the value for w given by the second equation into the first equation gives

$$-z = \lambda^2 z.$$

התנאים השקולים במסקנה להלן, יחד עם מסקנות עמוקות רבות באלגברה ליניארית, תקפים רק בהקשר של מרחבים וקטוריים מממד סופי.

5.7 תנאים שקולים להיותו ערך עצמי

נניח ש־ V מממד סופי, $T \in \mathcal{L}(V)$, ו־ $\lambda \in \mathbf{F}$. אזי התנאים הבאים שקולים.

תזכורת: $I \in \mathcal{L}(V)$ הוא אופרטור הזהות. לפיכך $Iv = v$ לכל $v \in V$.

א. λ הוא ערך עצמי של T.

ב. $T - \lambda I$ אינו חד־חד־ערכי.

ג. $T - \lambda I$ אינו על.

ד. $T - \lambda I$ אינו הפיך.

הוכחה התנאים (א) ו־ (ב) שקולים, מאחר שהמשוואה $Tv = \lambda v$ שקולה למשוואה $(T - \lambda I)v = 0$. התנאים (ב), (ג), ו־ (ד) שקולים לפי 3.65. ■

5.8 הגדרה: וקטור עצמי

נניח ש־ $T \in \mathcal{L}(V)$ ו־ $\lambda \in \mathbf{F}$ ערך עצמי של T. וקטור $v \in V$ נקרא וקטור עצמי של T השייך ל־ λ אם $v \neq 0$ וגם $Tv = \lambda v$.

במילים אחרות, וקטור $v \in V$ השונה מאפס הוא וקטור עצמי של האופרטור $T \in \mathcal{L}(V)$ אם־ורק־אם Tv הוא כפולת סקלר של v. מאחר ש־ $Tv = \lambda v$ אם־ורק־אם $(T - \lambda I)v = 0$, וקטור $v \in V$ כאשר $v \neq 0$ הוא וקטור עצמי של T השייך ל־ λ אם־ורק־אם $v \in \operatorname{null}(T - \lambda I)$.

5.9 דוגמה: ערכים עצמיים וקטורים עצמיים

נניח ש־ $T \in \mathcal{L}(\mathbf{F}^2)$ מוגדר על ידי $T(w, z) = (-z, w)$.

א. תחילה נבחן את המקרה שבו $\mathbf{F} = \mathbf{R}$. אזי T הוא סיבוב נגד כיוון השעון של $90°$ סביב ראשית הצירים ב־ $\mathbf{R}^2$. לאופרטור יש ערך עצמי אם־ורק־אם קיים וקטור השונה מאפס בתחומו, הנשלח על ידי האופרטור לכפולת סקלר של עצמו. סיבוב $90°$ נגד כיוון השעון של וקטור השונה מאפס ב־ $\mathbf{R}^2$ אינו יכול להיות שווה לכפולה בסקלר של עצמו. מסקנה: אם $\mathbf{F} = \mathbf{R}$, אזי ל־ T אין ערכים עצמיים (ולפיכך גם לא וקטורים עצמיים).

ב. כעת נבחן את המקרה שבו $\mathbf{F} = \mathbf{C}$. כדי למצוא ערכים עצמיים של T, עלינו למצוא סקלרים λ כך שעבור $T(w, z) = \lambda(w, z)$ יש פתרונות השונים מ־ $w = z = 0$. המשוואה $T(w, z) = \lambda(w, z)$ שקולה לצמד המשוואות

$$-z = \lambda w, \quad w = \lambda z. \tag{5.10}$$

בהצבת ערכו של w הנתון על ידי המשוואה השנייה במשוואה הראשונה, נקבל

$$-z = \lambda^2 z.$$

Now z cannot equal 0 [otherwise 5.10 implies that $w = 0$; we are looking for solutions to 5.10 such that (w, z) is not the 0 vector], so the equation above leads to the equation

$$-1 = \lambda^2.$$

The solutions to this equation are $\lambda = i$ and $\lambda = -i$.

You can verify that i and $-i$ are eigenvalues of T. Indeed, the eigenvectors corresponding to the eigenvalue i are the vectors of the form $(w, -wi)$, with $w \in \mathbf{C}$ and $w \neq 0$. Furthermore, the eigenvectors corresponding to the eigenvalue $-i$ are the vectors of the form (w, wi), with $w \in \mathbf{C}$ and $w \neq 0$.

In the next proof, we again use the equivalence

$$Tv = \lambda v \iff (T - \lambda I)v = 0.$$

5.11 *linearly independent eigenvectors*

Suppose $T \in \mathcal{L}(V)$. Then every list of eigenvectors of T corresponding to distinct eigenvalues of T is linearly independent.

Proof Suppose the desired result is false. Then there exists a smallest positive integer m such that there exists a linearly dependent list $v_1, \dots, v_m$ of eigenvectors of T corresponding to distinct eigenvalues $\lambda_1, \dots, \lambda_m$ of T (note that $m \geq 2$ because an eigenvector is, by definition, nonzero). Thus there exist $a_1, \dots, a_m \in \mathbf{F}$, none of which are 0 (because of the minimality of m), such that

$$a_1 v_1 + \cdots + a_m v_m = 0.$$

Apply $T - \lambda_m I$ to both sides of the equation above, getting

$$a_1(\lambda_1 - \lambda_m)v_1 + \cdots + a_{m-1}(\lambda_{m-1} - \lambda_m)v_{m-1} = 0.$$

Because the eigenvalues $\lambda_1, \dots, \lambda_m$ are distinct, none of the coefficients above equal 0. Thus $v_1, \dots, v_{m-1}$ is a linearly dependent list of $m - 1$ eigenvectors of T corresponding to distinct eigenvalues, contradicting the minimality of m. This contradiction completes the proof. ∎

The result above leads to a short proof of the result below, which puts an upper bound on the number of distinct eigenvalues that an operator can have.

5.12 *operator cannot have more eigenvalues than dimension of vector space*

Suppose V is finite-dimensional. Then each operator on V has at most $\dim V$ distinct eigenvalues.

Proof Let $T \in \mathcal{L}(V)$. Suppose $\lambda_1, \dots, \lambda_m$ are distinct eigenvalues of T. Let $v_1, \dots, v_m$ be corresponding eigenvectors. Then 5.11 implies that the list $v_1, \dots, v_m$ is linearly independent. Thus $m \leq \dim V$ (see 2.22), as desired. ∎

כעת z אינו יכול להיות שווה 0 [אחרת 5.10 גורר ש־ $w = 0$; ואנו מחפשים פתרונות עבור 5.10 כך ש־ (w, z) אינו וקטור האפס], ולכן המשוואה לעיל מובילה אותנו למשוואה הבאה

$$-1 = \lambda^2.$$

הפתרונות עבור המשוואה הזו הם $\lambda = i$ ו־ $\lambda = -i$.

תוכלו לוודא ש־ i ו־ $-i$ ערכים עצמיים של T. אכן, הווקטורים העצמיים השייכים לערך העצמי i הם הווקטורים מהצורה $(w, -wi)$, כאשר $w \in \mathbf{C}$ ו־ $w \neq 0$. בנוסף, הווקטורים העצמיים השייכים לערך העצמי $-i$ הם הווקטורים מהצורה (w, wi), כאשר $w \in \mathbf{C}$ ו־ $w \neq 0$.

בהוכחה הבאה, נעשה שוב שימוש בתנאי השקול

$$Tv = \lambda v \iff (T - \lambda I)v = 0.$$

5.11 וקטורים עצמיים בלתי תלויים ליניארית

נניח ש־ $T \in \mathcal{L}(V)$. אזי כל רשימת וקטורים עצמיים של T השייכים לערכים עצמיים שונים של T בלתי תלויה ליניארית.

הוכחה נניח שהמסקנה אינה נכונה. אזי קיים מספר שלם חיובי m הקטן ביותר כך שקיימת רשימה תלויה ליניארית $v_1, \dots, v_m$ של וקטורים עצמיים של T השייכים לערכים עצמיים שונים של T (נשים לב כי $m \geq 2$ מאחר שווקטור עצמי, בהגדרה, שונה מאפס). לפיכך קיימים $a_1, \dots, a_m \in \mathbf{F}$, שאף אחד מהם אינו 0 (לאור המינימליות של m), כך ש־

$$a_1 v_1 + \cdots + a_m v_m = 0.$$

נפעיל את ההעתקה $T - \lambda_m I$ על שני אגפי המשוואה לעיל, ונקבל

$$a_1(\lambda_1 - \lambda_m)v_1 + \cdots + a_{m-1}(\lambda_{m-1} - \lambda_m)v_{m-1} = 0.$$

מאחר שהערכים העצמיים $\lambda_1, \dots, \lambda_m$ שונים זה מזה, אף אחד מהמקדמים אינו שווה 0. לפיכך הרשימה $v_1, \dots, v_{m-1}$ היא רשימה תלויה ליניארית המכילה $m - 1$ וקטורים עצמיים של T השייכים לערכים עצמיים שונים, בסתירה למינימליות של m. סתירה זו משלימה את ההוכחה. ■

המסקנה לעיל מובילה להוכחה קצרה למסקנה להלן, אשר קובעת חסם עליון למספר הערכים העצמיים השונים שיכולים להיות לאופרטור.

5.12 מספר הערכים העצמיים של אופרטור לא יעלה על ממד המרחב הווקטורי

נניח ש־ V מממד סופי. אזי לכל אופרטור מעל V יש לכל היותר $\dim V$ ערכים עצמיים שונים.

הוכחה יהי $T \in \mathcal{L}(V)$. נניח ש־ $\lambda_1, \dots, \lambda_m$ ערכים עצמיים שונים של T. יהיו $v_1, \dots, v_m$ וקטורים עצמיים השייכים לערכים עצמיים אלו. אזי 5.11 גוררת שהרשימה $v_1, \dots, v_m$ בלתי תלויה ליניארית. לפיכך $m \leq \dim V$ (ראו 2.22), כפי שרצינו. ■

Polynomials Applied to Operators

The main reason that a richer theory exists for operators (which map a vector space into itself) than for more general linear maps is that operators can be raised to powers. In this subsection we define that notion and the concept of applying a polynomial to an operator. This concept will be the key tool that we use in the next section when we prove that every operator on a nonzero finite-dimensional complex vector space has an eigenvalue.

If T is an operator, then TT makes sense (see 3.7) and is also an operator on the same vector space as T. We usually write T^2 instead of TT. More generally, we have the following definition of T^m.

5.13 notation: T^m

Suppose $T \in \mathcal{L}(V)$ and m is a positive integer.

- $T^m \in \mathcal{L}(V)$ is defined by $T^m = \underbrace{T \cdots T}_{m \text{ times}}$.
- T^0 is defined to be the identity operator I on V.
- If T is invertible with inverse T^{-1}, then $T^{-m} \in \mathcal{L}(V)$ is defined by

$$T^{-m} = (T^{-1})^m.$$

You should verify that if T is an operator, then

$$T^m T^n = T^{m+n} \quad \text{and} \quad (T^m)^n = T^{mn},$$

where m and n are arbitrary integers if T is invertible and are nonnegative integers if T is not invertible.

Having defined powers of an operator, we can now define what it means to apply a polynomial to an operator.

5.14 notation: $p(T)$

Suppose $T \in \mathcal{L}(V)$ and $p \in \mathcal{P}(\mathbf{F})$ is a polynomial given by

$$p(z) = a_0 + a_1 z + a_2 z^2 + \cdots + a_m z^m$$

for all $z \in \mathbf{F}$. Then $p(T)$ is the operator on V defined by

$$p(T) = a_0 I + a_1 T + a_2 T^2 + \cdots + a_m T^m.$$

This is a new use of the symbol p because we are applying p to operators, not just elements of $\mathbf{F}$. The idea here is that to evaluate $p(T)$, we simply replace z with T in the expression defining p. Note that the constant term a_0 in $p(z)$ becomes the operator $a_0 I$ (which is a reasonable choice because $a_0 = a_0 z^0$ and thus we should replace a_0 with $a_0 T^0$, which equals $a_0 I$).

יישום פולינומים על אופרטורים

הסיבה העיקרית לכך שקיימת תאוריה עשירה יותר עבור אופרטורים (המעתיקים מרחב וקטורי אל עצמו) מזו הקיימת עבור העתקות ליניאריות כלליות יותר, היא שניתן להעלות אופרטורים בחזקה. בתת־יחידה זו נגדיר מושג זה ואת רעיון הפעלת פולינום על אופרטור. רעיון זה יהיה הכלי המרכזי שבו נשתמש ביחידה הבאה, כאשר נוכיח כי לכל אופרטור מעל מרחב וקטורי מרוכב מממד סופי שאינו מרחב האפס, קיים ערך עצמי.

אם T אופרטור, אזי יש משמעות ל־ TT (ראו 3.7), וגם הוא, בדומה ל־ T, אופרטור מעל אותו מרחב וקטורי. בדרך כלל נרשום T^2 במקום TT. באופן כללי יותר, נתונה להלן ההגדרה עבור T^m.

5.13 סימון: T^m

נניח ש־ $T \in \mathcal{L}(V)$ ו־ m מספר שלם חיובי.

- $T^m \in \mathcal{L}(V)$ מוגדר על ידי $T^m = \underbrace{T \cdots T}_{m \text{ פעמים}}$.
- T^0 מוגדר כאופרטור הזהות I מעל V.
- אם T הפיך וההופכי שלו הוא T^{-1}, אזי $T^{-m} \in \mathcal{L}(V)$ מוגדר על ידי $T^{-m} = (T^{-1})^m$.

עליכם לוודא שאם T אופרטור, אזי

$$(T^m)^n = T^{mn}, \quad \text{וגם} \quad T^m T^n = T^{m+n},$$

כאשר m ו־ n מספרים שלמים שרירותיים אם T הפיך ומספרים שלמים אי־שליליים אם T אינו הפיך.

משהגדרנו חזקות עבור אופרטורים, נוכל כעת להגדיר את משמעות הפעלת פולינום על אופרטור.

5.14 סימון: $p(T)$

נניח ש־ $T \in \mathcal{L}(V)$ ו־ $p \in \mathcal{P}(\mathbf{F})$ פולינום הנתון על ידי

$$p(z) = a_0 + a_1 z + a_2 z^2 + \cdots + a_m z^m$$

לכל $z \in \mathbf{F}$. אזי $p(T)$ הוא אופרטור מעל V המוגדר על ידי

$$p(T) = a_0 I + a_1 T + a_2 T^2 + \cdots + a_m T^m.$$

זהו שימוש חדש בסימון p שכן אנו מפעילים את p על אופרטורים, לא רק על איברים ב־ $\mathbf{F}$. הרעיון כאן הוא שכדי לחשב את $p(T)$, אנו פשוט מחליפים את z ב־ T בביטוי שבו אנו מגדירים את p. נשים לב כי האיבר הקבוע a_0 ב־ $p(z)$ הופך לאופרטור $a_0 I$ (שזו בחירה הגיונית מאחר ש־ $a_0 = a_0 z^0$ ולפיכך עלינו להחליף את a_0 ב־ $a_0 T^0$, השווה ל־ $a_0 I$).

5.15 example: *a polynomial applied to the differentiation operator*

Suppose $D \in \mathcal{L}(\mathcal{P}(\mathbf{R}))$ is the differentiation operator defined by $Dq = q'$ and p is the polynomial defined by $p(x) = 7 - 3x + 5x^2$. Then $p(D) = 7I - 3D + 5D^2$. Thus

$$(p(D))q = 7q - 3q' + 5q''$$

for every $q \in \mathcal{P}(\mathbf{R})$.

If we fix an operator $T \in \mathcal{L}(V)$, then the function from $\mathcal{P}(\mathbf{F})$ to $\mathcal{L}(V)$ given by $p \mapsto p(T)$ is linear, as you should verify.

5.16 definition: *product of polynomials*

If $p, q \in \mathcal{P}(\mathbf{F})$, then $pq \in \mathcal{P}(\mathbf{F})$ is the polynomial defined by

$$(pq)(z) = p(z)q(z)$$

for all $z \in \mathbf{F}$.

The order does not matter in taking products of polynomials of a single operator, as shown by (b) in the next result.

5.17 *multiplicative properties*

Suppose $p, q \in \mathcal{P}(\mathbf{F})$ and $T \in \mathcal{L}(V)$. Then

(a) $(pq)(T) = p(T)q(T)$;

(b) $p(T)q(T) = q(T)p(T)$.

Informal proof: When a product of polynomials is expanded using the distributive property, it does not matter whether the symbol is z or T.

Proof

(a) Suppose $p(z) = \sum_{j=0}^{m} a_j z^j$ and $q(z) = \sum_{k=0}^{n} b_k z^k$ for all $z \in \mathbf{F}$. Then

$$(pq)(z) = \sum_{j=0}^{m} \sum_{k=0}^{n} a_j b_k z^{j+k}.$$

Thus

$$\begin{aligned}(pq)(T) &= \sum_{j=0}^{m} \sum_{k=0}^{n} a_j b_k T^{j+k} \\ &= \Big(\sum_{j=0}^{m} a_j T^j\Big)\Big(\sum_{k=0}^{n} b_k T^k\Big) \\ &= p(T)q(T).\end{aligned}$$

(b) Using (a) twice, we have $p(T)q(T) = (pq)(T) = (qp)(T) = q(T)p(T)$. ∎

5.15 דוגמה: פולינום מופעל על אופרטור גזירה

נניח ש־ $D \in \mathcal{L}(\mathcal{P}(\mathbf{R}))$ הוא אופרטור הגזירה המוגדר על ידי $Dq = q'$ ו־ p פולינום המוגדר על ידי $p(x) = 7 - 3x + 5x^2$. אזי $p(D) = 7I - 3D + 5D^2$. לפיכך

$$(p(D))q = 7q - 3q' + 5q''$$

לכל $q \in \mathcal{P}(\mathbf{R})$.

אם נְקַבֵּעַ אופרטור $T \in \mathcal{L}(V)$, אזי הפונקציה מ־ $\mathcal{P}(\mathbf{F})$ אל $\mathcal{L}(V)$ הנתונה על ידי $p \mapsto p(T)$ ליניארית, כפי שעליכם לוודא.

5.16 הגדרה: מכפלת פולינומים

אם $p, q \in \mathcal{P}(\mathbf{F})$, אזי $pq \in \mathcal{P}(\mathbf{F})$ הוא הפולינום המוגדר על ידי

$$(pq)(z) = p(z)q(z)$$

לכל $z \in \mathbf{F}$.

כשמכפילים פולינומים של אותו אופרטור אין חשיבות לסדר הגורמים, כפי שנראה בסעיף (ב) במסקנה הבאה.

5.17 תכונות מכפלה

נניח ש־ $p, q \in \mathcal{P}(\mathbf{F})$ ו־ $T \in \mathcal{L}(V)$. אזי

א. $(pq)(T) = p(T)q(T)$

ב. $p(T)q(T) = q(T)p(T)$.

הוכחה בלתי רשמית: אם נפתח סוגריים למכפלת פולינומים תוך שימוש בתכונת הדיסטריבוטיביות, לא יהיה הבדל אם הסמל הוא z או T.

הוכחה

א. נניח ש־ $p(z) = \sum_{j=0}^{m} a_j z^j$ וגם $q(z) = \sum_{k=0}^{n} b_k z^k$ לכל $z \in \mathbf{F}$. אזי

$$(pq)(z) = \sum_{j=0}^{m} \sum_{k=0}^{n} a_j b_k z^{j+k}.$$

לפיכך

$$\begin{aligned}(pq)(T) &= \sum_{j=0}^{m} \sum_{k=0}^{n} a_j b_k T^{j+k} \\ &= \left(\sum_{j=0}^{m} a_j T^j\right)\left(\sum_{k=0}^{n} b_k T^k\right) \\ &= p(T)q(T).\end{aligned}$$

ב. בהפעלת (א) פעמיים, נקבל $p(T)q(T) = (pq)(T) = (qp)(T) = q(T)p(T)$. ■

We observed earlier that if $T \in \mathcal{L}(V)$, then the subspaces $\operatorname{null} T$ and $\operatorname{range} T$ are invariant under T (see 5.4). Now we show that the null space and the range of every polynomial of T are also invariant under T.

5.18 *null space and range of $p(T)$ are invariant under T*

Suppose $T \in \mathcal{L}(V)$ and $p \in \mathcal{P}(\mathbf{F})$. Then $\operatorname{null} p(T)$ and $\operatorname{range} p(T)$ are invariant under T.

Proof Suppose $u \in \operatorname{null} p(T)$. Then $p(T)u = 0$. Thus

$$(p(T))(Tu) = (p(T)\,T)(u) = (T\,p(T))(u) = T(p(T)u) = T(0) = 0.$$

Hence $Tu \in \operatorname{null} p(T)$. Thus $\operatorname{null} p(T)$ is invariant under T, as desired.

Suppose $u \in \operatorname{range} p(T)$. Then there exists $v \in V$ such that $u = p(T)v$. Thus

$$Tu = T(p(T)v) = p(T)(Tv).$$

Hence $Tu \in \operatorname{range} p(T)$. Thus $\operatorname{range} p(T)$ is invariant under T, as desired. ∎

Exercises 5A

1 Suppose $T \in \mathcal{L}(V)$ and U is a subspace of V.

(a) Prove that if $U \subseteq \operatorname{null} T$, then U is invariant under T.

(b) Prove that if $\operatorname{range} T \subseteq U$, then U is invariant under T.

2 Suppose that $T \in \mathcal{L}(V)$ and $V_1, \dots, V_m$ are subspaces of V invariant under T. Prove that $V_1 + \cdots + V_m$ is invariant under T.

3 Suppose $T \in \mathcal{L}(V)$. Prove that the intersection of every collection of subspaces of V invariant under T is invariant under T.

4 Prove or give a counterexample: If V is finite-dimensional and U is a subspace of V that is invariant under every operator on V, then $U = \{0\}$ or $U = V$.

5 Suppose $T \in \mathcal{L}(\mathbf{R}^2)$ is defined by $T(x, y) = (-3y, x)$. Find the eigenvalues of T.

6 Define $T \in \mathcal{L}(\mathbf{F}^2)$ by $T(w, z) = (z, w)$. Find all eigenvalues and eigenvectors of T.

7 Define $T \in \mathcal{L}(\mathbf{F}^3)$ by $T(z_1, z_2, z_3) = (2z_2, 0, 5z_3)$. Find all eigenvalues and eigenvectors of T.

8 Suppose $P \in \mathcal{L}(V)$ is such that $P^2 = P$. Prove that if λ is an eigenvalue of P, then $\lambda = 0$ or $\lambda = 1$.

מוקדם יותר ראינו שאם $T \in \mathcal{L}(V)$, אזי תתי־המרחב $\operatorname{null} T$ ו־ $\operatorname{range} T$ שמורים תחת T (ראו 5.4). כעת נראה שמרחב האפס והטווח של כל פולינום של T גם הם שמורים תחת T.

5.18 מרחב האפס והטווח של $p(T)$ שמורים תחת T

נניח ש־ $T \in \mathcal{L}(V)$ וגם $p \in \mathcal{P}(\mathbf{F})$. אזי $\operatorname{null} p(T)$ ו־ $\operatorname{range} p(T)$ שמורים תחת T.

הוכחה נניח ש־ $u \in \operatorname{null} p(T)$. אזי $p(T)u = 0$. לפיכך

$$(p(T))(Tu) = (p(T)\,T)(u) = (T\,p(T))(u) = T(p(T)u) = T(0) = 0.$$

לכן $Tu \in \operatorname{null} p(T)$. לפיכך $\operatorname{null} p(T)$ שמור תחת T, כפי שרצינו.

נניח ש־ $u \in \operatorname{range} p(T)$. אזי קיים $v \in V$ כך ש־ $u = p(T)v$. לפיכך

$$Tu = T(p(T)v) = p(T)(Tv).$$

לכן $Tu \in \operatorname{range} p(T)$. לפיכך $\operatorname{range} p(T)$ שמור תחת T, כפי שרצינו. ■

תרגילים הו

1 נניח ש־ $T \in \mathcal{L}(V)$ ו־ U תת־מרחב של V.

א. הוכיחו שאם $U \subseteq \operatorname{null} T$, אזי U שמור תחת T.

ב. הוכיחו שאם $\operatorname{range} T \subseteq U$, אזי U שמור תחת T.

2 נניח ש־ $T \in \mathcal{L}(V)$ וגם $V_1, \dots, V_m$ תתי־מרחב של V השמורים תחת T. הוכיחו ש־ $V_1 + \cdots + V_m$ שמור תחת T.

3 נניח ש־ $T \in \mathcal{L}(V)$. הוכיחו שהחיתוך של כל אוסף תתי־מרחב של V השמורים תחת T שמור תחת T.

4 הוכיחו או הביאו דוגמה נגדית: אם V מממד סופי ו־ U תת־מרחב של V השמור תחת כל אופרטור מעל V, אזי $U = \{0\}$ או $U = V$.

5 נניח ש־ $T \in \mathcal{L}(\mathbf{R}^2)$ מוגדר על ידי $T(x, y) = (-3y, x)$. מצאו את הערכים העצמיים של T.

6 הגדירו $T \in \mathcal{L}(\mathbf{F}^2)$ על ידי $T(w, z) = (z, w)$. מצאו את כל הערכים העצמיים והווקטורים העצמיים של T.

7 הגדירו $T \in \mathcal{L}(\mathbf{F}^3)$ על ידי $T(z_1, z_2, z_3) = (2z_2, 0, 5z_3)$. מצאו את כל הערכים העצמיים והווקטורים העצמיים של T.

8 נניח ש־ $P \in \mathcal{L}(V)$ הוא כך ש־ $P^2 = P$. הוכיחו שאם λ ערך עצמי של P, אזי $\lambda = 0$ או $\lambda = 1$.

9 Define $T: \mathcal{P}(\mathbf{R}) \to \mathcal{P}(\mathbf{R})$ by $Tp = p'$. Find all eigenvalues and eigenvectors of T.

10 Define $T \in \mathcal{L}(\mathcal{P}_4(\mathbf{R}))$ by $(Tp)(x) = xp'(x)$ for all $x \in \mathbf{R}$. Find all eigenvalues and eigenvectors of T.

11 Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and $\alpha \in \mathbf{F}$. Prove that there exists $\delta > 0$ such that $T - \lambda I$ is invertible for all $\lambda \in \mathbf{F}$ such that $0 < |\alpha - \lambda| < \delta$.

12 Suppose $V = U \oplus W$, where U and W are nonzero subspaces of V. Define $P \in \mathcal{L}(V)$ by $P(u + w) = u$ for each $u \in U$ and each $w \in W$. Find all eigenvalues and eigenvectors of P.

13 Suppose $T \in \mathcal{L}(V)$. Suppose $S \in \mathcal{L}(V)$ is invertible.

(a) Prove that T and $S^{-1}TS$ have the same eigenvalues.

(b) What is the relationship between the eigenvectors of T and the eigenvectors of $S^{-1}TS$?

14 Give an example of an operator on $\mathbf{R}^4$ that has no (real) eigenvalues.

15 Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and $\lambda \in \mathbf{F}$. Show that λ is an eigenvalue of T if and only if λ is an eigenvalue of the dual operator $T' \in \mathcal{L}(V')$.

16 Suppose $v_1, \dots, v_n$ is a basis of V and $T \in \mathcal{L}(V)$. Prove that if λ is an eigenvalue of T, then

$$|\lambda| \le n \max\{|\mathcal{M}(T)_{j,k}| : 1 \le j, k \le n\},$$

where $\mathcal{M}(T)_{j,k}$ denotes the entry in row j, column k of the matrix of T with respect to the basis $v_1, \dots, v_n$.

See Exercise 19 in Section 6A for a different bound on $|\lambda|$.

17 Suppose $\mathbf{F} = \mathbf{R}$, $T \in \mathcal{L}(V)$, and $\lambda \in \mathbf{R}$. Prove that λ is an eigenvalue of T if and only if λ is an eigenvalue of the complexification $T_{\mathbf{C}}$.

See Exercise 33 in Section 3B for the definition of $T_{\mathbf{C}}$.

18 Suppose $\mathbf{F} = \mathbf{R}$, $T \in \mathcal{L}(V)$, and $\lambda \in \mathbf{C}$. Prove that λ is an eigenvalue of the complexification $T_{\mathbf{C}}$ if and only if $\overline{\lambda}$ is an eigenvalue of $T_{\mathbf{C}}$.

19 Show that the forward shift operator $T \in \mathcal{L}(\mathbf{F}^\infty)$ defined by

$$T(z_1, z_2, \dots) = (0, z_1, z_2, \dots)$$

has no eigenvalues.

20 Define the backward shift operator $S \in \mathcal{L}(\mathbf{F}^\infty)$ by

$$S(z_1, z_2, z_3, \dots) = (z_2, z_3, \dots).$$

(a) Show that every element of $\mathbf{F}$ is an eigenvalue of S.

(b) Find all eigenvectors of S.

9 הגדירו $T: \mathcal{P}(\mathbf{R}) \to \mathcal{P}(\mathbf{R})$ על ידי $Tp = p'$. מצאו את כל הערכים העצמיים והווקטורים העצמיים של T.

10 הגדירו $T \in \mathcal{L}(\mathcal{P}_4(\mathbf{R}))$ על ידי $(Tp)(x) = xp'(x)$ לכל $x \in \mathbf{R}$. מצאו את כל הערכים העצמיים והווקטורים העצמיים של T.

11 נניח ש־V מממד סופי, $T \in \mathcal{L}(V)$, ו־$\alpha \in \mathbf{F}$. הוכיחו שקיים $\delta > 0$ כך ש־$T - \lambda I$ הפיך לכל $\lambda \in \mathbf{F}$ כך ש־$0 < |\alpha - \lambda| < \delta$.

12 נניח ש־$V = U \oplus W$, כאשר U ו־W תתי־מרחב של V השונים מאפס. הגדירו $P \in \mathcal{L}(V)$ על ידי $P(u + w) = u$ לכל $u \in U$ ולכל $w \in W$. מצאו את כל הערכים העצמיים והווקטורים העצמיים של P.

13 נניח ש־$T \in \mathcal{L}(V)$. נניח ש־$S \in \mathcal{L}(V)$ הפיך.

א. הוכיחו שלאופרטור T ולו־$S^{-1}TS$ אותם ערכים עצמיים.

ב. מה הקשר בין הווקטורים העצמיים של T לווקטורים העצמיים של $S^{-1}TS$?

14 הביאו דוגמה לאופרטור מעל $\mathbf{R}^4$ שאין לו ערכים עצמיים (ממשיים).

15 נניח ש־V מממד סופי, $T \in \mathcal{L}(V)$, ו־$\lambda \in \mathbf{F}$. הראו ש־λ ערך עצמי של T אם־ורק־אם λ ערך עצמי של האופרטור הדואלי $T' \in \mathcal{L}(V')$.

16 נניח ש־$v_1, \dots, v_n$ בסיס של V ו־$T \in \mathcal{L}(V)$. הוכיחו שאם λ ערך עצמי של T, אזי

$$|\lambda| \le n \max\{|\mathcal{M}(T)_{j,k}| : 1 \le j, k \le n\},$$

כאשר $\mathcal{M}(T)_{j,k}$ מסמן את האיבר בשורה j, עמודה k של המטריצה של T ביחס לבסיס $v_1, \dots, v_n$.

ראו תרגיל 19 ביחידה 11 לחסם שונה על $|\lambda|$.

17 נניח ש־$T \in \mathcal{L}(V)$, $\mathbf{F} = \mathbf{R}$, ו־$\lambda \in \mathbf{R}$. הוכיחו ש־λ ערך עצמי של T אם־ורק־אם λ ערך עצמי של המירכוב $T_{\mathbf{C}}$.

ראו תרגיל 33 ביחידה 13 להגדרת $T_{\mathbf{C}}$.

18 נניח ש־$T \in \mathcal{L}(V)$, $\mathbf{F} = \mathbf{R}$, ו־$\lambda \in \mathbf{C}$. הוכיחו ש־λ ערך עצמי של המירכוב $T_{\mathbf{C}}$ אם־ורק־אם $\overline{\lambda}$ ערך עצמי של $T_{\mathbf{C}}$.

19 הראו שלאופרטור ההסטה לפנים $T \in \mathcal{L}(\mathbf{F}^\infty)$ המוגדר על ידי

$$T(z_1, z_2, \dots) = (0, z_1, z_2, \dots)$$

אין ערכים עצמיים.

20 הגדירו את אופרטור ההסטה לאחור $S \in \mathcal{L}(\mathbf{F}^\infty)$ על ידי

$$S(z_1, z_2, z_3, \dots) = (z_2, z_3, \dots).$$

א. הראו שכל איבר של $\mathbf{F}$ הוא ערך עצמי של S.

ב. מצאו את כל הווקטורים העצמיים של S.

21 Suppose $T \in \mathcal{L}(V)$ is invertible.

(a) Suppose $\lambda \in \mathbf{F}$ with $\lambda \neq 0$. Prove that λ is an eigenvalue of T if and only if $\frac{1}{\lambda}$ is an eigenvalue of T^{-1}.

(b) Prove that T and T^{-1} have the same eigenvectors.

22 Suppose $T \in \mathcal{L}(V)$ and there exist nonzero vectors u and w in V such that

$$Tu = 3w \quad \text{and} \quad Tw = 3u.$$

Prove that 3 or -3 is an eigenvalue of T.

23 Suppose V is finite-dimensional and $S, T \in \mathcal{L}(V)$. Prove that ST and TS have the same eigenvalues.

24 Suppose A is an n-by-n matrix with entries in $\mathbf{F}$. Define $T \in \mathcal{L}(\mathbf{F}^n)$ by $Tx = Ax$, where elements of $\mathbf{F}^n$ are thought of as n-by-1 column vectors.

(a) Suppose the sum of the entries in each row of A equals 1. Prove that 1 is an eigenvalue of T.

(b) Suppose the sum of the entries in each column of A equals 1. Prove that 1 is an eigenvalue of T.

25 Suppose $T \in \mathcal{L}(V)$ and u, w are eigenvectors of T such that $u + w$ is also an eigenvector of T. Prove that u and w are eigenvectors of T corresponding to the same eigenvalue.

26 Suppose $T \in \mathcal{L}(V)$ is such that every nonzero vector in V is an eigenvector of T. Prove that T is a scalar multiple of the identity operator.

27 Suppose that V is finite-dimensional and $k \in \{1, \dots, \dim V - 1\}$. Suppose $T \in \mathcal{L}(V)$ is such that every subspace of V of dimension k is invariant under T. Prove that T is a scalar multiple of the identity operator.

28 Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Prove that T has at most $1 + \dim \operatorname{range} T$ distinct eigenvalues.

29 Suppose $T \in \mathcal{L}(\mathbf{R}^3)$ and -4, 5, and $\sqrt{7}$ are eigenvalues of T. Prove that there exists $x \in \mathbf{R}^3$ such that $Tx - 9x = (-4, 5, \sqrt{7})$.

30 Suppose $T \in \mathcal{L}(V)$ and $(T - 2I)(T - 3I)(T - 4I) = 0$. Suppose λ is an eigenvalue of T. Prove that $\lambda = 2$ or $\lambda = 3$ or $\lambda = 4$.

31 Give an example of $T \in \mathcal{L}(\mathbf{R}^2)$ such that $T^4 = -I$.

32 Suppose $T \in \mathcal{L}(V)$ has no eigenvalues and $T^4 = I$. Prove that $T^2 = -I$.

33 Suppose $T \in \mathcal{L}(V)$ and m is a positive integer.

(a) Prove that T is injective if and only if T^m is injective.

(b) Prove that T is surjective if and only if T^m is surjective.

21 נניח ש־ $T \in \mathcal{L}(V)$ הפיך.

א. נניח ש־ $\lambda \in \mathbf{F}$ ו־ $\lambda \neq 0$. הוכיחו ש־ λ ערך עצמי של T אם־ורק־אם $\frac{1}{\lambda}$ ערך עצמי של T^{-1}.

ב. הוכיחו שלאופרטורים T ו־ T^{-1} יש אותם וקטורים עצמיים.

22 נניח ש־ $T \in \mathcal{L}(V)$ ושקיימים וקטורים השונים מאפס u ו־ w ב־ V כך ש־

$$Tu = 3w \quad \text{וגם} \quad Tw = 3u.$$

הוכיחו ש־ 3 או -3 ערך עצמי של T.

23 נניח ש־ V מממד סופי ו־ $S, T \in \mathcal{L}(V)$. הוכיחו שלאופרטורים ST ו־ TS יש אותם ערכים עצמיים.

24 נניח ש־ A מטריצה $n \times n$ עם איברים ב־ $\mathbf{F}$. הגדירו $T \in \mathcal{L}(\mathbf{F}^n)$ על ידי $Tx = Ax$, כאשר אנו חושבים על איברים ב־ $\mathbf{F}^n$ כווקטורי עמודה $n \times 1$.

א. נניח שסכום האיברים בכל שורה של A שווה 1. הוכיחו ש־ 1 ערך עצמי של T.

ב. נניח שסכום האיברים בכל עמודה של A שווה 1. הוכיחו ש־ 1 ערך עצמי של T.

25 נניח ש־ $T \in \mathcal{L}(V)$ ו־ u, w וקטורים עצמיים של T כך ש־ $u + w$ גם הוא וקטור עצמי של T. הוכיחו ש־ u ו־ w וקטורים עצמיים של T השייכים לאותו ערך עצמי.

26 נניח ש־ $T \in \mathcal{L}(V)$ כך שכל וקטור השונה מאפס ב־ V הוא וקטור עצמי של T. הוכיחו ש־ T כפולת סקלר של אופרטור הזהות.

27 נניח ש־ V מממד סופי ו־ $k \in \{1, \dots, \dim V - 1\}$. נניח ש־ $T \in \mathcal{L}(V)$ כך שכל תת־מרחב של V מממד k שמור תחת T. הוכיחו ש־ T כפולת סקלר של אופרטור הזהות.

28 נניח ש־ V מממד סופי ו־ $T \in \mathcal{L}(V)$. הוכיחו כי מספר הערכים העצמיים השונים של T הוא לכל היותר $1 + \dim \operatorname{range} T$.

29 נניח ש־ $T \in \mathcal{L}(\mathbf{R}^3)$ ו־ $-4, 5$, ו־ $\sqrt{7}$ ערכים עצמיים של T. הוכיחו כי קיים $x \in \mathbf{R}^3$ כך ש־ $Tx - 9x = (-4, 5, \sqrt{7})$.

30 נניח ש־ $T \in \mathcal{L}(V)$ ו־ $(T - 2I)(T - 3I)(T - 4I) = 0$. נניח ש־ λ ערך עצמי של T. הוכיחו ש־ $\lambda = 2$ או $\lambda = 3$ או $\lambda = 4$.

31 הביאו דוגמה ל־ $T \in \mathcal{L}(\mathbf{R}^2)$ כך ש־ $T^4 = -I$.

32 נניח שלאופרטור $T \in \mathcal{L}(V)$ אין ערכים עצמיים וגם $T^4 = I$. הוכיחו ש־ $T^2 = -I$.

33 נניח ש־ $T \in \mathcal{L}(V)$ ו־ m מספר שלם חיובי.

א. הוכיחו ש־ T חד־חד־ערכי אם־ורק־אם T^m חד־חד־ערכי.

ב. הוכיחו ש־ T על אם־ורק־אם T^m על.

34 Suppose V is finite-dimensional and $v_1, \dots, v_m \in V$. Prove that the list $v_1, \dots, v_m$ is linearly independent if and only if there exists $T \in \mathcal{L}(V)$ such that $v_1, \dots, v_m$ are eigenvectors of T corresponding to distinct eigenvalues.

35 Suppose that $\lambda_1, \dots, \lambda_n$ is a list of distinct real numbers. Prove that the list $e^{\lambda_1 x}, \dots, e^{\lambda_n x}$ is linearly independent in the vector space of real-valued functions on $\mathbf{R}$.

Hint: Let $V = \operatorname{span}(e^{\lambda_1 x}, \dots, e^{\lambda_n x})$*, and define an operator* $D \in \mathcal{L}(V)$ *by* $Df = f'$*. Find eigenvalues and eigenvectors of* D*.*

36 Suppose that $\lambda_1, \dots, \lambda_n$ is a list of distinct positive numbers. Prove that the list $\cos(\lambda_1 x), \dots, \cos(\lambda_n x)$ is linearly independent in the vector space of real-valued functions on $\mathbf{R}$.

37 Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Define $\mathcal{A} \in \mathcal{L}(\mathcal{L}(V))$ by

$$\mathcal{A}(S) = TS$$

for each $S \in \mathcal{L}(V)$. Prove that the set of eigenvalues of T equals the set of eigenvalues of $\mathcal{A}$.

38 Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and U is a subspace of V invariant under T. The *quotient operator* $T/U \in \mathcal{L}(V/U)$ is defined by

$$(T/U)(v + U) = Tv + U$$

for each $v \in V$.

(a) Show that the definition of T/U makes sense (which requires using the condition that U is invariant under T) and show that T/U is an operator on V/U.

(b) Show that each eigenvalue of T/U is an eigenvalue of T.

39 Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Prove that T has an eigenvalue if and only if there exists a subspace of V of dimension $\dim V - 1$ that is invariant under T.

40 Suppose $S, T \in \mathcal{L}(V)$ and S is invertible. Suppose $p \in \mathcal{P}(\mathbf{F})$ is a polynomial. Prove that

$$p(STS^{-1}) = Sp(T)S^{-1}.$$

41 Suppose $T \in \mathcal{L}(V)$ and U is a subspace of V invariant under T. Prove that U is invariant under $p(T)$ for every polynomial $p \in \mathcal{P}(\mathbf{F})$.

42 Define $T \in \mathcal{L}(\mathbf{F}^n)$ by $T(x_1, x_2, x_3, \dots, x_n) = (x_1, 2x_2, 3x_3, \dots, nx_n)$.

(a) Find all eigenvalues and eigenvectors of T.

(b) Find all subspaces of $\mathbf{F}^n$ that are invariant under T.

43 Suppose that V is finite-dimensional, $\dim V > 1$, and $T \in \mathcal{L}(V)$. Prove that $\{p(T) : p \in \mathcal{P}(\mathbf{F})\} \neq \mathcal{L}(V)$.

34 נניח ש־ V מממד סופי ו־ $v_1, \dots, v_m \in V$. הוכיחו ש־ $v_1, \dots, v_m$ רשימה בלתי תלויה ליניארית אם־ורק־אם קיים $T \in \mathcal{L}(V)$ כך ש־ $v_1, \dots, v_m$ וקטורים עצמיים של T השייכים לערכים עצמיים שונים.

35 נניח ש־ $\lambda_1, \dots, \lambda_n$ רשימת מספרים ממשיים שונים. הוכיחו ש־ $e^{\lambda_1 x}, \dots, e^{\lambda_n x}$ רשימה בלתי תלויה ליניארית במרחב הווקטורי של הפונקציות הממשיות מעל $\mathbf{R}$.

רמז: יהי $V = \operatorname{span}(e^{\lambda_1 x}, \dots, e^{\lambda_n x})$, והגדירו אופרטור $D \in \mathcal{L}(V)$ על ידי $Df = f'$. מצאו את הערכים העצמיים והווקטורים העצמיים של D.

36 נניח ש־ $\lambda_1, \dots, \lambda_n$ היא רשימה של מספרים חיוביים שונים. הוכיחו כי הרשימה $\cos(\lambda_1 x), \dots, \cos(\lambda_n x)$ בלתי תלויה ליניארית במרחב הווקטורי של הפונקציות הממשיות מעל $\mathbf{R}$.

37 נניח ש־ V מממד סופי ו־ $T \in \mathcal{L}(V)$. הגדירו $\mathcal{A} \in \mathcal{L}(\mathcal{L}(V))$ על ידי

$$\mathcal{A}(S) = TS$$

לכל $S \in \mathcal{L}(V)$. הוכיחו כי קבוצת הערכים העצמיים של T שווה לקבוצת הערכים העצמיים של $\mathcal{A}$.

38 נניח ש־ V מממד סופי, $T \in \mathcal{L}(V)$, ו־ U תת־מרחב של V שמור תחת T. אופרטור המנה $T/U \in \mathcal{L}(V/U)$ מוגדר על ידי

$$(T/U)(v + U) = Tv + U$$

לכל $v \in V$.

א. הראו שהגדרת T/U הגיונית (דבר הדורש שימוש בתנאי ש־ U שמור תחת T) והראו ש־ T/U אופרטור מעל V/U.

ב. הראו שכל ערך עצמי של T/U הוא ערך עצמי של T.

39 נניח ש־ V מממד סופי ו־ $T \in \mathcal{L}(V)$. הוכיחו שלאופרטור T יש ערך עצמי אם־ורק־אם קיים תת־מרחב של V שממדו $\dim V - 1$ ושמור תחת T.

40 נניח ש־ $S, T \in \mathcal{L}(V)$ ו־ S הפיך. נניח ש־ $p \in \mathcal{P}(\mathbf{F})$ פולינום. הוכיחו ש־

$$p(STS^{-1}) = Sp(T)S^{-1}.$$

41 נניח ש־ $T \in \mathcal{L}(V)$ ו־ U תת־מרחב של V השמור תחת T. הוכיחו ש־ U שמור תחת $p(T)$ לכל פולינום $p \in \mathcal{P}(\mathbf{F})$.

42 הגדירו $T \in \mathcal{L}(\mathbf{F}^n)$ על ידי $T(x_1, x_2, x_3, \dots, x_n) = (x_1, 2x_2, 3x_3, \dots, nx_n)$.

א. מצאו את כל הערכים העצמיים והווקטורים העצמיים של T.

ב. מצאו את כל תתי־המרחב של $\mathbf{F}^n$ השמורים תחת T.

43 נניח ש־ V מממד סופי, $\dim V > 1$, ו־ $T \in \mathcal{L}(V)$. הוכיחו ש־

$$\{p(T) : p \in \mathcal{P}(\mathbf{F})\} \neq \mathcal{L}(V).$$

5B The Minimal Polynomial

Existence of Eigenvalues on Complex Vector Spaces

Now we come to one of the central results about operators on finite-dimensional complex vector spaces.

5.19 *existence of eigenvalues*

Every operator on a finite-dimensional nonzero complex vector space has an eigenvalue.

Proof Suppose V is a finite-dimensional complex vector space of dimension $n > 0$ and $T \in \mathcal{L}(V)$. Choose $v \in V$ with $v \neq 0$. Then

$$v, Tv, T^2v, \dots, T^nv$$

is not linearly independent, because V has dimension n and this list has length $n + 1$. Hence some linear combination (with not all the coefficients equal to 0) of the vectors above equals 0. Thus there exists a nonconstant polynomial p of smallest degree such that

$$p(T)v = 0.$$

By the first version of the fundamental theorem of algebra (see 4.12), there exists $\lambda \in \mathbf{C}$ such that $p(\lambda) = 0$. Hence there exists a polynomial $q \in \mathcal{P}(\mathbf{C})$ such that

$$p(z) = (z - \lambda)q(z)$$

for every $z \in \mathbf{C}$ (see 4.6). This implies (using 5.17) that

$$0 = p(T)v = (T - \lambda I)(q(T)v).$$

Because q has smaller degree than p, we know that $q(T)v \neq 0$. Thus the equation above implies that λ is an eigenvalue of T with eigenvector $q(T)v$. ∎

The proof above makes crucial use of the fundamental theorem of algebra. The comment following Exercise 16 helps explain why the fundamental theorem of algebra is so tightly connected to the result above.

The hypothesis in the result above that $\mathbf{F} = \mathbf{C}$ cannot be replaced with the hypothesis that $\mathbf{F} = \mathbf{R}$, as shown by Example 5.9. The next example shows that the finite-dimensional hypothesis in the result above also cannot be deleted.

5.20 example: *an operator on a complex vector space with no eigenvalues*

Define $T \in \mathcal{L}(\mathcal{P}(\mathbf{C}))$ by $(Tp)(z) = zp(z)$. If $p \in \mathcal{P}(\mathbf{C})$ is a nonzero polynomial, then the degree of Tp is one more than the degree of p, and thus Tp cannot equal a scalar multiple of p. Hence T has no eigenvalues.

Because $\mathcal{P}(\mathbf{C})$ is infinite-dimensional, this example does not contradict the result above.

2ה הפולינום המינימלי

קיומם של ערכים עצמיים מעל מרחב וקטורי מרוכב

כעת אנו ניגשים אל אחת המסקנות המרכזיות לגבי אופרטורים מעל מרחב וקטורי מרוכב מממד סופי.

5.19 קיומם של ערכים עצמיים

לכל אופרטור מעל מרחב וקטורי מרוכב מממד סופי השונה מאפס יש ערך עצמי.

הוכחה נניח ש־ V מרחב וקטורי מרוכב מממד סופי שממדו $n > 0$ ו־ $T \in \mathcal{L}(V)$. נבחר $v \in V$ עם $v \neq 0$. נקבל ש־

$$v, Tv, T^2v, \dots, T^nv$$

אינה בלתי תלויה ליניארית, שכן ממדו של V הוא n וברשימה זו $n+1$ איברים. לכן, צירוף ליניארי כלשהו (כאשר לא כל המקדמים שווים 0) של הווקטורים לעיל שווה 0. לפיכך קיים פולינום שאינו קבוע p מהמעלה הנמוכה ביותר כך ש־

$$p(T)v = 0.$$

לפי המשפט היסודי של האלגברה בגרסתו הראשונה (ראו 4.12), קיים $\lambda \in \mathbf{C}$ כך ש־ $p(\lambda) = 0$. לכן קיים פולינום $q \in \mathcal{P}(\mathbf{C})$ כך ש־

$$p(z) = (z - \lambda)q(z)$$

לכל $z \in \mathbf{C}$ (ראו 4.6). מכאן נובע (תוך שימוש ב־ 5.17) ש־

$$0 = p(T)v = (T - \lambda I)(q(T)v).$$

מאחר שמעלת q נמוכה מזו של p, אנו יודעים ש־ $q(T)v \neq 0$. לפיכך מהמשוואה לעיל נובע ש־ λ ערך עצמי של T עם וקטור עצמי $q(T)v$. ■

ההוכחה לעיל עושה שימוש מרכזי במשפט היסודי של האלגברה. ההערה שאחרי תרגיל 16 מסייעת להבין מדוע למשפט היסודי של האלגברה קשר כה הדוק למסקנה לעיל.

את ההנחה במסקנה לעיל כי $\mathbf{F} = \mathbf{C}$ לא ניתן להחליף בהנחה ש־ $\mathbf{F} = \mathbf{R}$, כפי שראינו בדוגמה 5.9. בדוגמה להלן נראה שגם ההנחה כי מדובר במרחב מממד סופי הכרחית.

5.20 דוגמה: אופרטור מעל מרחב וקטורי מרוכב ללא ערכים עצמיים

נגדיר $T \in \mathcal{L}(\mathcal{P}(\mathbf{C}))$ על ידי $(Tp)(z) = zp(z)$. אם $p \in \mathcal{P}(\mathbf{C})$ פולינום השונה מאפס, אזי מעלת Tp גדולה באחת ממעלת p, ולפיכך Tp אינו יכול להיות שווה לכפולת סקלר של p. לכן ל־ T אין ערכים עצמיים.

מאחר ש־ $\mathcal{P}(\mathbf{C})$ מממד אינסופי, דוגמה זו אינה בסתירה למסקנה לעיל.

Eigenvalues and the Minimal Polynomial

In this subsection we introduce an important polynomial associated with each operator. We begin with the following definition.

5.21 definition: *monic polynomial*

A *monic polynomial* is a polynomial whose highest-degree coefficient equals 1.

For example, the polynomial $2 + 9z^2 + z^7$ is a monic polynomial of degree 7.

5.22 *existence, uniqueness, and degree of minimal polynomial*

Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Then there is a unique monic polynomial $p \in \mathcal{P}(\mathbf{F})$ of smallest degree such that $p(T) = 0$. Furthermore, $\deg p \le \dim V$.

Proof If $\dim V = 0$, then I is the zero operator on V and thus we take p to be the constant polynomial 1.

Now use induction on $\dim V$. Thus assume that $\dim V > 0$ and that the desired result is true for all operators on all vector spaces of smaller dimension. Let $u \in V$ be such that $u \ne 0$. The list $u, Tu, \dots, T^{\dim V}u$ has length $1 + \dim V$ and thus is linearly dependent. By the linear dependence lemma (2.19), there is a smallest positive integer $m \le \dim V$ such that $T^m u$ is a linear combination of $u, Tu, \dots, T^{m-1}u$. Thus there exist scalars $c_0, c_1, c_2, \dots, c_{m-1} \in \mathbf{F}$ such that

$$c_0 u + c_1 Tu + \cdots + c_{m-1}T^{m-1}u + T^m u = 0. \tag{5.23}$$

Define a monic polynomial $q \in \mathcal{P}_m(\mathbf{F})$ by

$$q(z) = c_0 + c_1 z + \cdots + c_{m-1}z^{m-1} + z^m.$$

Then 5.23 implies that $q(T)u = 0$.

If k is a nonnegative integer, then

$$q(T)(T^k u) = T^k(q(T)u) = T^k(0) = 0.$$

The linear dependence lemma (2.19) shows that $u, Tu, \dots, T^{m-1}u$ is linearly independent. Thus the equation above implies that $\dim \operatorname{null} q(T) \ge m$. Hence

$$\dim \operatorname{range} q(T) = \dim V - \dim \operatorname{null} q(T) \le \dim V - m.$$

Because $\operatorname{range} q(T)$ is invariant under T (by 5.18), we can apply our induction hypothesis to the operator $T|_{\operatorname{range} q(T)}$ on the vector space $\operatorname{range} q(T)$. Thus there is a monic polynomial $s \in \mathcal{P}(\mathbf{F})$ with

$$\deg s \le \dim V - m \quad \text{and} \quad s(T|_{\operatorname{range} q(T)}) = 0.$$

Hence for all $v \in V$ we have

$$((sq)(T))(v) = s(T)(q(T)v) = 0$$

because $q(T)v \in \operatorname{range} q(T)$ and $s(T)|_{\operatorname{range} q(T)} = s(T|_{\operatorname{range} q(T)}) = 0$. Thus sq is a monic polynomial such that $\deg sq \le \dim V$ and $(sq)(T) = 0$.

ערכים עצמיים והפולינום המינימלי

בתת־יחידה זו אנו מציגים פולינום חשוב המשויך לכל אופרטור. נפתח בהגדרה הבאה.

5.21 הגדרה: פולינום מתוקן

פולינום מתוקן הוא פולינום שערך המקדם המוביל בו הוא 1.

לדוגמה, הפולינום $2 + 9z^2 + z^7$ הוא פולינום מתוקן שמעלתו 7.

5.22 קיום, יחידות, ומעלת פולינום מינימלי

נניח ש־ V מממד סופי ו־ $T \in \mathcal{L}(V)$. אזי קיים פולינום מתוקן יחיד $p \in \mathcal{P}(\mathbf{F})$ מהמעלה הנמוכה ביותר כך ש־ $p(T) = 0$. בנוסף, $\deg p \le \dim V$.

הוכחה אם $\dim V = 0$, אזי I הוא אופרטור האפס מעל V ולפיכך נבחר את p כפולינום הקבוע 1.

כעת נשתמש באינדוקציה על $\dim V$. לפיכך נניח ש־ $\dim V > 0$ ונניח שהתוצאה מתקיימת עבור כל האופרטורים מעל כל המרחבים הווקטוריים מממד קטן יותר. יהי $u \in V$ כך ש־ $u \neq 0$. אורך הרשימה $u, Tu, \dots, T^{\dim V}u$ הוא $1 + \dim V$, ולפיכך היא תלויה ליניארית. לפי למת התלות הליניארית (2.19), קיים מספר שלם חיובי $m \le \dim V$ הקטן ביותר כך ש־ $T^m u$ צירוף ליניארי של $u, Tu, \dots, T^{m-1}u$. לפיכך קיימים סקלרים $c_0, c_1, c_2, \dots, c_{m-1} \in \mathbf{F}$ כך ש־

$$c_0 u + c_1 Tu + \cdots + c_{m-1}T^{m-1}u + T^m u = 0. \tag{5.23}$$

נגדיר פולינום מתוקן $q \in \mathcal{P}_m(\mathbf{F})$ על ידי

$$q(z) = c_0 + c_1 z + \cdots + c_{m-1}z^{m-1} + z^m.$$

אזי 5.23 גוררת ש־ $q(T)u = 0$.

אם k מספר שלם אי־שלילי, אזי

$$q(T)(T^k u) = T^k(q(T)u) = T^k(0) = 0.$$

למת התלות הליניארית (2.19) מראה לנו ש־ $u, Tu, \dots, T^{m-1}u$ בלתי תלויה ליניארית. לפיכך המשוואה לעיל גוררת ש־ $\dim \operatorname{null} q(T) \ge m$. לכן

$$\dim \operatorname{range} q(T) = \dim V - \dim \operatorname{null} q(T) \le \dim V - m.$$

מאחר ש־ $\operatorname{range} q(T)$ שמור תחת T (לפי 5.18), נוכל להחיל את הנחת האינדוקציה שלנו על האופרטור $T|_{\operatorname{range} q(T)}$ מעל המרחב הווקטורי $\operatorname{range} q(T)$. לפיכך קיים פולינום מתוקן $s \in \mathcal{P}(\mathbf{F})$ עם

$$s(T|_{\operatorname{range} q(T)}) = 0 \quad \text{וגם} \quad \deg s \le \dim V - m.$$

לכן, לכל $v \in V$ נקבל

$$((sq)(T))(v) = s(T)(q(T)v) = 0$$

מאחר ש־ $q(T)v \in \operatorname{range} q(T)$ ו־ $s(T)|_{\operatorname{range} q(T)} = s(T|_{\operatorname{range} q(T)}) = 0$. לפיכך sq פולינום מתוקן כך ש־ $\deg sq \le \dim V$ ובנוסף $(sq)(T) = 0$.

The paragraph above shows that there is a monic polynomial of degree at most $\dim V$ that when applied to T gives the 0 operator. Thus there is a monic polynomial of smallest degree with this property, completing the existence part of this result.

Let $p \in \mathcal{P}(\mathbf{F})$ be a monic polynomial of smallest degree such that $p(T) = 0$. To prove the uniqueness part of the result, suppose $r \in \mathcal{P}(\mathbf{F})$ is a monic polynomial of the same degree as p and $r(T) = 0$. Then $(p - r)(T) = 0$ and also $\deg(p - r) < \deg p$. If $p - r$ were not equal to 0, then we could divide $p - r$ by the coefficient of the highest-order term in $p - r$ to get a monic polynomial (of smaller degree than p) that when applied to T gives the 0 operator. Thus $p - r = 0$, as desired. ∎

The previous result justifies the following definition.

5.24 definition: *minimal polynomial*

Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Then the *minimal polynomial* of T is the unique monic polynomial $p \in \mathcal{P}(\mathbf{F})$ of smallest degree such that $p(T) = 0$.

To compute the minimal polynomial of an operator $T \in \mathcal{L}(V)$, we need to find the smallest positive integer m such that the equation

$$c_0 I + c_1 T + \cdots + c_{m-1} T^{m-1} = -T^m$$

has a solution $c_0, c_1, \ldots, c_{m-1} \in \mathbf{F}$. If we pick a basis of V and replace T in the equation above with the matrix of T, then the equation above can be thought of as a system of $(\dim V)^2$ linear equations in the m unknowns $c_0, c_1, \ldots, c_{m-1} \in \mathbf{F}$. Gaussian elimination or another fast method of solving systems of linear equations can tell us whether a solution exists, testing successive values $m = 1, 2, \ldots$ until a solution exists. By 5.22, a solution exists for some smallest positive integer $m \le \dim V$. The minimal polynomial of T is then $c_0 + c_1 z + \cdots + c_{m-1} z^{m-1} + z^m$.

Even faster (usually), pick $v \in V$ with $v \ne 0$ and consider the equation

$$c_0 v + c_1 Tv + \cdots + c_{\dim V - 1} T^{\dim V - 1} v = -T^{\dim V} v. \tag{5.25}$$

Use a basis of V to convert the equation above to a system of $\dim V$ linear equations in $\dim V$ unknowns $c_0, c_1, \ldots, c_{\dim V - 1}$. If this system of equations has a unique solution $c_0, c_1, \ldots, c_{\dim V - 1}$ (as happens most of the time), then the scalars $c_0, c_1, \ldots, c_{\dim V - 1}, 1$ are the coefficients of the minimal polynomial of T (because 5.22 states that the degree of the minimal polynomial is at most $\dim V$).

Consider operators on $\mathbf{R}^4$ (thought of as 4-by-4 matrices with respect to the standard basis), and take $v = (1, 0, 0, 0)$ in the paragraph above. The faster method described above works on over 99.8% of the 4-by-4 matrices with integer entries in the interval $[-10, 10]$ and on over 99.999% of the 4-by-4 matrices with integer entries in $[-100, 100]$.

These estimates are based on testing millions of random matrices.

בפסקה לעיל ראינו שקיים פולינום מתוקן שמעלתו לכל היותר $\dim V$ והפעלתו על T מפיקה את אופרטור האפס. לפיכך קיים פולינום מתוקן מהמעלה הנמוכה ביותר עם תכונה זו, ובכך השלמנו את הוכחת חלק הקיום של מסקנה זו.

יהי $p \in \mathcal{P}(\mathbf{F})$ פולינום מתוקן מהמעלה הנמוכה ביותר כך ש־ $p(T) = 0$. להוכחת חלק היְחִידוּת של המסקנה, נניח ש־ $r \in \mathcal{P}(\mathbf{F})$ פולינום מתוקן שמעלתו זהה לזו של p ובנוסף, $r(T) = 0$. אזי $(p - r)(T) = 0$ ובנוסף $\deg(p - r) < \deg p$. אם $p - r$ לא היה שווה 0, היינו יכולים לחלק את $p - r$ במקדם המוביל של $p - r$ כדי לקבל פולינום מתוקן (שמעלתו נמוכה מזו של p) שכאשר מופעל על T מפיק את אופרטור האפס. לפיכך $p - r = 0$, כפי שרצינו. ∎

המסקנה הקודמת מצדיקה את ההגדרה להלן.

5.24 הגדרה: פולינום מינימלי

נניח ש־ V מממד סופי ו־ $T \in \mathcal{L}(V)$. אזי הפולינום המינימלי של T הוא הפולינום המתוקן היחיד $p \in \mathcal{P}(\mathbf{F})$ מהמעלה הנמוכה ביותר כך ש־ $p(T) = 0$.

לחישוב הפולינום המינימלי של אופרטור $T \in \mathcal{L}(V)$, עלינו למצוא את המספר השלם החיובי הקטן ביותר m כך שלמשוואה

$$c_0 I + c_1 T + \cdots + c_{m-1} T^{m-1} = -T^m$$

יש פתרון $c_0, c_1, \ldots, c_{m-1} \in \mathbf{F}$. אם נבחר בסיס של V ונחליף את T במשוואה לעיל במטריצה של T, אזי נוכל לחשוב על המשוואה לעיל כמערכת של $(\dim V)^2$ משוואות ליניאריות עם m נעלמים $c_0, c_1, \ldots, c_{m-1} \in \mathbf{F}$. שיטת החילוץ של גאוס או כל שיטה מהירה אחרת לפתרון מערכת משוואות ליניאריות תוכל לתת מענה לשאלה האם קיים פתרון, תוך בדיקת ערכים עוקבים $m = 1, 2, \ldots$ עד למציאת פתרון. לפי 5.22, קיים מספר שלם חיובי קטן ביותר $m \le \dim V$ שעבורו קיים פתרון. הפולינום המינימלי של T הוא אם כן $c_0 + c_1 z + \cdots + c_{m-1} z^{m-1} + z^m$.

שיטה מהירה אף יותר (ברוב המקרים): נבחר $v \in V$ כך ש־ $v \neq 0$ ונתבונן במשוואה

$$c_0 v + c_1 T v + \cdots + c_{\dim V - 1} T^{\dim V - 1} v = -T^{\dim V} v. \tag{5.25}$$

נשתמש בבסיס של V להמרת המשוואה למערכת של $\dim V$ משוואות ליניאריות עם $\dim V$ נעלמים $c_0, c_1, \ldots, c_{\dim V - 1}$. אם למערכת משוואות זו יש פתרון יחיד $c_0, c_1, \ldots, c_{\dim V - 1}$ (כפי שקורה ברוב המקרים), אזי הסקלרים $c_0, c_1, \ldots, c_{\dim V - 1}, 1$ הם המקדמים של הפולינום המינימלי של T (מאחר ש־ 5.22 קובעת שמעלת הפולינום המינימלי היא לכל היותר $\dim V$).

הערכות אלו מבוססות על בדיקת מיליוני מטריצות אקראיות.

נתבונן באופרטורים מעל $\mathbf{R}^4$ (כשאנו חושבים עליהם כמטריצות 4×4 ביחס לבסיס הסטנדרטי), ונקח את $v = (1, 0, 0, 0)$ בפסקה לעיל. השיטה המהירה שתוארה לעיל מצליחה עבור יותר מ־ 99.8% מהמטריצות 4×4 עם איברים שלמים בקטע $[-10, 10]$ וביותר מ־ 99.999% מהמטריצות 4×4 עם איברים שלמים בקטע $[-100, 100]$.

The next example illustrates the faster procedure discussed above.

5.26 example: *minimal polynomial of an operator on* $\mathbf{F}^5$

Suppose $T \in \mathcal{L}(\mathbf{F}^5)$ and

$$\mathcal{M}(T) = \begin{pmatrix} 0 & 0 & 0 & 0 & -3 \\ 1 & 0 & 0 & 0 & 6 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

with respect to the standard basis e_1, e_2, e_3, e_4, e_5. Taking $v = e_1$ for 5.25, we have

$$\begin{aligned} Te_1 &= e_2, & T^4e_1 &= T(T^3e_1) = Te_4 = e_5, \\ T^2e_1 &= T(Te_1) = Te_2 = e_3, & T^5e_1 &= T(T^4e_1) = Te_5 = -3e_1 + 6e_2. \\ T^3e_1 &= T(T^2e_1) = Te_3 = e_4, \end{aligned}$$

Thus $3e_1 - 6Te_1 = -T^5e_1$. The list $e_1, Te_1, T^2e_1, T^3e_1, T^4e_1$, which equals the list e_1, e_2, e_3, e_4, e_5, is linearly independent, so no other linear combination of this list equals $-T^5e_1$. Hence the minimal polynomial of T is $3 - 6z + z^5$.

Recall that by definition, eigenvalues of operators on V and zeros of polynomials in $\mathcal{P}(\mathbf{F})$ must be elements of $\mathbf{F}$. In particular, if $\mathbf{F} = \mathbf{R}$, then eigenvalues and zeros must be real numbers.

5.27 *eigenvalues are the zeros of the minimal polynomial*

Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$.

(a) The zeros of the minimal polynomial of T are the eigenvalues of T.

(b) If V is a complex vector space, then the minimal polynomial of T has the form

$$(z - \lambda_1) \cdots (z - \lambda_m),$$

where $\lambda_1, \ldots, \lambda_m$ is a list of all eigenvalues of T, possibly with repetitions.

Proof Let p be the minimal polynomial of T.

(a) First suppose $\lambda \in \mathbf{F}$ is a zero of p. Then p can be written in the form

$$p(z) = (z - \lambda)q(z),$$

where q is a monic polynomial with coefficients in $\mathbf{F}$ (see 4.6). Because $p(T) = 0$, we have

$$0 = (T - \lambda I)(q(T)v)$$

for all $v \in V$. Because $\deg q = (\deg p) - 1$ and p is the minimal polynomial of T, there exists at least one vector $v \in V$ such that $q(T)v \neq 0$. The equation above thus implies that λ is an eigenvalue of T, as desired.

הדוגמה הבאה ממחישה את השיטה המהירה שתוארה לעיל.

5.26 דוגמה: פולינום מינימלי של אופרטור מעל $\mathbf{F}^5$

נניח ש־ $T \in \mathcal{L}(\mathbf{F}^5)$ ו־

$$\mathcal{M}(T) = \begin{pmatrix} 0 & 0 & 0 & 0 & -3 \\ 1 & 0 & 0 & 0 & 6 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

ביחס לבסיס הסטנדרטי e_1, e_2, e_3, e_4, e_5. אם נבחר $v = e_1$ עבור 5.25, נקבל

$$Te_1 = e_2, \qquad T^4 e_1 = T(T^3 e_1) = Te_4 = e_5,$$
$$T^2 e_1 = T(Te_1) = Te_2 = e_3, \qquad T^5 e_1 = T(T^4 e_1) = Te_5 = -3e_1 + 6e_2.$$
$$T^3 e_1 = T(T^2 e_1) = Te_3 = e_4,$$

לפיכך $3e_1 - 6Te_1 = -T^5 e_1$. הרשימה $e_1, Te_1, T^2 e_1, T^3 e_1, T^4 e_1$, השווה לרשימה e_1, e_2, e_3, e_4, e_5, בלתי תלויה ליניארית, כך שאין צירוף ליניארי אחר של רשימה זו השווה ל־ $-T^5 e_1$. לכן הפולינום המינימלי של T הוא $3 - 6z + z^5$.

נזכיר כי לפי ההגדרה, ערכים עצמיים של אופרטור מעל V ושורשים של פולינומים ב־ $\mathcal{P}(\mathbf{F})$ חייבים להיות איברים ב־ $\mathbf{F}$. בפרט, אם $\mathbf{F} = \mathbf{R}$, אזי הערכים העצמיים והשורשים חייבים להיות מספרים ממשיים.

5.27 ערכים עצמיים הם שורשי הפולינום המינימלי

נניח ש־ V מממד סופי ו־ $T \in \mathcal{L}(V)$.

א. השורשים של הפולינום המינימלי של T הם הערכים העצמיים של T.

ב. אם V מרחב וקטורי מרוכב, אזי הפולינום המינימלי של T הוא מהצורה
$$(z - \lambda_1) \cdots (z - \lambda_m),$$
כאשר $\lambda_1, \ldots, \lambda_m$ היא רשימת כל הערכים העצמיים של T, עם אפשרות לחזרות.

הוכחה יהי p הפולינום המינימלי של T.

א. ראשית נניח ש־ $\lambda \in \mathbf{F}$ שורש של p. אזי נוכל לכתוב את p באופן הבא
$$p(z) = (z - \lambda)q(z),$$
כאשר q פולינום מתוקן עם מקדמים ב־ $\mathbf{F}$ (ראו 4.6). מאחר ש־ $p(T) = 0$, נקבל
$$0 = (T - \lambda I)(q(T)v)$$
לכל $v \in V$. מאחר ש־ $\deg q = (\deg p) - 1$ ו־ p הפולינום המינימלי של T, קיים לפחות וקטור אחד $v \in V$ כך ש־ $q(T)v \neq 0$. לפיכך מהמשוואה לעיל נובע ש־ λ ערך עצמי של T, כפי שרצינו.

To prove that every eigenvalue of T is a zero of p, now suppose $\lambda \in \mathbf{F}$ is an eigenvalue of T. Thus there exists $v \in V$ with $v \neq 0$ such that $Tv = \lambda v$. Repeated applications of T to both sides of this equation show that $T^k v = \lambda^k v$ for every nonnegative integer k. Thus

$$p(T)v = p(\lambda)v.$$

Because p is the minimal polynomial of T, we have $p(T)v = 0$. Hence the equation above implies that $p(\lambda) = 0$. Thus λ is a zero of p, as desired.

(b) To get the desired result, use (a) and the second version of the fundamental theorem of algebra (see 4.13). ∎

A nonzero polynomial has at most as many distinct zeros as its degree (see 4.8). Thus (a) of the previous result, along with the result that the minimal polynomial of an operator on V has degree at most $\dim V$, gives an alternative proof of 5.12, which states that an operator on V has at most $\dim V$ distinct eigenvalues.

Every monic polynomial is the minimal polynomial of some operator, as shown by Exercise 16, which generalizes Example 5.26. Thus 5.27(a) shows that finding exact expressions for the eigenvalues of an operator is equivalent to the problem of finding exact expressions for the zeros of a polynomial (and thus is not possible for some operators).

5.28 example: *An operator whose eigenvalues cannot be found exactly*

Let $T \in \mathcal{L}(\mathbf{C}^5)$ be the operator defined by

$$T(z_1, z_2, z_3, z_4, z_5) = (-3z_5, z_1 + 6z_5, z_2, z_3, z_4).$$

The matrix of T with respect to the standard basis of $\mathbf{C}^5$ is the 5-by-5 matrix in Example 5.26. As we showed in that example, the minimal polynomial of T is the polynomial

$$3 - 6z + z^5.$$

No zero of the polynomial above can be expressed using rational numbers, roots of rational numbers, and the usual rules of arithmetic (a proof of this would take us considerably beyond linear algebra). Because the zeros of the polynomial above are the eigenvalues of T [by 5.27(a)], we cannot find an exact expression for any eigenvalue of T in any familiar form.

Numerical techniques, which we will not discuss here, show that the zeros of the polynomial above, and thus the eigenvalues of T, are approximately the following five complex numbers:

$$-1.67, \quad 0.51, \quad 1.40, \quad -0.12 + 1.59i, \quad -0.12 - 1.59i.$$

Note that the two nonreal zeros of this polynomial are complex conjugates of each other, as we expect for a polynomial with real coefficients (see 4.14).

כדי להוכיח שכל ערך עצמי של T הוא שורש של p, נניח כעת ש־ $\lambda \in \mathbf{F}$ ערך עצמי של T. לפיכך קיים $v \in V$ עם $v \neq 0$ כך ש־ $Tv = \lambda v$. הפעלות חוזרות של T על שני אגפי המשוואה מראות ש־ $T^k v = \lambda^k v$ לכל k שלם אי־שלילי. לפיכך

$$p(T)v = p(\lambda)v.$$

מאחר ש־ p הוא הפולינום המינימלי של T, קיבלנו ש־ $p(T)v = 0$. לכן מהמשוואה לעיל נובע ש־ $p(\lambda) = 0$. לפיכך λ שורש של p, כפי שרצינו.

ב. לקבלת התוצאה הרצויה, נשתמש ב־ (א) ובגרסה השנייה של המשפט היסודי של האלגברה (ראו 4.13). ■

לפולינום שאינו פולינום האפס יש לכל היותר מספר שורשים שונים השווה למעלתו (ראו 4.8). לפיכך סעיף (א) של המסקנה הקודמת, בצירוף המסקנה שמעלת הפולינום המינימלי של אופרטור מעל V היא לכל היותר $\dim V$, מספק הוכחה חלופית עבור 5.12, הקובעת כי לאופרטור מעל V יש לכל היותר $\dim V$ ערכים עצמיים שונים.

כל פולינום מתוקן הוא פולינום מינימלי של אופרטור כלשהו, כפי שמודגם בתרגיל 16, המכליל את דוגמה 5.26. לפיכך 5.27(א) מלמד אותנו שמציאת נוסחאות מדויקות עבור ערכים עצמיים של אופרטורים שקולה לבעיית מציאת נוסחאות מדויקות לשורשי פולינום (ולפיכך בלתי אפשרית עבור חלק מהאופרטורים).

5.28 דוגמה: אופרטור שעבורו לא ניתן למצוא ערכים עצמיים באופן מדויק

יהי $T \in \mathcal{L}(\mathbf{C}^5)$ האופרטור המוגדר על ידי

$$T(z_1, z_2, z_3, z_4, z_5) = (-3z_5, z_1 + 6z_5, z_2, z_3, z_4).$$

המטריצה של T ביחס לבסיס הסטנדרטי של $\mathbf{C}^5$ היא המטריצה 5×5 בדוגמה 5.26. כפי שהראינו בדוגמה ההיא, הפולינום המינימלי של T הוא הפולינום

$$3 - 6z + z^5.$$

לא ניתן לבטא אף אחד משורשי פולינום זה תוך שימוש במספרים רציונליים, שורשים של מספרים רציונליים, וכללי האריתמטיקה הרגילים (הוכחה לכך תחרוג משמעותית מעבר לאלגברה ליניארית). מאחר ששורשי הפולינום לעיל הם הערכים העצמיים של T [לפי 5.27(א)], אין ביכולתנו למצוא ביטוי מדויק לערכו של אף אחד מהערכים העצמיים של T בשום צורה מוכרת.

טכניקות נומריות, שלא נפרט כאן, מראות ששורשי הפולינום לעיל, ולפיכך הערכים העצמיים של T, הם חמשת המספרים המרוכבים הבאים בקירוב:

$$-1.67, \quad 0.51, \quad 1.40, \quad -0.12 + 1.59i, \quad -0.12 - 1.59i.$$

שימו לב כי שני השורשים הלא־ממשיים של הפולינום הם מרוכבים צמודים זה של זה, כפי שאנו מצפים מפולינום עם מקדמים ממשיים (ראו 4.14).

The next result completely characterizes the polynomials that when applied to an operator give the 0 operator.

5.29 $q(T) = 0 \iff q$ *is a polynomial multiple of the minimal polynomial*

Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and $q \in \mathcal{P}(\mathbf{F})$. Then $q(T) = 0$ if and only if q is a polynomial multiple of the minimal polynomial of T.

Proof Let p denote the minimal polynomial of T.

First suppose $q(T) = 0$. By the division algorithm for polynomials (4.9), there exist polynomials $s, r \in \mathcal{P}(\mathbf{F})$ such that

5.30 $$q = ps + r$$

and $\deg r < \deg p$. We have

$$0 = q(T) = p(T)s(T) + r(T) = r(T).$$

The equation above implies that $r = 0$ (otherwise, dividing r by its highest-degree coefficient would produce a monic polynomial that when applied to T gives 0; this polynomial would have a smaller degree than the minimal polynomial, which would be a contradiction). Thus 5.30 becomes the equation $q = ps$. Hence q is a polynomial multiple of p, as desired.

To prove the other direction, now suppose q is a polynomial multiple of p. Thus there exists a polynomial $s \in \mathcal{P}(\mathbf{F})$ such that $q = ps$. We have

$$q(T) = p(T)s(T) = 0\, s(T) = 0,$$

as desired. ∎

The next result is a nice consequence of the result above.

5.31 *minimal polynomial of a restriction operator*

Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and U is a subspace of V that is invariant under T. Then the minimal polynomial of T is a polynomial multiple of the minimal polynomial of $T|_U$.

Proof Suppose p is the minimal polynomial of T. Thus $p(T)v = 0$ for all $v \in V$. In particular,

$$p(T)u = 0 \text{ for all } u \in U.$$

Thus $p(T|_U) = 0$. Now 5.29, applied to the operator $T|_U$ in place of T, implies that p is a polynomial multiple of the minimal polynomial of $T|_U$. ∎

See Exercise 25 for a result about quotient operators that is analogous to the result above.

The next result shows that the constant term of the minimal polynomial of an operator determines whether the operator is invertible.

המסקנה הבאה מאפיינת באופן מלא את הפולינומים המפיקים את אופרטור האפס כאשר הם מופעלים על אופרטור.

5.29 $q(T) = 0$ אם־ורק־אם q כפולת פולינום של הפולינום המינימלי

נניח ש־ V מממד סופי, $T \in \mathcal{L}(V)$, ו־ $q \in \mathcal{P}(\mathbf{F})$. אזי $q(T) = 0$ אם־ורק־אם q הוא כפולת פולינום של הפולינום המינימלי של T.

הוכחה יהי p הפולינום המינימלי של T.

תחילה נניח ש־ $q(T) = 0$. לפי האלגוריתם לחלוקת פולינומים (4.9), קיימים פולינומים $s, r \in \mathcal{P}(\mathbf{F})$ כך ש־

$$q = ps + r \tag{5.30}$$

ובנוסף $\deg r < \deg p$. קיבלנו אם כן

$$0 = q(T) = p(T)s(T) + r(T) = r(T).$$

ממשוואה זו נובע ש־ $r = 0$ (אחרת, חלוקת r במקדם המוביל תפיק פולינום מתוקן שכאשר יופעל על T ייתן 0; לפולינום זה תהיה מעלה נמוכה יותר מזו של הפולינום המינימלי, בסתירה להגדרה). לפיכך 5.30 הופכת למשוואה $q = ps$. לכן q כפולת פולינום של p, כפי שרצינו.

להוכחת הכיוון השני, נניח כעת ש־ q כפולת פולינום של p. לפיכך קיים פולינום $s \in \mathcal{P}(\mathbf{F})$ כך ש־ $q = ps$. קיבלנו אם כן

$$q(T) = p(T)s(T) = 0\, s(T) = 0,$$

כפי שרצינו. ■

המסקנה הבאה היא תוצאה נחמדה של המסקנה לעיל.

5.31 פולינום מינימלי של אופרטור מצומצם

נניח ש־ V מממד סופי, $T \in \mathcal{L}(V)$, ו־ U תת־מרחב של V שמור תחת T. אזי הפולינום המינימלי של T הוא כפולת פולינום של הפולינום המינימלי של $T|_U$.

הוכחה נניח ש־ p הפולינום המינימלי של T. לפיכך $p(T)v = 0$ לכל $v \in V$. בפרט,

$$p(T)u = 0 \quad \text{לכל } u \in U.$$

לפיכך $p(T|_U) = 0$. כעת 5.29, מופעל על האופרטור $T|_U$ במקום על T, גורר ש־ p כפולת פולינום של הפולינום המינימלי של $T|_U$. ■

ראו תרגיל 25 לתוצאה הנוגעת לאופרטורי מנה האנלוגית למסקנה זו.

במסקנה הבאה נראה שהאיבר הקבוע של הפולינום המינימלי של אופרטור קובע האם האופרטור הפיך.

5.32 *T not invertible $\iff$ constant term of minimal polynomial of T is 0*

Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Then T is not invertible if and only if the constant term of the minimal polynomial of T is 0.

Proof Suppose $T \in \mathcal{L}(V)$ and p is the minimal polynomial of T. Then

$$\begin{aligned} T \text{ is not invertible} &\iff 0 \text{ is an eigenvalue of } T \\ &\iff 0 \text{ is a zero of } p \\ &\iff \text{the constant term of } p \text{ is } 0, \end{aligned}$$

where the first equivalence holds by 5.7, the second equivalence holds by 5.27(a), and the last equivalence holds because the constant term of p equals $p(0)$.

Eigenvalues on Odd-Dimensional Real Vector Spaces

The next result will be the key tool that we use to show that every operator on an odd-dimensional real vector space has an eigenvalue.

5.33 *even-dimensional null space*

Suppose $\mathbf{F} = \mathbf{R}$ and V is finite-dimensional. Suppose also that $T \in \mathcal{L}(V)$ and $b, c \in \mathbf{R}$ with $b^2 < 4c$. Then $\dim \operatorname{null}(T^2 + bT + cI)$ is an even number.

Proof Recall that $\operatorname{null}(T^2 + bT + cI)$ is invariant under T (by 5.18). By replacing V with $\operatorname{null}(T^2 + bT + cI)$ and replacing T with T restricted to $\operatorname{null}(T^2 + bT + cI)$, we can assume that $T^2 + bT + cI = 0$; we now need to prove that $\dim V$ is even.

Suppose $\lambda \in \mathbf{R}$ and $v \in V$ are such that $Tv = \lambda v$. Then

$$0 = (T^2 + bT + cI)v = (\lambda^2 + b\lambda + c)v = \left(\left(\lambda + \tfrac{b}{2}\right)^2 + c - \tfrac{b^2}{4}\right)v.$$

The term in large parentheses above is a positive number. Thus the equation above implies that $v = 0$. Hence we have shown that T has no eigenvectors.

Let U be a subspace of V that is invariant under T and has the largest dimension among all subspaces of V that are invariant under T and have even dimension. If $U = V$, then we are done; otherwise assume there exists $w \in V$ such that $w \notin U$.

Let $W = \operatorname{span}(w, Tw)$. Then W is invariant under T because $T(Tw) = -bTw - cw$. Furthermore, $\dim W = 2$ because otherwise w would be an eigenvector of T. Now

$$\dim(U + W) = \dim U + \dim W - \dim(U \cap W) = \dim U + 2,$$

where $U \cap W = \{0\}$ because otherwise $U \cap W$ would be a one-dimensional subspace of V that is invariant under T (impossible because T has no eigenvectors).

Because $U + W$ is invariant under T, the equation above shows that there exists a subspace of V invariant under T of even dimension larger than $\dim U$. Thus the assumption that $U \neq V$ was incorrect. Hence V has even dimension.

5.32 T אינו הפיך אם־ורק־אם האיבר הקבוע של הפולינום המינימלי של T הוא 0

נניח ש־ V מממד סופי ו־ $T \in \mathcal{L}(V)$. אזי T אינו הפיך אם־ורק־אם האיבר הקבוע של הפולינום המינימלי של T הוא 0.

הוכחה נניח ש־ $T \in \mathcal{L}(V)$ ו־ p הפולינום המינימלי של T. אזי

$$\begin{aligned} T \text{ אינו הפיך} &\iff 0 \text{ ערך עצמי של } T \\ &\iff 0 \text{ שורש של } p \\ &\iff \text{האיבר הקבוע של } p \text{ הוא } 0, \end{aligned}$$

כאשר השקילות הראשונה מתקיימת לפי 5.7, השקילות השנייה מתקיימת לפי 5.27(א), והשקילות האחרונה מתקיימת מאחר שהאיבר הקבוע של p שווה ל־ $p(0)$. ■

ערכים עצמיים מעל מרחבים וקטוריים ממשיים מממד אי־זוגי

המסקנה הבאה תשמש ככלי מרכזי שבו נשתמש כאשר נראה כי לכל אופרטור מעל מרחב וקטורי ממשי מממד אי־זוגי יש ערך עצמי.

5.33 מרחב אפס מממד זוגי

נניח ש־ $\mathbf{F} = \mathbf{R}$ ו־ V מממד סופי. נניח בנוסף ש־ $T \in \mathcal{L}(V)$ ו־ $b, c \in \mathbf{R}$ כאשר $b^2 < 4c$. אזי $\dim \operatorname{null}(T^2 + bT + cI)$ מספר זוגי.

הוכחה נזכיר כי $\operatorname{null}(T^2 + bT + cI)$ שמור תחת T (לפי 5.18). בהחלפת V ב־ $\operatorname{null}(T^2 + bT + cI)$ ובהחלפת T ב־ T מצומצם ל־ $\operatorname{null}(T^2 + bT + cI)$, נוכל להניח כי $T^2 + bT + cI = 0$; כעת עלינו להוכיח כי $\dim V$ מספר זוגי.

נניח ש־ $\lambda \in \mathbf{R}$ ו־ $v \in V$ כך ש־ $Tv = \lambda v$. אזי

$$0 = (T^2 + bT + cI)v = (\lambda^2 + b\lambda + c)v = \left(\left(\lambda + \tfrac{b}{2}\right)^2 + c - \tfrac{b^2}{4}\right)v.$$

הביטוי בסוגריים הגדולים לעיל הוא מספר חיובי. לפיכך מהמשוואה נובע ש־ $v = 0$. לכן הראינו של־ T אין וקטורים עצמיים.

יהי U תת־מרחב של V השמור תחת T מהממד הגבוה ביותר מבין כל תתי־המרחב של V השמורים תחת T וממדם מספר זוגי. אם $U = V$, בזאת סיימנו; אחרת, נניח שקיים $w \in V$ כך ש־ $w \notin U$.

יהי $W = \operatorname{span}(w, Tw)$. אזי W שמור תחת T שכן $T(Tw) = -bTw - cw$. בנוסף, $\dim W = 2$ שכן אחרת w היה וקטור עצמי של T. כעת נקבל

$$\dim(U + W) = \dim U + \dim W - \dim(U \cap W) = \dim U + 2,$$

כאשר $U \cap W = \{0\}$ שכן אחרת $U \cap W$ היה תת־מרחב חד־ממדי של V השמור תחת T (דבר בלתי אפשרי לאור העובדה שאין ל־ T וקטורים עצמיים).

מאחר ש־ $U + W$ שמור תחת T, מהמשוואה לעיל אנו רואים כי קיים תת־מרחב של V השמור תחת T מממד זוגי הגדול מ־ $\dim U$. לפיכך ההנחה ש־ $U \neq V$ הייתה שגויה. לכן V מממד זוגי. ■

The next result states that on odd-dimensional vector spaces, every operator has an eigenvalue. We already know this result for finite-dimensional complex vector spaces (without the odd hypothesis). Thus in the proof below, we will assume that $\mathbf{F} = \mathbf{R}$.

5.34 *operators on odd-dimensional vector spaces have eigenvalues*

Every operator on an odd-dimensional vector space has an eigenvalue.

Proof Suppose $\mathbf{F} = \mathbf{R}$ and V is finite-dimensional. Let $n = \dim V$, and suppose n is an odd number. Let $T \in \mathcal{L}(V)$. We will use induction on n in steps of size two to show that T has an eigenvalue. To get started, note that the desired result holds if $\dim V = 1$ because then every nonzero vector in V is an eigenvector of T.

Now suppose that $n \geq 3$ and the desired result holds for all operators on all odd-dimensional vector spaces of dimension less than n. Let p denote the minimal polynomial of T. If p is a polynomial multiple of $x - \lambda$ for some $\lambda \in \mathbf{R}$, then λ is an eigenvalue of T [by 5.27(a)] and we are done. Thus we can assume that there exist $b, c \in \mathbf{R}$ such that $b^2 < 4c$ and p is a polynomial multiple of $x^2 + bx + c$ (see 4.16).

There exists a monic polynomial $q \in \mathcal{P}(\mathbf{R})$ such that $p(x) = q(x)(x^2 + bx + c)$ for all $x \in \mathbf{R}$. Now

$$0 = p(T) = (q(T))(T^2 + bT + cI),$$

which means that $q(T)$ equals 0 on $\operatorname{range}(T^2 + bT + cI)$. Because $\deg q < \deg p$ and p is the minimal polynomial of T, this implies that $\operatorname{range}(T^2 + bT + cI) \neq V$.

The fundamental theorem of linear maps (3.21) tells us that

$$\dim V = \dim \operatorname{null}(T^2 + bT + cI) + \dim \operatorname{range}(T^2 + bT + cI).$$

Because $\dim V$ is odd (by hypothesis) and $\dim \operatorname{null}(T^2 + bT + cI)$ is even (by 5.33), the equation above shows that $\dim \operatorname{range}(T^2 + bT + cI)$ is odd.

Hence $\operatorname{range}(T^2 + bT + cI)$ is a subspace of V that is invariant under T (by 5.18) and that has odd dimension less than $\dim V$. Our induction hypothesis now implies that T restricted to $\operatorname{range}(T^2 + bT + cI)$ has an eigenvalue, which means that T has an eigenvalue. ∎

See Exercise 23 in Section 8B and Exercise 10 in Section 9C for alternative proofs of the result above.

Exercises 5B

1 Suppose $T \in \mathcal{L}(V)$. Prove that 9 is an eigenvalue of T^2 if and only if 3 or -3 is an eigenvalue of T.

2 Suppose V is a complex vector space and $T \in \mathcal{L}(V)$ has no eigenvalues. Prove that every subspace of V invariant under T is either $\{0\}$ or infinite-dimensional.

המסקנה הבאה קובעת כי במרחב וקטורי מממד אי־זוגי, לכל אופרטור יש ערך עצמי. מסקנה זו ידועה לנו כבר עבור מרחבים וקטוריים מרוכבים מממד סופי (ללא הנחת הממד האי־זוגי). לפיכך, בהוכחה להלן, נניח ש־ $\mathbf{F} = \mathbf{R}$.

5.34 לאופרטורים מעל מרחב וקטורי מממד אי־זוגי יש ערכים עצמיים

לכל אופרטור מעל מרחב וקטורי מממד אי־זוגי יש ערך עצמי.

הוכחה נניח ש־ $\mathbf{F} = \mathbf{R}$ ו־ V מממד סופי. יהי $n = \dim V$, ונניח ש־ n מספר אי־זוגי. יהי $T \in \mathcal{L}(V)$. נשתמש באינדוקציה על n בדילוגים של שניים כדי להראות כי ל־ T יש ערך עצמי. ראשית, נשים לב כי אם $\dim V = 1$ התוצאה הרצויה מתקיימת, שכן אז, כל וקטור השונה מאפס ב־ V הוא וקטור עצמי של T.

כעת נניח ש־ $n \geq 3$ ושהתוצאה הרצויה מתקיימת עבור כל האופרטורים מעל כל המרחבים הווקטוריים מממד אי־זוגי שממדם קטן מ־ n. יהי p הפולינום המינימלי של T. אם p כפולת פולינום של $x - \lambda$ עבור $\lambda \in \mathbf{R}$ כלשהו, אזי λ ערך עצמי של T [לפי 5.27(א)] ובזאת סיימנו. לפיכך נוכל להניח כי קיימים $b, c \in \mathbf{R}$ כך ש־ $b^2 < 4c$ ו־ p כפולת פולינום של $x^2 + bx + c$ (ראו 4.16).

קיים פולינום מתוקן $q \in \mathcal{P}(\mathbf{R})$ כך ש־ $p(x) = q(x)(x^2 + bx + c)$ לכל $x \in \mathbf{R}$. כעת נקבל

$$0 = p(T) = (q(T))(T^2 + bT + cI),$$

משמעות הדבר היא ש־ $q(T)$ שווה 0 מעל $\operatorname{range}(T^2 + bT + cI)$. מאחר ש־ $\deg q < \deg p$ ו־ p הוא הפולינום המינימלי של T, מכאן נובע ש־ $\operatorname{range}(T^2 + bT + cI) \neq V$.

המשפט היסודי של ההעתקות הליניאריות (3.21) קובע כי

$$\dim V = \dim \operatorname{null}(T^2 + bT + cI) + \dim \operatorname{range}(T^2 + bT + cI).$$

מאחר ש־ $\dim V$ אי־זוגי (לפי ההנחה) ו־ $\dim \operatorname{null}(T^2 + bT + cI)$ זוגי (לפי 5.33), המשוואה לעיל מראה ש־ $\dim \operatorname{range}(T^2 + bT + cI)$ אי־זוגי.

לכן $\operatorname{range}(T^2 + bT + cI)$ תת־מרחב של V השמור תחת T (לפי 5.18) וממדו אי־זוגי הקטן מ־ $\dim V$. לפי הנחת האינדוקציה, לאופרטור T כאשר הוא מצומצם ל־ $\operatorname{range}(T^2 + bT + cI)$ יש ערך עצמי, ומשמעות הדבר היא כי יש ל־ T ערך עצמי. ■

ראו תרגיל 23 ביחידה ח2 ותרגיל 10 ביחידה ט3 להוכחות חלופיות למסקנה לעיל.

תרגילים ה2

1 נניח ש־ $T \in \mathcal{L}(V)$. הוכיחו ש־ 9 ערך עצמי של T^2 אם־ורק־אם 3 או -3 ערך עצמי של T.

2 נניח ש־ V מרחב וקטורי מרוכב ו־ $T \in \mathcal{L}(V)$ אופרטור ללא ערכים עצמיים. הוכיחו כי כל תת־מרחב של V השמור תחת T הוא או $\{0\}$ או מממד אינסופי.

3 Suppose n is an integer with $n > 1$ and $T \in \mathcal{L}(\mathbf{F}^n)$ is defined by

$$T(x_1, \dots, x_n) = (x_1 + \cdots + x_n, \dots, x_1 + \cdots + x_n).$$

(a) Find all eigenvalues and eigenvectors of T.
(b) Find the minimal polynomial of T.

The matrix of T with respect to the standard basis of $\mathbf{F}^n$ consists of all 1's.

4 Suppose $\mathbf{F} = \mathbf{C}$, $T \in \mathcal{L}(V)$, $p \in \mathcal{P}(\mathbf{C})$ is a nonconstant polynomial, and $\alpha \in \mathbf{C}$. Prove that α is an eigenvalue of $p(T)$ if and only if $\alpha = p(\lambda)$ for some eigenvalue λ of T.

5 Give an example of an operator on $\mathbf{R}^2$ that shows the result in Exercise 4 does not hold if $\mathbf{C}$ is replaced with $\mathbf{R}$.

6 Suppose $T \in \mathcal{L}(\mathbf{F}^2)$ is defined by $T(w, z) = (-z, w)$. Find the minimal polynomial of T.

7 (a) Give an example of $S, T \in \mathcal{L}(\mathbf{F}^2)$ such that the minimal polynomial of ST does not equal the minimal polynomial of TS.
(b) Suppose V is finite-dimensional and $S, T \in \mathcal{L}(V)$. Prove that if at least one of S, T is invertible, then the minimal polynomial of ST equals the minimal polynomial of TS.

Hint: Show that if S is invertible and $p \in \mathcal{P}(\mathbf{F})$, then $p(TS) = S^{-1}p(ST)S$.

8 Suppose $T \in \mathcal{L}(\mathbf{R}^2)$ is the operator of counterclockwise rotation by 1°. Find the minimal polynomial of T.

Because $\dim \mathbf{R}^2 = 2$, the degree of the minimal polynomial of T is at most 2. Thus the minimal polynomial of T is not the tempting polynomial $x^{180} + 1$, even though $T^{180} = -I$.

9 Suppose $T \in \mathcal{L}(V)$ is such that with respect to some basis of V, all entries of the matrix of T are rational numbers. Explain why all coefficients of the minimal polynomial of T are rational numbers.

10 Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and $v \in V$. Prove that

$$\operatorname{span}(v, Tv, \dots, T^m v) = \operatorname{span}(v, Tv, \dots, T^{\dim V - 1} v)$$

for all integers $m \geq \dim V - 1$.

11 Suppose V is a two-dimensional vector space, $T \in \mathcal{L}(V)$, and the matrix of T with respect to some basis of V is $\begin{pmatrix} a & c \\ b & d \end{pmatrix}$.

(a) Show that $T^2 - (a + d)T + (ad - bc)I = 0$.
(b) Show that the minimal polynomial of T equals

$$\begin{cases} z - a & \text{if } b = c = 0 \text{ and } a = d, \\ z^2 - (a + d)z + (ad - bc) & \text{otherwise.} \end{cases}$$

3 נניח ש־ n מספר שלם עם $n > 1$ ו־ $T \in \mathcal{L}(\mathbf{F}^n)$ מוגדר על ידי
$$T(x_1, \dots, x_n) = (x_1 + \cdots + x_n, \dots, x_1 + \cdots + x_n).$$
א. מצאו את כל הערכים העצמיים והווקטורים העצמיים של T.
ב. מצאו את הפולינום המינימלי של T.

המטריצה של T ביחס לבסיס הסטנדרטי של $\mathbf{F}^n$ איבריה כולם 1־ים.

4 נניח ש־ $\mathbf{F} = \mathbf{C}$, $T \in \mathcal{L}(V)$, $p \in \mathcal{P}(\mathbf{C})$ פולינום שאינו קבוע, ו־ $\alpha \in \mathbf{C}$. הוכיחו ש־ α ערך עצמי של $p(T)$ אם־ורק־אם $\alpha = p(\lambda)$ עבור λ ערך עצמי כלשהו של T.

5 הביאו דוגמה לאופרטור מעל $\mathbf{R}^2$ המראה כי התוצאה בתרגיל 4 אינה מתקיימת אם נחליף את $\mathbf{C}$ ב־ $\mathbf{R}$.

6 נניח ש־ $T \in \mathcal{L}(\mathbf{F}^2)$ מוגדר על ידי $T(w, z) = (-z, w)$. מצאו את הפולינום המינימלי של T.

7 א. הביאו דוגמה של $S, T \in \mathcal{L}(\mathbf{F}^2)$ כך שהפולינום המינימלי של ST אינו שווה לפולינום המינימלי של TS.
ב. נניח ש־ V מממד סופי ו־ $S, T \in \mathcal{L}(V)$. הוכיחו שאם לפחות אחד מ־ S, T הפיך, אזי הפולינום המינימלי של ST שווה לפולינום המינימלי של TS.

רמז: הראו שאם S הפיך ו־ $p \in \mathcal{P}(\mathbf{F})$, אזי $p(TS) = S^{-1}p(ST)S$.

8 נניח ש־ $T \in \mathcal{L}(\mathbf{R}^2)$ אופרטור המסובב נגד כיוון השעון ב־ $1°$. מצאו את הפולינום המינימלי של T.

מאחר ש־ $\dim \mathbf{R}^2 = 2$, מעלת הפולינום המינימלי של T היא לכל היותר 2. לפיכך הפולינום המינימלי של T אינו זה שהיינו מתפתים לחשוב עליו $x^{180} + 1$, אף ש־ $T^{180} = -I$.

9 נניח ש־ $T \in \mathcal{L}(V)$ כך שביחס לבסיס כלשהו של V, כל איברי המטריצה של T מספרים רציונליים. הסבירו מדוע כל המקדמים של הפולינום המינימלי של T מספרים רציונליים.

10 נניח ש־ V מממד סופי, $T \in \mathcal{L}(V)$, ו־ $v \in V$. הוכיחו ש־
$$\operatorname{span}(v, Tv, \dots, T^m v) = \operatorname{span}(v, Tv, \dots, T^{\dim V - 1} v)$$
לכל המספרים השלמים $m \geq \dim V - 1$.

11 נניח ש־ V מרחב וקטורי דו־ממדי, $T \in \mathcal{L}(V)$, והמטריצה של T ביחס לבסיס כלשהו של V היא $\begin{pmatrix} a & c \\ b & d \end{pmatrix}$.

א. הראו ש־ $T^2 - (a + d)T + (ad - bc)I = 0$.
ב. הראו שהפולינום המינימלי של T שווה ל־
$$\begin{cases} z - a & \text{if } b = c = 0 \text{ and } a = d, \\ z^2 - (a + d)z + (ad - bc) & \text{otherwise.} \end{cases}$$

אלגברה ליניארית בדרך הישר, מהדורה רביעית, מחבר: שלדון אקסלר
Linear Algebra Done Right, fourth edition, by Sheldon Axler

12 Define $T \in \mathcal{L}(\mathbf{F}^n)$ by $T(x_1, x_2, x_3, \dots, x_n) = (x_1, 2x_2, 3x_3, \dots, nx_n)$. Find the minimal polynomial of T.

13 Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and $p \in \mathcal{P}(\mathbf{F})$. Prove that there exists a unique $r \in \mathcal{P}(\mathbf{F})$ such that $p(T) = r(T)$ and $\deg r$ is less than the degree of the minimal polynomial of T.

14 Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$ has minimal polynomial $4 + 5z - 6z^2 - 7z^3 + 2z^4 + z^5$. Find the minimal polynomial of T^{-1}.

15 Suppose V is a finite-dimensional complex vector space with $\dim V > 0$ and $T \in \mathcal{L}(V)$. Define $f : \mathbf{C} \to \mathbf{R}$ by

$$f(\lambda) = \dim \operatorname{range}(T - \lambda I).$$

Prove that f is not a continuous function.

16 Suppose $a_0, \dots, a_{n-1} \in \mathbf{F}$. Let T be the operator on $\mathbf{F}^n$ whose matrix (with respect to the standard basis) is

$$\begin{pmatrix} 0 & & & & -a_0 \\ 1 & 0 & & & -a_1 \\ & 1 & \ddots & & -a_2 \\ & & \ddots & & \vdots \\ & & & 0 & -a_{n-2} \\ & & & 1 & -a_{n-1} \end{pmatrix}.$$

Here all entries of the matrix are 0 except for all 1's on the line under the diagonal and the entries in the last column (some of which might also be 0). Show that the minimal polynomial of T is the polynomial

$$a_0 + a_1 z + \cdots + a_{n-1} z^{n-1} + z^n.$$

The matrix above is called the ***companion matrix*** *of the polynomial above. This exercise shows that every monic polynomial is the minimal polynomial of some operator. Hence a formula or an algorithm that could produce exact eigenvalues for each operator on each* $\mathbf{F}^n$ *could then produce exact zeros for each polynomial [by 5.27(a)]. Thus there is no such formula or algorithm. However, efficient numerical methods exist for obtaining very good approximations for the eigenvalues of an operator.*

17 Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and p is the minimal polynomial of T. Suppose $\lambda \in \mathbf{F}$. Show that the minimal polynomial of $T - \lambda I$ is the polynomial q defined by $q(z) = p(z + \lambda)$.

18 Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and p is the minimal polynomial of T. Suppose $\lambda \in \mathbf{F} \backslash \{0\}$. Show that the minimal polynomial of λT is the polynomial q defined by $q(z) = \lambda^{\deg p} p\left(\frac{z}{\lambda}\right)$.

12 הגדירו $T \in \mathcal{L}(\mathbf{F}^n)$ על ידי $T(x_1, x_2, x_3, \dots, x_n) = (x_1, 2x_2, 3x_3, \dots, nx_n)$. מצאו את הפולינום המינימלי של T.

13 נניח ש־ V מממד סופי, $T \in \mathcal{L}(V)$, ו־ $p \in \mathcal{P}(\mathbf{F})$. הוכיחו כי קיים פולינום יחיד $r \in \mathcal{P}(\mathbf{F})$ כך ש־ $p(T) = r(T)$ ו־ $\deg r$ קטן ממעלת הפולינום המינימלי של T.

14 נניח ש־ V מרחב וקטורי מממד סופי והפולינום המינימלי של $T \in \mathcal{L}(V)$ הוא $4 + 5z - 6z^2 - 7z^3 + 2z^4 + z^5$. מצאו את הפולינום המינימלי של T^{-1}.

15 נניח ש־ V מרחב וקטורי מרוכב מממד סופי, $\dim V > 0$ ו־ $T \in \mathcal{L}(V)$. נגדיר $f\colon \mathbf{C} \to \mathbf{R}$ על ידי

$$f(\lambda) = \dim \operatorname{range}(T - \lambda I).$$

הוכיחו ש־ f פונקציה שאינה רציפה.

16 נניח ש־ $a_0, \dots, a_{n-1} \in \mathbf{F}$. יהי T אופרטור מעל $\mathbf{F}^n$ שהמטריצה שלו (ביחס לבסיס הסטנדרטי) היא

$$\begin{pmatrix} 0 & & & & -a_0 \\ 1 & 0 & & & -a_1 \\ & 1 & \ddots & & -a_2 \\ & & \ddots & & \vdots \\ & & & 0 & -a_{n-2} \\ & & & 1 & -a_{n-1} \end{pmatrix}.$$

במטריצה זו, כל האיברים הם 0 מלבד כל ה־ 1־ים באלכסון שמתחת לאלכסון הראשי והאיברים בעמודה האחרונה (שייתכן שחלקם 0). הראו שהפולינום המינימלי של T הוא הפולינום

$$a_0 + a_1 z + \cdots + a_{n-1} z^{n-1} + z^n.$$

המטריצה לעיל נקראת **מטריצה מלווה** לפולינום לעיל. תרגיל זה מראה שכל פולינום מתוקן הוא פולינום מינימלי של אופרטור כלשהו. לכן, נוסחה או אלגוריתם המסוגלים למצוא באופן מדויק ערכים עצמיים לכל אופרטור מעל כל $\mathbf{F}^n$ היו יכולים למצוא באופן מדויק שורשים עבור כל פולינום [לפי 5.27(א)]. לפיכך אין נוסחה או אלגוריתם כאלו. בכל זאת, קיימים אלגוריתמים יעילים באנליזה נומרית המסוגלים למצוא קירובים טובים מאוד לערכים העצמיים של אופרטור.

17 נניח ש־ V מממד סופי, $T \in \mathcal{L}(V)$, ו־ p הפולינום המינימלי של T. נניח ש־ $\lambda \in \mathbf{F}$. הראו שהפולינום המינימלי של $T - \lambda I$ הוא הפולינום q המוגדר על ידי $q(z) = p(z + \lambda)$.

18 נניח ש־ V מממד סופי, $T \in \mathcal{L}(V)$, ו־ p הוא הפולינום המינימלי של T. נניח ש־ $\lambda \in \mathbf{F} \backslash \{0\}$. הראו שהפולינום המינימלי של λT הוא הפולינום q המוגדר על ידי $q(z) = \lambda^{\deg p} p\left(\frac{z}{\lambda}\right)$.

19 Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Let $\mathcal{E}$ be the subspace of $\mathcal{L}(V)$ defined by

$$\mathcal{E} = \{q(T) : q \in \mathcal{P}(\mathbf{F})\}.$$

Prove that $\dim \mathcal{E}$ equals the degree of the minimal polynomial of T.

20 Suppose $T \in \mathcal{L}(\mathbf{F}^4)$ is such that the eigenvalues of T are $3, 5, 8$. Prove that $(T - 3I)^2(T - 5I)^2(T - 8I)^2 = 0$.

21 Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Prove that the minimal polynomial of T has degree at most $1 + \dim \operatorname{range} T$.

If $\dim \operatorname{range} T < \dim V - 1$, then this exercise gives a better upper bound than 5.22 for the degree of the minimal polynomial of T.

22 Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Prove that T is invertible if and only if $I \in \operatorname{span}(T, T^2, \dots, T^{\dim V})$.

23 Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Let $n = \dim V$. Prove that if $v \in V$, then $\operatorname{span}(v, Tv, \dots, T^{n-1}v)$ is invariant under T.

24 Suppose V is a finite-dimensional complex vector space. Suppose $T \in \mathcal{L}(V)$ is such that 5 and 6 are eigenvalues of T and that T has no other eigenvalues. Prove that $(T - 5I)^{\dim V - 1}(T - 6I)^{\dim V - 1} = 0$.

25 Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and U is a subspace of V that is invariant under T.

(a) Prove that the minimal polynomial of T is a polynomial multiple of the minimal polynomial of the quotient operator T/U.

(b) Prove that

$$(\text{minimal polynomial of } T|_U) \times (\text{minimal polynomial of } T/U)$$

is a polynomial multiple of the minimal polynomial of T.

The quotient operator T/U was defined in Exercise 38 in Section 5A.

26 Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and U is a subspace of V that is invariant under T. Prove that the set of eigenvalues of T equals the union of the set of eigenvalues of $T|_U$ and the set of eigenvalues of T/U.

27 Suppose $\mathbf{F} = \mathbf{R}$, V is finite-dimensional, and $T \in \mathcal{L}(V)$. Prove that the minimal polynomial of $T_{\mathbf{C}}$ equals the minimal polynomial of T.

The complexification $T_{\mathbf{C}}$ was defined in Exercise 33 of Section 3B.

28 Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Prove that the minimal polynomial of $T' \in \mathcal{L}(V')$ equals the minimal polynomial of T.

The dual map T' was defined in Section 3F.

29 Show that every operator on a finite-dimensional vector space of dimension at least two has an invariant subspace of dimension two.

Exercise 6 in Section 5C will give an improvement of this result when $\mathbf{F} = \mathbf{C}$.

19 נניח ש־ V מממד סופי ו־ $T \in \mathcal{L}(V)$. יהי $\mathcal{E}$ התת־מרחב של $\mathcal{L}(V)$ המוגדר על ידי

$$\mathcal{E} = \{q(T) : q \in \mathcal{P}(\mathbf{F})\}.$$

הוכיחו ש־ $\dim \mathcal{E}$ שווה למעלת הפולינום המינימלי של T.

20 נניח ש־ $T \in \mathcal{L}(\mathbf{F}^4)$ הוא כך שהערכים העצמיים של T הם $3, 5, 8$. הוכיחו ש־ $(T - 3I)^2(T - 5I)^2(T - 8I)^2 = 0$.

21 נניח ש־ V מממד סופי ו־ $T \in \mathcal{L}(V)$. הוכיחו שמעלת הפולינום המינימלי של T היא לכל היותר $1 + \dim \operatorname{range} T$.

אם $\dim \operatorname{range} T < \dim V - 1$, אזי תרגיל זה נותן חסם עליון טוב יותר מזה המוגדר ב־ 5.22 למעלת הפולינום המינימלי של T.

22 נניח ש־ V מממד סופי ו־ $T \in \mathcal{L}(V)$. הוכיחו ש־ T הפיך אם־ורק־אם $I \in \operatorname{span}(T, T^2, \dots, T^{\dim V})$.

23 נניח ש־ V מממד סופי ו־ $T \in \mathcal{L}(V)$. יהי $n = \dim V$. הוכיחו שאם $v \in V$, אזי $\operatorname{span}(v, Tv, \dots, T^{n-1}v)$ שמור תחת T.

24 נניח ש־ V מרחב וקטורי מרוכב מממד סופי. נניח ש־ $T \in \mathcal{L}(V)$ כך ש־ 5 ו־ 6 ערכים עצמיים של T וכי ל־ T אין ערכים עצמיים נוספים. הוכיחו כי $(T - 5I)^{\dim V - 1}(T - 6I)^{\dim V - 1} = 0$.

25 נניח ש־ V מממד סופי, $T \in \mathcal{L}(V)$, ו־ U תת־מרחב של V השמור תחת T.

א. הוכיחו שהפולינום המינימלי של T הוא כפולת פולינום של הפולינום המינימלי של אופרטור המנה T/U.

ב. הוכיחו כי אם p_1 הוא הפולינום המינימלי של $T|_U$ ו־ p_2 הוא הפולינום המינימלי של T/U, אזי $p_1 p_2$ הוא כפולת פולינום של הפולינום המינימלי של T.

אופרטור המנה T/U הוגדר בתרגיל 38 ביחידה ה1.

26 נניח ש־ V מממד סופי, $T \in \mathcal{L}(V)$, ו־ U תת־מרחב של V השמור תחת T. הוכיחו שקבוצת הערכים העצמיים של T שווה לאיחוד של קבוצת הערכים העצמיים של $T|_U$ וקבוצת הערכים העצמיים של T/U.

27 נניח ש־ V מממד סופי, $\mathbf{F} = \mathbf{R}$, ו־ $T \in \mathcal{L}(V)$. הוכיחו שהפולינום המינימלי של $T_{\mathbf{C}}$ שווה לפולינום המינימלי של T.

המירכוב $T_{\mathbf{C}}$ הוגדר בתרגיל 33 ביחידה ג2.

28 נניח ש־ V מממד סופי ו־ $T \in \mathcal{L}(V)$. הוכיחו שהפולינום המינימלי של $T' \in \mathcal{L}(V')$ שווה לפולינום המינימלי של T.

ההעתקה הדואלית T' הוגדרה ביחידה ג6.

29 הראו שלכל אופרטור מעל מרחב וקטורי מממד סופי שממדו לפחות שניים יש תת־מרחב שמור דו־ממדי.

תרגיל 6 ביחידה ה3 ישפר את התוצאה כאשר $\mathbf{F} = \mathbf{C}$.

5C Upper-Triangular Matrices

In Chapter 3 we defined the matrix of a linear map from a finite-dimensional vector space to another finite-dimensional vector space. That matrix depends on a choice of basis of each of the two vector spaces. Now that we are studying operators, which map a vector space to itself, the emphasis is on using only one basis.

5.35 definition: *matrix of an operator, $\mathcal{M}(T)$*

Suppose $T \in \mathcal{L}(V)$. The *matrix of* T with respect to a basis $v_1, \dots, v_n$ of V is the n-by-n matrix

$$\mathcal{M}(T) = \begin{pmatrix} A_{1,1} & \cdots & A_{1,n} \\ \vdots & & \vdots \\ A_{n,1} & \cdots & A_{n,n} \end{pmatrix}$$

whose entries $A_{j,k}$ are defined by

$$Tv_k = A_{1,k}v_1 + \cdots + A_{n,k}v_n.$$

The notation $\mathcal{M}(T, (v_1, \dots, v_n))$ is used if the basis is not clear from the context.

Operators have square matrices (meaning that the number of rows equals the number of columns), rather than the more general rectangular matrices that we considered earlier for linear maps.

The kth column of the matrix $\mathcal{M}(T)$ is formed from the coefficients used to write Tv_k as a linear combination of the basis $v_1, \dots, v_n$.

If T is an operator on $\mathbf{F}^n$ and no basis is specified, assume that the basis in question is the standard one (where the kth basis vector is 1 in the kth slot and 0 in all other slots). You can then think of the kth column of $\mathcal{M}(T)$ as T applied to the kth basis vector, where we identify n-by-1 column vectors with elements of $\mathbf{F}^n$.

5.36 example: *matrix of an operator with respect to standard basis*

Define $T \in \mathcal{L}(\mathbf{F}^3)$ by $T(x, y, z) = (2x + y, 5y + 3z, 8z)$. Then the matrix of T with respect to the standard basis of $\mathbf{F}^3$ is

$$\mathcal{M}(T) = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 5 & 3 \\ 0 & 0 & 8 \end{pmatrix},$$

as you should verify.

A central goal of linear algebra is to show that given an operator T on a finite-dimensional vector space V, there exists a basis of V with respect to which T has a reasonably simple matrix. To make this vague formulation a bit more precise, we might try to choose a basis of V such that $\mathcal{M}(T)$ has many 0's.

הג מטריצות משולשיות עליונות

בפרק ג הגדרנו את מטריצת ההעתקה הליניארית ממרחב וקטורי מממד סופי אחד למרחב וקטורי מממד סופי אחר. מטריצה זו תלויה בבחירת בסיס עבור כל אחד משני המרחבים הווקטוריים. כעת כשאנו לומדים על אופרטורים המעתיקים מרחב וקטורי לעצמו, הדגש הוא על בחירת בסיס אחד בלבד.

5.35 הגדרה: מטריצת אופרטור, $\mathcal{M}(T)$

נניח ש־ $T \in \mathcal{L}(V)$. המטריצה של T ביחס לבסיס $v_1, \dots, v_n$ של V היא המטריצה $n \times n$

$$\mathcal{M}(T) = \begin{pmatrix} A_{1,1} & \cdots & A_{1,n} \\ \vdots & & \vdots \\ A_{n,1} & \cdots & A_{n,n} \end{pmatrix}$$

שאיבריה $A_{j,k}$ מוגדרים על ידי

$$Tv_k = A_{1,k}v_1 + \cdots + A_{n,k}v_n.$$

נשתמש בסימון $\mathcal{M}(T, (v_1, \dots, v_n))$ כאשר הבסיס אינו ברור מההקשר.

מטריצות של אופרטורים הן מטריצות ריבועיות (כלומר מספר השורות שווה למספר העמודות), בשונה מהמטריצות המלבניות הכלליות יותר שאותן הגדרנו קודם לכן עבור העתקות ליניאריות.

העמודה ה־ k^{th} של המטריצה $\mathcal{M}(T)$ נבנית מהמקדמים שבהם השתמשנו כדי לבטא את Tv_k כצירוף ליניארי של הבסיס $v_1, \dots, v_n$.

אם T אופרטור מעל $\mathbf{F}^n$ ולא צוין בסיס, הניחו כי הבסיס המדובר הוא הבסיס הסטנדרטי (כאשר בווקטור הבסיס ה־ k^{th} יש 1 ברכיב ה־ k^{th} ו־ 0 ביתר הרכיבים). ניתן לחשוב על העמודה ה־ k^{th} של $\mathcal{M}(T)$ כהפעלת T על וקטור הבסיס ה־ k^{th}, כאשר אנו מזהים וקטורי עמודה $n \times 1$ עם איברים ב־ $\mathbf{F}^n$.

5.36 דוגמה: מטריצה של אופרטור ביחס לבסיס הסטנדרטי

נגדיר $T \in \mathcal{L}(\mathbf{F}^3)$ על ידי $T(x, y, z) = (2x + y, 5y + 3z, 8z)$. אזי המטריצה של T ביחס לבסיס הסטנדרטי של $\mathbf{F}^3$ היא

$$\mathcal{M}(T) = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 5 & 3 \\ 0 & 0 & 8 \end{pmatrix},$$

כפי שעליכם לוודא.

מטרה מרכזית באלגברה ליניארית היא להראות כי בהינתן אופרטור T מעל מרחב וקטורי מממד סופי V, קיים בסיס של V אשר ביחס אליו, ל־ T יש מטריצה פשוטה באופן יחסי. כדי להפוך ניסוח מעורפל זה למדויק מעט יותר, נוכל לנסות לבחור בסיס של V כך שבמטריצה $\mathcal{M}(T)$ יהיו הרבה אפסים.

If V is a finite-dimensional complex vector space, then we already know enough to show that there is a basis of V with respect to which the matrix of T has 0's everywhere in the first column, except possibly the first entry. In other words, there is a basis of V with respect to which the matrix of T looks like

$$\begin{pmatrix} \lambda & & \\ 0 & & \\ \vdots & * & \\ 0 & & \end{pmatrix};$$

here $*$ denotes the entries in all columns other than the first column. To prove this, let λ be an eigenvalue of T (one exists by 5.19) and let v be a corresponding eigenvector. Extend v to a basis of V. Then the matrix of T with respect to this basis has the form above. Soon we will see that we can choose a basis of V with respect to which the matrix of T has even more 0's.

5.37 definition: *diagonal of a matrix*

The *diagonal* of a square matrix consists of the entries on the line from the upper left corner to the bottom right corner.

For example, the diagonal of the matrix

$$\mathcal{M}(T) = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 5 & 3 \\ 0 & 0 & 8 \end{pmatrix}$$

from Example 5.36 consists of the entries $2, 5, 8$, which are shown in red in the matrix above.

5.38 definition: *upper-triangular matrix*

A square matrix is called *upper triangular* if all entries below the diagonal are 0.

For example, the 3-by-3 matrix above is upper triangular.

Typically we represent an upper-triangular matrix in the form

$$\begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix};$$

the 0 in the matrix above indicates that all entries below the diagonal in this n-by-n matrix equal 0. Upper-triangular matrices can be considered reasonably simple—if n is large, then at least almost half the entries in an n-by-n upper-triangular matrix are 0.

We often use $*$ to denote matrix entries that we do not know or that are irrelevant to the questions being discussed.

אם V מרחב וקטורי מרוכב מממד סופי, אזי יש בידינו מספיק כלים להראות כי קיים בסיס של V אשר ביחס אליו למטריצה של T יש אפסים בכל מקום בעמודה הראשונה, מלבד אולי באיבר הראשון. במילים אחרות, קיים בסיס של V אשר ביחס אליו המטריצה של T נראית כך

$$\begin{pmatrix} \lambda & & \\ 0 & & * \\ \vdots & & \\ 0 & & \end{pmatrix};$$

ה־ $*$ כאן מייצגת את כל האיברים במטריצה מלבד איברי העמודה הראשונה. כדי להוכיח זאת, יהי λ ערך עצמי של T (לפי 5.19 קיים אחד) ויהי v וקטור עצמי השייך לערך עצמי זה. נרחיב את v לבסיס של V. אזי המטריצה של T ביחס לבסיס זה היא מהצורה שתוארה לעיל. בקרוב נראה שאנו יכולים לבחור בסיס של V אשר ביחס אליו, למטריצה של T יהיו עוד יותר אפסים.

5.37 הגדרה: אלכסון המטריצה

האלכסון של מטריצה ריבועית מורכב מהאיברים שעל הקו הנמתח מהפינה השמאלית העליונה אל הפינה הימנית התחתונה.

לדוגמה, אלכסון המטריצה

$$\mathcal{M}(T) = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 5 & 3 \\ 0 & 0 & 8 \end{pmatrix}$$

מדוגמה 5.36 מורכב מהאיברים $2, 5, 8$, המוצגים באדום במטריצה לעיל.

5.38 הגדרה: מטריצה משולשית עליונה

מטריצה ריבועית נקראת מטריצה משולשית עליונה אם כל איבריה תחת האלכסון אפסים.

לדוגמה, המטריצה 3×3 לעיל היא משולשית עליונה.

ברוב המקרים, נציג מטריצה משולשית עליונה בצורה הבאה

$$\begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix};$$

ה־ 0 במטריצה לעיל מסמן כי כל האיברים תחת האלכסון במטריצה $n \times n$ זו אפסים. מטריצות משולשיות עליונות נחשבות יחסית פשוטות—אם n מספר גדול, אזי לפחות כמעט מחצית מאיברי מטריצה $n \times n$ משולשית עליונה הם אפסים.

אנו משתמשים לעתים קרובות ב־ $*$ לסמן איברי מטריצה שאין לנו ידע לגביהם או שאינם רלוונטיים לדיון.

The next result provides a useful connection between upper-triangular matrices and invariant subspaces.

5.39 *conditions for upper-triangular matrix*

Suppose $T \in \mathcal{L}(V)$ and $v_1, \dots, v_n$ is a basis of V. Then the following are equivalent.

(a) The matrix of T with respect to $v_1, \dots, v_n$ is upper triangular.

(b) $\operatorname{span}(v_1, \dots, v_k)$ is invariant under T for each $k = 1, \dots, n$.

(c) $Tv_k \in \operatorname{span}(v_1, \dots, v_k)$ for each $k = 1, \dots, n$.

Proof First suppose (a) holds. To prove that (b) holds, suppose $k \in \{1, \dots, n\}$. If $j \in \{1, \dots, n\}$, then

$$Tv_j \in \operatorname{span}(v_1, \dots, v_j)$$

because the matrix of T with respect to $v_1, \dots, v_n$ is upper triangular. Because $\operatorname{span}(v_1, \dots, v_j) \subseteq \operatorname{span}(v_1, \dots, v_k)$ if $j \le k$, we see that

$$Tv_j \in \operatorname{span}(v_1, \dots, v_k)$$

for each $j \in \{1, \dots, k\}$. Thus $\operatorname{span}(v_1, \dots, v_k)$ is invariant under T, completing the proof that (a) implies (b).

Now suppose (b) holds, so $\operatorname{span}(v_1, \dots, v_k)$ is invariant under T for each $k = 1, \dots, n$. In particular, $Tv_k \in \operatorname{span}(v_1, \dots, v_k)$ for each $k = 1, \dots, n$. Thus (b) implies (c).

Now suppose (c) holds, so $Tv_k \in \operatorname{span}(v_1, \dots, v_k)$ for each $k = 1, \dots, n$. This means that when writing each Tv_k as a linear combination of the basis vectors $v_1, \dots, v_n$, we need to use only the vectors $v_1, \dots, v_k$. Hence all entries under the diagonal of $\mathcal{M}(T)$ are 0. Thus $\mathcal{M}(T)$ is an upper-triangular matrix, completing the proof that (c) implies (a).

We have shown that (a) $\implies$ (b) $\implies$ (c) $\implies$ (a), which shows that (a), (b), and (c) are equivalent. ∎

The next result tells us that if $T \in \mathcal{L}(V)$ and with respect to some basis of V we have

$$\mathcal{M}(T) = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix},$$

then T satisfies a simple equation depending on $\lambda_1, \dots, \lambda_n$.

5.40 *equation satisfied by operator with upper-triangular matrix*

Suppose $T \in \mathcal{L}(V)$ and V has a basis with respect to which T has an upper-triangular matrix with diagonal entries $\lambda_1, \dots, \lambda_n$. Then

$$(T - \lambda_1 I) \cdots (T - \lambda_n I) = 0.$$

המסקנה הבאה מספקת קשר שימושי בין מטריצות משולשיות עליונות לבין תתי־מרחב שמורים.

5.39 תנאים למטריצה משולשית עליונה

נניח ש־ $T \in \mathcal{L}(V)$ ו־ $v_1, \dots, v_n$ בסיס של V. אזי התנאים הבאים שקולים.

א. המטריצה של T ביחס ל־ $v_1, \dots, v_n$ משולשית עליונה.

ב. $\operatorname{span}(v_1, \dots, v_k)$ שמור תחת T לכל $k = 1, \dots, n$.

ג. $Tv_k \in \operatorname{span}(v_1, \dots, v_k)$ לכל $k = 1, \dots, n$.

הוכחה ראשית נניח ש־ (א) מתקיים. להוכחה ש־ (ב) מתקיים, נניח ש־ $k \in \{1, \dots, n\}$. אם $j \in \{1, \dots, n\}$, אזי

$$Tv_j \in \operatorname{span}(v_1, \dots, v_j)$$

הואיל והמטריצה של T ביחס ל־ $v_1, \dots, v_n$ משולשית עליונה. שכן $\operatorname{span}(v_1, \dots, v_j) \subseteq \operatorname{span}(v_1, \dots, v_k)$ אם $j \le k$, אנו רואים כי

$$Tv_j \in \operatorname{span}(v_1, \dots, v_k)$$

לכל $j \in \{1, \dots, k\}$. לפיכך, $\operatorname{span}(v_1, \dots, v_k)$ שמור תחת T, ובזאת הושלמה ההוכחה ש־ (א) גורר את (ב).

כעת נניח ש־ (ב) מתקיים; אזי $\operatorname{span}(v_1, \dots, v_k)$ שמור תחת T לכל $k = 1, \dots, n$. בפרט, $Tv_k \in \operatorname{span}(v_1, \dots, v_k)$ לכל $k = 1, \dots, n$. לפיכך (ב) גורר את (ג).

כעת נניח ש־ (ג) מתקיים; אזי $Tv_k \in \operatorname{span}(v_1, \dots, v_k)$ לכל $k = 1, \dots, n$. משמעות הדבר היא כי בביטוי של כל Tv_k כצירוף ליניארי של וקטורי הבסיס $v_1, \dots, v_n$, אנו זקוקים רק לווקטורים $v_1, \dots, v_k$. לכן כל האיברים תחת האלכסון של $\mathcal{M}(T)$ הם אפסים. לפיכך $\mathcal{M}(T)$ מטריצה משולשית עליונה, ובזאת הושלמה ההוכחה ש־ (ג) גורר את (א).

הראינו כי (א) $\Leftarrow$ (ב) $\Leftarrow$ (ג) $\Leftarrow$ (א), מה שמראה כי התנאים (א), (ב), ו־ (ג) שקולים. ∎

המסקנה הבאה קובעת כי אם $T \in \mathcal{L}(V)$ וביחס לבסיס כלשהו של V יש לנו

$$\mathcal{M}(T) = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix},$$

אזי T מקיים משוואה פשוטה התלויה ב־ $\lambda_1, \dots, \lambda_n$.

5.40 משוואה שאופרטור עם מטריצה משולשית עליונה מקיים

נניח ש־ $T \in \mathcal{L}(V)$ ול־ V יש בסיס אשר ביחס אליו יש ל־ T מטריצה משולשית עליונה שאיברי האלכסון הם $\lambda_1, \dots, \lambda_n$. אזי

$$(T - \lambda_1 I) \cdots (T - \lambda_n I) = 0.$$

Proof Let $v_1, \dots, v_n$ denote a basis of V with respect to which T has an upper-triangular matrix with diagonal entries $\lambda_1, \dots, \lambda_n$. Then $Tv_1 = \lambda_1 v_1$, which means that $(T - \lambda_1 I)v_1 = 0$, which implies that $(T - \lambda_1 I) \cdots (T - \lambda_m I)v_1 = 0$ for $m = 1, \dots, n$ (using the commutativity of each $T - \lambda_j I$ with each $T - \lambda_k I$).

Note that $(T - \lambda_2 I)v_2 \in \operatorname{span}(v_1)$. Thus $(T - \lambda_1 I)(T - \lambda_2 I)v_2 = 0$ (by the previous paragraph), which implies that $(T - \lambda_1 I) \cdots (T - \lambda_m I)v_2 = 0$ for $m = 2, \dots, n$ (using the commutativity of each $T - \lambda_j I$ with each $T - \lambda_k I$).

Note that $(T - \lambda_3 I)v_3 \in \operatorname{span}(v_1, v_2)$. Thus by the previous paragraph, $(T - \lambda_1 I)(T - \lambda_2 I)(T - \lambda_3 I)v_3 = 0$, which implies that $(T - \lambda_1 I) \cdots (T - \lambda_m I)v_3 = 0$ for $m = 3, \dots, n$ (using the commutativity of each $T - \lambda_j I$ with each $T - \lambda_k I$).

Continuing this pattern, we see that $(T - \lambda_1 I) \cdots (T - \lambda_n I)v_k = 0$ for each $k = 1, \dots, n$. Thus $(T - \lambda_1 I) \cdots (T - \lambda_n I)$ is the 0 operator because it is 0 on each vector in a basis of V. ∎

Unfortunately no method exists for exactly computing the eigenvalues of an operator from its matrix. However, if we are fortunate enough to find a basis with respect to which the matrix of the operator is upper triangular, then the problem of computing the eigenvalues becomes trivial, as the next result shows.

5.41 *determination of eigenvalues from upper-triangular matrix*

Suppose $T \in \mathcal{L}(V)$ has an upper-triangular matrix with respect to some basis of V. Then the eigenvalues of T are precisely the entries on the diagonal of that upper-triangular matrix.

Proof Suppose $v_1, \dots, v_n$ is a basis of V with respect to which T has an upper-triangular matrix

$$\mathcal{M}(T) = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

Because $Tv_1 = \lambda_1 v_1$, we see that λ_1 is an eigenvalue of T.

Suppose $k \in \{2, \dots, n\}$. Then $(T - \lambda_k I)v_k \in \operatorname{span}(v_1, \dots, v_{k-1})$. Thus $T - \lambda_k I$ maps $\operatorname{span}(v_1, \dots, v_k)$ into $\operatorname{span}(v_1, \dots, v_{k-1})$. Because

$$\dim \operatorname{span}(v_1, \dots, v_k) = k \quad \text{and} \quad \dim \operatorname{span}(v_1, \dots, v_{k-1}) = k - 1,$$

this implies that $T - \lambda_k I$ restricted to $\operatorname{span}(v_1, \dots, v_k)$ is not injective (by 3.22). Thus there exists $v \in \operatorname{span}(v_1, \dots, v_k)$ such that $v \neq 0$ and $(T - \lambda_k I)v = 0$. Thus λ_k is an eigenvalue of T. Hence we have shown that every entry on the diagonal of $\mathcal{M}(T)$ is an eigenvalue of T.

To prove T has no other eigenvalues, let q be the polynomial defined by $q(z) = (z - \lambda_1) \cdots (z - \lambda_n)$. Then $q(T) = 0$ (by 5.40). Hence q is a polynomial multiple of the minimal polynomial of T (by 5.29). Thus every zero of the minimal polynomial of T is a zero of q. Because the zeros of the minimal polynomial of T are the eigenvalues of T (by 5.27), this implies that every eigenvalue of T is a zero of q. Hence the eigenvalues of T are all contained in the list $\lambda_1, \dots, \lambda_n$. ∎

הוכחה יהי $v_1, \dots, v_n$ בסיס של V אשר ביחס אליו, ל־ T יש מטריצה משולשית עליונה עם איברי אלכסון $\lambda_1, \dots, \lambda_n$. אזי $Tv_1 = \lambda_1 v_1$, משמעות הדבר היא ש־ $(T - \lambda_1 I)v_1 = 0$, מכאן נובע ש־ $(T - \lambda_1 I) \cdots (T - \lambda_m I)v_1 = 0$ עבור $m = 1, \dots, n$ (תוך הפעלת התכונה הקומוטטיבית של כל $T - \lambda_j I$ עם כל $T - \lambda_k I$).

נשים לב כי $(T - \lambda_2 I)v_2 \in \operatorname{span}(v_1)$. לפיכך $(T - \lambda_1 I)(T - \lambda_2 I)v_2 = 0$ (לפי הפסקה הקודמת), מכאן נובע ש־ $(T - \lambda_1 I) \cdots (T - \lambda_m I)v_2 = 0$ עבור $m = 2, \dots, n$ (תוך הפעלת התכונה הקומוטטיבית של כל $T - \lambda_j I$ עם כל $T - \lambda_k I$).

נשים לב כי $(T - \lambda_3 I)v_3 \in \operatorname{span}(v_1, v_2)$. לפיכך, לפי הפסקה הקודמת, $(T - \lambda_1 I)(T - \lambda_2 I)(T - \lambda_3 I)v_3 = 0$, מכאן נובע ש־ $(T - \lambda_1 I) \cdots (T - \lambda_m I)v_3 = 0$ עבור $m = 3, \dots, n$ (תוך הפעלת התכונה הקומוטטיבית של כל $T - \lambda_j I$ עם כל $T - \lambda_k I$).

נמשיך את התבנית הזו ונראה ש־ $(T - \lambda_1 I) \cdots (T - \lambda_n I)v_k = 0$ לכל $k = 1, \dots, n$. לפיכך $(T - \lambda_1 I) \cdots (T - \lambda_n I)$ הוא אופרטור האפס שכן הוא שווה 0 על כל וקטור בבסיס של V. ■

למרבה הצער לא קיימת שיטה לחישוב מדויק של ערכים עצמיים של אופרטור על פי המטריצה שלו. בכל זאת, אם התמזל מזלנו ומצאנו בסיס אשר ביחס אליו מטריצת האופרטור היא משולשית עליונה, אזי בעיית חישוב הערכים העצמיים של האופרטור הופכת לטריוויאלית, כפי שמראה המסקנה הבאה.

5.41 קביעת ערכים עצמיים על פי מטריצה משולשית עליונה

נניח ש־ $T \in \mathcal{L}(V)$ יש מטריצה משולשית עליונה ביחס לבסיס כלשהו של V. אזי הערכים העצמיים של T הם בדיוק איברי האלכסון של אותה מטריצה משולשית עליונה.

הוכחה נניח ש־ $v_1, \dots, v_n$ בסיס של V אשר ביחס אליו ל־ T יש מטריצה משולשית עליונה

$$\mathcal{M}(T) = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

מאחר ש־ $Tv_1 = \lambda_1 v_1$, אנו רואים ש־ λ_1 ערך עצמי של T.

נניח ש־ $k \in \{2, \dots, n\}$. אזי $(T - \lambda_k I)v_k \in \operatorname{span}(v_1, \dots, v_{k-1})$. לפיכך $T - \lambda_k I$ מעתיק את $\operatorname{span}(v_1, \dots, v_k)$ אל $\operatorname{span}(v_1, \dots, v_{k-1})$. מאחר ש־

$$\dim \operatorname{span}(v_1, \dots, v_{k-1}) = k - 1, \quad \text{וגם} \quad \dim \operatorname{span}(v_1, \dots, v_k) = k$$

מכאן נובע ש־ $T - \lambda_k I$ מצומצם ל־ $\operatorname{span}(v_1, \dots, v_k)$ אינו חד־חד־ערכי (לפי 3.22). לפיכך, קיים $v \in \operatorname{span}(v_1, \dots, v_k)$ כך ש־ $v \neq 0$ ו־ $(T - \lambda_k I)v = 0$. לפיכך λ_k ערך עצמי של T. לכן הראינו כי כל איבר באלכסון המטריצה $\mathcal{M}(T)$ הוא ערך עצמי של T.

להוכחה כי ל־ T אין ערכים עצמיים אחרים, יהי q הפולינום המוגדר על ידי $q(z) = (z - \lambda_1) \cdots (z - \lambda_n)$. לכן $q(T) = 0$ (לפי 5.40). לכן q כפולת פולינום של הפולינום המינימלי של T (לפי 5.29). לפיכך כל שורש של הפולינום המינימלי של T הוא שורש של q. מאחר שהשורשים של הפולינום המינימלי של T הם הערכים העצמיים של T (לפי 5.27), מכאן נובע שכל ערך עצמי של T הוא שורש של q. לכן כל הערכים העצמיים של T כלולים ברשימה $\lambda_1, \dots, \lambda_n$. ■

5.42 example: *eigenvalues via an upper-triangular matrix*

Define $T \in \mathcal{L}(\mathbf{F}^3)$ by $T(x, y, z) = (2x + y, 5y + 3z, 8z)$. The matrix of T with respect to the standard basis is

$$\mathcal{M}(T) = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 5 & 3 \\ 0 & 0 & 8 \end{pmatrix}.$$

Now 5.41 implies that the eigenvalues of T are 2, 5, and 8.

The next example illustrates 5.44: an operator has an upper-triangular matrix with respect to some basis if and only if the minimal polynomial of the operator is the product of polynomials of degree 1.

5.43 example: *whether T has an upper-triangular matrix can depend on $\mathbf{F}$*

Define $T \in \mathcal{L}(\mathbf{F}^4)$ by

$$T(z_1, z_2, z_3, z_4) = (-z_2, z_1, 2z_1 + 3z_3, z_3 + 3z_4).$$

Thus with respect to the standard basis of $\mathbf{F}^4$, the matrix of T is

$$\begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 2 & 0 & 3 & 0 \\ 0 & 0 & 1 & 3 \end{pmatrix}.$$

You can ask a computer to verify that the minimal polynomial of T is the polynomial p defined by

$$p(z) = 9 - 6z + 10z^2 - 6z^3 + z^4.$$

First consider the case $\mathbf{F} = \mathbf{R}$. Then the polynomial p factors as

$$p(z) = (z^2 + 1)(z - 3)(z - 3),$$

with no further factorization of $z^2 + 1$ as the product of two polynomials of degree 1 with real coefficients. Thus 5.44 states that there does not exist a basis of $\mathbf{R}^4$ with respect to which T has an upper-triangular matrix.

Now consider the case $\mathbf{F} = \mathbf{C}$. Then the polynomial p factors as

$$p(z) = (z - i)(z + i)(z - 3)(z - 3),$$

where all factors above have the form $z - \lambda_k$. Thus 5.44 states that there is a basis of $\mathbf{C}^4$ with respect to which T has an upper-triangular matrix. Indeed, you can verify that with respect to the basis $(4 - 3i, -3 - 4i, -3 + i, 1)$, $(4 + 3i, -3 + 4i, -3 - i, 1)$, $(0, 0, 0, 1)$, $(0, 0, 1, 0)$ of $\mathbf{C}^4$, the operator T has the upper-triangular matrix

$$\begin{pmatrix} i & 0 & 0 & 0 \\ 0 & -i & 0 & 0 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

5.42 דוגמה: ערכים עצמיים באמצעות מטריצה משולשית עליונה

נגדיר $T \in \mathcal{L}(\mathbf{F}^3)$ על ידי $T(x,y,z) = (2x+y, 5y+3z, 8z)$. המטריצה של T ביחס לבסיס הסטנדרטי היא

$$\mathcal{M}(T) = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 5 & 3 \\ 0 & 0 & 8 \end{pmatrix}.$$

ממסקנה 5.41 נובע שהערכים העצמיים של T הם 2, 5, ו־ 8.

הדוגמה הבאה ממחישה את 5.44: לאופרטור יש מטריצה משולשית עליונה ביחס לבסיס כלשהו אם־ורק־אם הפולינום המינימלי של האופרטור הוא מכפלה של פולינומים ממעלה 1.

5.43 דוגמה: השאלה אם ל־ T יש מטריצה משולשית עליונה עשויה להיות תלויה ב־ $\mathbf{F}$

נגדיר $T \in \mathcal{L}(\mathbf{F}^4)$ על ידי

$$T(z_1, z_2, z_3, z_4) = (-z_2, z_1, 2z_1 + 3z_3, z_3 + 3z_4).$$

לפיכך, ביחס לבסיס הסטנדרטי של $\mathbf{F}^4$, המטריצה של T היא

$$\begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 2 & 0 & 3 & 0 \\ 0 & 0 & 1 & 3 \end{pmatrix}.$$

תוכלו להיעזר במחשב כדי לוודא שהפולינום המינימלי של T הוא הפולינום p המוגדר על ידי

$$p(z) = 9 - 6z + 10z^2 - 6z^3 + z^4.$$

תחילה נשקול את המקרה של $\mathbf{F} = \mathbf{R}$. אזי הפולינום p מתפרק כך

$$p(z) = (z^2+1)(z-3)(z-3),$$

ללא אפשרות לפרק עוד את z^2+1 למכפלה של שני פולינומים ממעלה 1 עם מקדמים ממשיים. לפיכך 5.44 קובעת שלא קיים בסיס של $\mathbf{R}^4$ אשר ביחס אליו, ל־ T יש מטריצה משולשית עליונה.

כעת נשקול את המקרה של $\mathbf{F} = \mathbf{C}$. אזי הפולינום p מתפרק כך

$$p(z) = (z-i)(z+i)(z-3)(z-3),$$

כאשר כל הגורמים לעיל הם מהצורה $z - \lambda_k$. לפיכך 5.44 קובעת כי קיים בסיס של $\mathbf{C}^4$ אשר ביחס אליו, יש ל־ T מטריצה משולשית עליונה. אכן, תוכלו לוודא כי ביחס לבסיס $(4-3i, -3-4i, -3+i, 1), (4+3i, -3+4i, -3-i, 1), (0,0,0,1), (0,0,1,0)$ של $\mathbf{C}^4$, לאופרטור T יש מטריצה משולשית עליונה

$$\begin{pmatrix} i & 0 & 0 & 0 \\ 0 & -i & 0 & 0 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

5.44 *necessary and sufficient condition to have an upper-triangular matrix*

Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Then T has an upper-triangular matrix with respect to some basis of V if and only if the minimal polynomial of T equals $(z - \lambda_1) \cdots (z - \lambda_m)$ for some $\lambda_1, \ldots, \lambda_m \in \mathbf{F}$.

Proof First suppose T has an upper-triangular matrix with respect to some basis of V. Let $\alpha_1, \ldots, \alpha_n$ denote the diagonal entries of that matrix. Define a polynomial $q \in \mathcal{P}(\mathbf{F})$ by

$$q(z) = (z - \alpha_1) \cdots (z - \alpha_n).$$

Then $q(T) = 0$, by 5.40. Hence q is a polynomial multiple of the minimal polynomial of T, by 5.29. Thus the minimal polynomial of T equals $(z - \lambda_1) \cdots (z - \lambda_m)$ for some $\lambda_1, \ldots, \lambda_m \in \mathbf{F}$ with $\{\lambda_1, \ldots, \lambda_m\} \subseteq \{\alpha_1, \ldots, \alpha_n\}$.

To prove the implication in the other direction, now suppose the minimal polynomial of T equals $(z - \lambda_1) \cdots (z - \lambda_m)$ for some $\lambda_1, \ldots, \lambda_m \in \mathbf{F}$. We will use induction on m. To get started, if $m = 1$ then $z - \lambda_1$ is the minimal polynomial of T, which implies that $T = \lambda_1 I$, which implies that the matrix of T (with respect to any basis of V) is upper triangular.

Now suppose $m > 1$ and the desired result holds for all smaller positive integers. Let

$$U = \operatorname{range}(T - \lambda_m I).$$

Then U is invariant under T [this is a special case of 5.18 with $p(z) = z - \lambda_m$]. Thus $T|_U$ is an operator on U.

If $u \in U$, then $u = (T - \lambda_m I)v$ for some $v \in V$ and

$$(T - \lambda_1 I) \cdots (T - \lambda_{m-1} I)u = (T - \lambda_1 I) \cdots (T - \lambda_m I)v = 0.$$

Hence $(z - \lambda_1) \cdots (z - \lambda_{m-1})$ is a polynomial multiple of the minimal polynomial of $T|_U$, by 5.29. Thus the minimal polynomial of $T|_U$ is the product of at most $m - 1$ terms of the form $z - \lambda_k$.

By our induction hypothesis, there is a basis $u_1, \ldots, u_M$ of U with respect to which $T|_U$ has an upper-triangular matrix. Thus for each $k \in \{1, \ldots, M\}$, we have (using 5.39)

$$Tu_k = (T|_U)(u_k) \in \operatorname{span}(u_1, \ldots, u_k). \tag{5.45}$$

Extend $u_1, \ldots, u_M$ to a basis $u_1, \ldots, u_M, v_1, \ldots, v_N$ of V. If $k \in \{1, \ldots, N\}$, then

$$Tv_k = (T - \lambda_m I)v_k + \lambda_m v_k.$$

The definition of U shows that $(T - \lambda_m I)v_k \in U = \operatorname{span}(u_1, \ldots, u_M)$. Thus the equation above shows that

$$Tv_k \in \operatorname{span}(u_1, \ldots, u_M, v_1, \ldots, v_k). \tag{5.46}$$

From 5.45 and 5.46, we conclude (using 5.39) that T has an upper-triangular matrix with respect to the basis $u_1, \ldots, u_M, v_1, \ldots, v_N$ of V, as desired. ∎

5.44 תנאי הכרחי ומספיק לקיומה של מטריצה משולשית עליונה

נניח ש־ V מממד סופי ו־ $T \in \mathcal{L}(V)$. אזי יש ל־ T מטריצה משולשית עליונה ביחס לבסיס כלשהו של V אם־ורק־אם הפולינום המינימלי של T שווה ל־ $(z-\lambda_1)\cdots(z-\lambda_m)$ עבור $\lambda_1,\dots,\lambda_m \in \mathbf{F}$ כלשהם.

הוכחה ראשית נניח כי יש ל־ T מטריצה משולשית עליונה ביחס לבסיס כלשהו של V. יהיו $\alpha_1,\dots,\alpha_n$ איברי האלכסון של אותה מטריצה. נגדיר פולינום $q \in \mathcal{P}(\mathbf{F})$ על ידי

$$q(z) = (z-\alpha_1)\cdots(z-\alpha_n).$$

אזי $q(T)=0$, לפי 5.40. לכן q כפולת פולינום של הפולינום המינימלי של T, לפי 5.29. לפיכך הפולינום המינימלי של T שווה ל־ $(z-\lambda_1)\cdots(z-\lambda_m)$ עבור $\lambda_1,\dots,\lambda_m \in \mathbf{F}$ כלשהם, עם $\{\lambda_1,\dots,\lambda_m\} \subseteq \{\alpha_1,\dots,\alpha_n\}$.

כדי להוכיח את הגרירה בכיוון השני, נניח כעת כי הפולינום המינימלי של T שווה ל־ $(z-\lambda_1)\cdots(z-\lambda_m)$ עבור $\lambda_1,\dots,\lambda_m \in \mathbf{F}$. נוכיח באינדוקציה על m. כצעד ראשון, אם $m=1$ אזי $z-\lambda_1$ הוא הפולינום המינימלי של T, מה שגורר כי $T=\lambda_1 I$, ומכאן שהמטריצה של T (ביחס לכל בסיס של V) היא משולשית עליונה.

כעת נניח ש־ $m>1$ והתוצאה הרצויה מתקיימת עבור כל מספר שלם חיובי הקטן מ־ m. יהי

$$U = \operatorname{range}(T-\lambda_m I).$$

אזי U שמור תחת T [זהו מקרה מיוחד של 5.18 עם $p(z)=z-\lambda_m$]. לפיכך $T|_U$ אופרטור מעל U.

אם $u \in U$, אזי $u=(T-\lambda_m I)v$ עבור $v \in V$ כלשהו, וגם

$$(T-\lambda_1 I)\cdots(T-\lambda_{m-1}I)u = (T-\lambda_1 I)\cdots(T-\lambda_m I)v = 0.$$

לכן $(z-\lambda_1)\cdots(z-\lambda_{m-1})$ כפולת פולינום של הפולינום המינימלי של $T|_U$, לפי 5.29. לפיכך הפולינום המינימלי של $T|_U$ הוא מכפלה של לכל היותר $m-1$ גורמים מהצורה $z-\lambda_k$.

לפי הנחת האינדוקציה, קיים בסיס $u_1,\dots,u_M$ של U אשר ביחס אליו יש ל־ $T|_U$ מטריצה משולשית עליונה. לפיכך, לכל $k \in \{1,\dots,M\}$, נקבל (בשימוש ב־ 5.39)

$$Tu_k = (T|_U)(u_k) \in \operatorname{span}(u_1,\dots,u_k). \tag{5.45}$$

נשלים את $u_1,\dots,u_M$ לכדי בסיס $u_1,\dots,u_M,v_1,\dots,v_N$ של V. אם $k \in \{1,\dots,N\}$, אזי

$$Tv_k = (T-\lambda_m I)v_k + \lambda_m v_k.$$

מההגדרה של U אנו מסיקים כי $(T-\lambda_m I)v_k \in U = \operatorname{span}(u_1,\dots,u_M)$. לפיכך מהמשוואה לעיל אנו רואים כי

$$Tv_k \in \operatorname{span}(u_1,\dots,u_M,v_1,\dots,v_k). \tag{5.46}$$

מ־ 5.45 ו־ 5.46, נסיק (תוך שימוש ב־ 5.39) כי ל־ T יש מטריצה משולשית עליונה ביחס לבסיס $u_1,\dots,u_M,v_1,\dots,v_N$ של V, כפי שרצינו. ■

The set of numbers $\{\lambda_1, \dots, \lambda_m\}$ from the previous result equals the set of eigenvalues of T (because the set of zeros of the minimal polynomial of T equals the set of eigenvalues of T, by 5.27), although the list $\lambda_1, \dots, \lambda_m$ in the previous result may contain repetitions.

In Chapter 8 we will improve even the wonderful result below; see 8.37 and 8.46.

5.47 *if $\mathbf{F} = \mathbf{C}$, then every operator on V has an upper-triangular matrix*

Suppose V is a finite-dimensional complex vector space and $T \in \mathcal{L}(V)$. Then T has an upper-triangular matrix with respect to some basis of V.

Proof The desired result follows immediately from 5.44 and the second version of the fundamental theorem of algebra (see 4.13). ∎

For an extension of the result above to two operators S and T such that

$$ST = TS,$$

see 5.80. Also, for an extension to more than two operators, see Exercise 9(b) in Section 5E.

Caution: If an operator $T \in \mathcal{L}(V)$ has an upper-triangular matrix with respect to some basis $v_1, \dots, v_n$ of V, then the eigenvalues of T are exactly the entries on the diagonal of $\mathcal{M}(T)$, as shown by 5.41, and furthermore v_1 is an eigenvector of T. However, $v_2, \dots, v_n$ need not be eigenvectors of T. Indeed, a basis vector v_k is an eigenvector of T if and only if all entries in the kth column of the matrix of T are 0, except possibly the kth entry.

You may recall from a previous course that every matrix of numbers can be changed to a matrix in what is called row echelon form. If one begins with a square matrix, the matrix in row echelon form will be an upper-triangular matrix. Do not confuse this upper-triangular matrix with the upper-triangular matrix of an operator with respect to some basis whose existence is proclaimed by 5.47 (if $\mathbf{F} = \mathbf{C}$)—there is no connection between these upper-triangular matrices.

The row echelon form of the matrix of an operator does not give us a list of the eigenvalues of the operator. In contrast, an upper-triangular matrix with respect to some basis gives us a list of all the eigenvalues of the operator. However, there is no method for computing exactly such an upper-triangular matrix, even though 5.47 guarantees its existence if $\mathbf{F} = \mathbf{C}$.

Exercises 5C

1 Prove or give a counterexample: If $T \in \mathcal{L}(V)$ and T^2 has an upper-triangular matrix with respect to some basis of V, then T has an upper-triangular matrix with respect to some basis of V.

קבוצת המספרים $\{\lambda_1, \dots, \lambda_m\}$ מהמסקנה לעיל שווה לקבוצת הערכים העצמיים של T (שכן קבוצת השורשים של הפולינום המינימלי של T שווה לקבוצת הערכים העצמיים של T, לפי 5.27), אף שייתכן שברשימה $\lambda_1, \dots, \lambda_m$ במסקנה הקודמת יש איברים החוזרים על עצמם. בפרק ח אנו נשפר אפילו את המסקנה הנפלאה להלן; ראו 8.37 ו־ 8.46.

5.47 אם $\mathbf{F} = \mathbf{C}$, אזי לכל אופרטור מעל V יש מטריצה משולשית עליונה

נניח ש־ V מרחב וקטורי מרוכב מממד סופי ו־ $T \in \mathcal{L}(V)$. אזי יש ל־ T מטריצה משולשית עליונה ביחס לבסיס כלשהו של V.

הוכחה התוצאה הרצויה נובעת באופן מיידי מ־ 5.44 ומהגרסה השנייה של המשפט היסודי של האלגברה (ראו 4.13). ■

להרחבת המסקנה לעיל עבור שני אופרטורים S ו־ T כך ש־

$$ST = TS,$$

ראו 5.80. בנוסף, להרחבה ליותר משני אופרטורים, ראו תרגיל 9(ב) ביחידה ה5.

זְהִירוּת: אם לאופרטור $T \in \mathcal{L}(V)$ יש מטריצה משולשית עליונה ביחס לבסיס $v_1, \dots, v_n$ כלשהו של V, אזי הערכים העצמיים של T הם בדיוק האיברים באלכסון המטריצה $\mathcal{M}(T)$, כפי שהראינו ב־ 5.41, ויתר על כן, v_1 הוא וקטור עצמי של T. אולם, $v_2, \dots, v_n$ אינם בהכרח וקטורים עצמיים של T. אכן וקטור בסיס v_k הוא וקטור עצמי של T אם־ורק־אם כל האיברים בעמודה ה־ k^{th} של המטריצה של T הם 0, מלבד אולי האיבר ה־ k^{th}.

ייתכן שאתם זוכרים מקורס קודם כי כל מטריצה ניתנת לשינוי למטריצה בצורה הקרויה צורת מדרגות באמצעות פעולות אלמנטריות. דירוג מטריצה ריבועית מוביל למטריצה משולשית עליונה. אל תבלבלו את מטריצה משולשית עליונה זו עם המטריצה המשולשית העליונה של אופרטור ביחס לבסיס כלשהו, שאת קיומה מבטיחה 5.47 (אם $\mathbf{F} = \mathbf{C}$)—אין קשר בין מטריצות משולשיות עליונות אלו.

יישום שיטת דירוג מטריצות על מטריצת האופרטור לא יניב לנו את רשימת הערכים העצמיים של האופרטור. בניגוד לכך, מטריצה משולשית עליונה ביחס לבסיס כלשהו מספקת לנו רשימה של כל הערכים העצמיים של האופרטור. עם זאת, לא קיימת שיטה לחישוב מדויק של מטריצה משולשית עליונה שכזו, זאת אף ש־ 5.47 מבטיחה את קיומה אם $\mathbf{F} = \mathbf{C}$.

תרגילים ה3

1 הוכיחו או הביאו דוגמה נגדית: אם $T \in \mathcal{L}(V)$ ויש ל־ T^2 מטריצה משולשית עליונה ביחס לבסיס כלשהו של V, אזי יש ל־ T מטריצה משולשית עליונה ביחס לבסיס כלשהו של V.

2 Suppose A and B are upper-triangular matrices of the same size, with $\alpha_1, \dots, \alpha_n$ on the diagonal of A and $\beta_1, \dots, \beta_n$ on the diagonal of B.

(a) Show that $A + B$ is an upper-triangular matrix with $\alpha_1 + \beta_1, \dots, \alpha_n + \beta_n$ on the diagonal.

(b) Show that AB is an upper-triangular matrix with $\alpha_1\beta_1, \dots, \alpha_n\beta_n$ on the diagonal.

The results in this exercise are used in the proof of 5.81.

3 Suppose $T \in \mathcal{L}(V)$ is invertible and $v_1, \dots, v_n$ is a basis of V with respect to which the matrix of T is upper triangular, with $\lambda_1, \dots, \lambda_n$ on the diagonal. Show that the matrix of T^{-1} is also upper triangular with respect to the basis $v_1, \dots, v_n$, with

$$\frac{1}{\lambda_1}, \dots, \frac{1}{\lambda_n}$$

on the diagonal.

4 Give an example of an operator whose matrix with respect to some basis contains only 0's on the diagonal, but the operator is invertible.

This exercise and the exercise below show that 5.41 fails without the hypothesis that an upper-triangular matrix is under consideration.

5 Give an example of an operator whose matrix with respect to some basis contains only nonzero numbers on the diagonal, but the operator is not invertible.

6 Suppose $\mathbf{F} = \mathbf{C}$, V is finite-dimensional, and $T \in \mathcal{L}(V)$. Prove that if $k \in \{1, \dots, \dim V\}$, then V has a k-dimensional subspace invariant under T.

7 Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and $v \in V$.

(a) Prove that there exists a unique monic polynomial p_v of smallest degree such that $p_v(T)v = 0$.

(b) Prove that the minimal polynomial of T is a polynomial multiple of p_v.

8 Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and there exists a nonzero vector $v \in V$ such that $T^2v + 2Tv = -2v$.

(a) Prove that if $\mathbf{F} = \mathbf{R}$, then there does not exist a basis of V with respect to which T has an upper-triangular matrix.

(b) Prove that if $\mathbf{F} = \mathbf{C}$ and A is an upper-triangular matrix that equals the matrix of T with respect to some basis of V, then $-1 + i$ or $-1 - i$ appears on the diagonal of A.

9 Suppose B is a square matrix with complex entries. Prove that there exists an invertible square matrix A with complex entries such that $A^{-1}BA$ is an upper-triangular matrix.

2 נניח ש־ A ו־ B מטריצות משולשיות עליונות מאותו הגודל, עם $\alpha_1, \dots, \alpha_n$ באלכסון A ו־ $\beta_1, \dots, \beta_n$ באלכסון B.

א. הראו ש־ $A + B$ מטריצה משולשית עליונה עם $\alpha_1 + \beta_1, \dots, \alpha_n + \beta_n$ כאיברי האלכסון.

ב. הראו ש־ AB מטריצה משולשית עליונה עם $\alpha_1\beta_1, \dots, \alpha_n\beta_n$ כאיברי האלכסון.

בהוכחת 5.81 נעשה שימוש בתוצאות בתרגיל זה.

3 נניח ש־ $T \in \mathcal{L}(V)$ הפיך ו־ $v_1, \dots, v_n$ בסיס של V אשר ביחס אליו המטריצה של T משולשית עליונה, עם $\lambda_1, \dots, \lambda_n$ כאיברי האלכסון. הראו שהמטריצה של T^{-1} גם היא משולשית עליונה ביחס לבסיס $v_1, \dots, v_n$, עם

$$\frac{1}{\lambda_1}, \dots, \frac{1}{\lambda_n}$$

איברי האלכסון.

4 הביאו דוגמה לאופרטור שהמטריצה שלו ביחס לבסיס כלשהו מכילה באלכסון רק אפסים, אך האופרטור הפיך.

תרגיל זה והתרגיל הבא מראים כי מסקנה 5.41 אינה תקפה ללא ההנחה כי נקודת המוצא היא קיומה של מטריצה משולשית עליונה.

5 הביאו דוגמה לאופרטור שהמטריצה שלו ביחס לבסיס כלשהו מכילה באלכסון רק איברים שאינם אפסים, אך האופרטור בלתי הפיך.

6 נניח ש־ $\mathbf{F} = \mathbf{C}$, V מממד סופי, ו־ $T \in \mathcal{L}(V)$. הוכיחו שאם $k \in \{1, \dots, \dim V\}$, אזי יש ל־ V תת־מרחב k־ממדי השמור תחת T.

7 נניח ש־ V מממד סופי, $T \in \mathcal{L}(V)$, ו־ $v \in V$.

א. הוכיחו כי קיים פולינום מתוקן יחיד p_v מהמעלה הנמוכה ביותר כך ש־ $p_v(T)v = 0$.

ב. הוכיחו שהפולינום המינימלי של T הוא כפולת פולינום של p_v.

8 נניח ש־ V מממד סופי, $T \in \mathcal{L}(V)$, וקיים וקטור $v \in V$ השונה מאפס, כך ש־ $T^2v + 2Tv = -2v$.

א. הוכיחו שאם $\mathbf{F} = \mathbf{R}$, אזי לא קיים בסיס של V אשר ביחס אליו יש ל־ T מטריצה משולשית עליונה.

ב. הוכיחו שאם $\mathbf{F} = \mathbf{C}$ ו־ A מטריצה משולשית עליונה השווה למטריצה של T ביחס לבסיס כלשהו של V, אזי $-1 + i$ או $-1 - i$ מופיע באלכסונה של A.

9 נניח ש־ B מטריצה ריבועית עם איברים מרוכבים. הוכיחו שקיימת מטריצה ריבועית הפיכה A עם איברים מרוכבים כך ש־ $A^{-1}BA$ מטריצה משולשית עליונה.

אלגברה ליניארית בדרך הישר, מהדורה רביעית, מחבר: שלדון אקסלר
Linear Algebra Done Right, fourth edition, by Sheldon Axler

10 Suppose $T \in \mathcal{L}(V)$ and $v_1, \dots, v_n$ is a basis of V. Show that the following are equivalent.

(a) The matrix of T with respect to $v_1, \dots, v_n$ is lower triangular.

(b) $\operatorname{span}(v_k, \dots, v_n)$ is invariant under T for each $k = 1, \dots, n$.

(c) $Tv_k \in \operatorname{span}(v_k, \dots, v_n)$ for each $k = 1, \dots, n$.

A square matrix is called ***lower triangular*** *if all entries above the diagonal are* 0.

11 Suppose $\mathbf{F} = \mathbf{C}$ and V is finite-dimensional. Prove that if $T \in \mathcal{L}(V)$, then there exists a basis of V with respect to which T has a lower-triangular matrix.

12 Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$ has an upper-triangular matrix with respect to some basis of V, and U is a subspace of V that is invariant under T.

(a) Prove that $T|_U$ has an upper-triangular matrix with respect to some basis of U.

(b) Prove that the quotient operator T/U has an upper-triangular matrix with respect to some basis of V/U.

The quotient operator T/U was defined in Exercise 38 in Section 5A.

13 Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Suppose there exists a subspace U of V that is invariant under T such that $T|_U$ has an upper-triangular matrix with respect to some basis of U and also T/U has an upper-triangular matrix with respect to some basis of V/U. Prove that T has an upper-triangular matrix with respect to some basis of V.

14 Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Prove that T has an upper-triangular matrix with respect to some basis of V if and only if the dual operator T' has an upper-triangular matrix with respect to some basis of the dual space V'.

10 נניח ש־ $T \in \mathcal{L}(V)$ ו־ $v_1, \dots, v_n$ בסיס של V. הראו שהתנאים הבאים שקולים.

א. המטריצה של T ביחס ל־ $v_1, \dots, v_n$ היא מטריצה משולשית תחתונה.

ב. $\operatorname{span}(v_k, \dots, v_n)$ שמור תחת T לכל $k = 1, \dots, n$.

ג. $Tv_k \in \operatorname{span}(v_k, \dots, v_n)$ לכל $k = 1, \dots, n$.

מטריצה ריבועית נקראת **מטריצה משולשית תחתונה** אם כל איבריה מעל האלכסון הם 0.

11 נניח ש־ $\mathbf{F} = \mathbf{C}$ ו־ V מממד סופי. הוכיחו שאם $T \in \mathcal{L}(V)$, אזי קיים בסיס של V אשר ביחס אליו יש ל־ T מטריצה משולשית תחתונה.

12 נניח ש־ V מממד סופי, יש ל־ $T \in \mathcal{L}(V)$ מטריצה משולשית עליונה ביחס לבסיס כלשהו של V, ו־ U תת־מרחב של V השמור תחת T.

א. הוכיחו שיש ל־ $T|_U$ מטריצה משולשית עליונה ביחס לבסיס כלשהו של U.

ב. הוכיחו שיש לאופרטור המנה T/U מטריצה משולשית עליונה ביחס לבסיס כלשהו של V/U.

אופרטור המנה T/U הוגדר בתרגיל 38 ביחידה הו1.

13 נניח ש־ V מממד סופי ו־ $T \in \mathcal{L}(V)$. נניח כי קיים תת־מרחב U של V השמור תחת T כך שיש ל־ $T|_U$ מטריצה משולשית עליונה ביחס לבסיס כלשהו של U ובנוסף, יש ל־ T/U מטריצה משולשית עליונה ביחס לבסיס כלשהו של V/U. הוכיחו שיש ל־ T מטריצה משולשית עליונה ביחס לבסיס כלשהו של V.

14 נניח ש־ V מממד סופי ו־ $T \in \mathcal{L}(V)$. הוכיחו שיש ל־ T מטריצה משולשית עליונה ביחס לבסיס כלשהו של V אם־ורק־אם יש לאופרטור הדואלי T' מטריצה משולשית עליונה ביחס לבסיס כלשהו של המרחב הדואלי V'.

5D Diagonalizable Operators

Diagonal Matrices

5.48 definition: *diagonal matrix*

A *diagonal matrix* is a square matrix that is 0 everywhere except possibly on the diagonal.

5.49 example: *diagonal matrix*

$$\begin{pmatrix} 8 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 5 \end{pmatrix}$$

is a diagonal matrix.

If an operator has a diagonal matrix with respect to some basis, then the entries on the diagonal are precisely the eigenvalues of the operator; this follows from 5.41 (or find an easier direct proof for diagonal matrices).

Every diagonal matrix is upper triangular. Diagonal matrices typically have many more 0's than most upper-triangular matrices of the same size.

5.50 definition: *diagonalizable*

An operator on V is called *diagonalizable* if the operator has a diagonal matrix with respect to some basis of V.

5.51 example: *diagonalization may require a different basis*

Define $T \in \mathcal{L}(\mathbf{R}^2)$ by

$$T(x, y) = (41x + 7y, -20x + 74y).$$

The matrix of T with respect to the standard basis of $\mathbf{R}^2$ is

$$\begin{pmatrix} 41 & 7 \\ -20 & 74 \end{pmatrix},$$

which is not a diagonal matrix. However, T is diagonalizable. Specifically, the matrix of T with respect to the basis $(1, 4), (7, 5)$ is

$$\begin{pmatrix} 69 & 0 \\ 0 & 46 \end{pmatrix}$$

because $T(1, 4) = (69, 276) = 69(1, 4)$ and $T(7, 5) = (322, 230) = 46(7, 5)$.

4ה אופרטורים לכסינים

מטריצות אלכסוניות

5.48 הגדרה: מטריצה אלכסונית

מטריצה אלכסונית היא מטריצה ריבועית שכל איבריה אפסים מלבד אולי באלכסון.

5.49 דוגמה: מטריצה אלכסונית

$$\begin{pmatrix} 8 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 5 \end{pmatrix}$$

היא מטריצה אלכסונית.

כל מטריצה אלכסונית היא משולשית עליונה. במטריצות אלכסוניות יש בדרך כלל הרבה יותר אפסים מאשר ברוב המטריצות המשולשיות העליונות באותו הגודל.

אם לאופרטור יש מטריצה אלכסונית ביחס לבסיס כלשהו, אזי האיברים באלכסון המטריצה הם בדיוק הערכים העצמיים של האופרטור; הדבר נובע מ־ 5.41 (לחילופין מצאו הוכחה ישירה קלה יותר למטריצות אלכסוניות).

5.50 הגדרה: לכסין

אופרטור מעל V נקרא לכסין אם לאופרטור יש מטריצה אלכסונית ביחס לבסיס כלשהו של V.

5.51 דוגמה: ייתכן שלכסון ידרוש בסיס אחר

נגדיר $T \in \mathcal{L}(\mathbf{R}^2)$ על ידי

$$T(x, y) = (41x + 7y, -20x + 74y).$$

המטריצה של T ביחס לבסיס הסטנדרטי של $\mathbf{R}^2$ היא

$$\begin{pmatrix} 41 & 7 \\ -20 & 74 \end{pmatrix},$$

שאינה מטריצה אלכסונית. אולם, T לכסין. ליתר דיוק, המטריצה של T ביחס לבסיס $(1, 4), (7, 5)$ היא

$$\begin{pmatrix} 69 & 0 \\ 0 & 46 \end{pmatrix}$$

מאחר ש־ $T(1, 4) = (69, 276) = 69(1, 4)$ ו־ $T(7, 5) = (322, 230) = 46(7, 5)$.

For $\lambda \in \mathbf{F}$, we will find it convenient to have a name and a notation for the set of vectors that an operator T maps to λ times the vector.

5.52 definition: *eigenspace*, $E(\lambda, T)$

Suppose $T \in \mathcal{L}(V)$ and $\lambda \in \mathbf{F}$. The *eigenspace* of T corresponding to λ is the subspace $E(\lambda, T)$ of V defined by

$$E(\lambda, T) = \operatorname{null}(T - \lambda I) = \{v \in V : Tv = \lambda v\}.$$

Hence $E(\lambda, T)$ is the set of all eigenvectors of T corresponding to λ, along with the 0 vector.

For $T \in \mathcal{L}(V)$ and $\lambda \in \mathbf{F}$, the set $E(\lambda, T)$ is a subspace of V because the null space of each linear map on V is a subspace of V. The definitions imply that λ is an eigenvalue of T if and only if $E(\lambda, T) \neq \{0\}$.

5.53 example: *eigenspaces of an operator*

Suppose the matrix of an operator $T \in \mathcal{L}(V)$ with respect to a basis v_1, v_2, v_3 of V is the matrix in Example 5.49. Then

$$E(8, T) = \operatorname{span}(v_1), \quad E(5, T) = \operatorname{span}(v_2, v_3).$$

If λ is an eigenvalue of an operator $T \in \mathcal{L}(V)$, then T restricted to $E(\lambda, T)$ is just the operator of multiplication by λ.

5.54 *sum of eigenspaces is a direct sum*

Suppose $T \in \mathcal{L}(V)$ and $\lambda_1, \ldots, \lambda_m$ are distinct eigenvalues of T. Then

$$E(\lambda_1, T) + \cdots + E(\lambda_m, T)$$

is a direct sum. Furthermore, if V is finite-dimensional, then

$$\dim E(\lambda_1, T) + \cdots + \dim E(\lambda_m, T) \leq \dim V.$$

Proof To show that $E(\lambda_1, T) + \cdots + E(\lambda_m, T)$ is a direct sum, suppose

$$v_1 + \cdots + v_m = 0,$$

where each v_k is in $E(\lambda_k, T)$. Because eigenvectors corresponding to distinct eigenvalues are linearly independent (by 5.11), this implies that each v_k equals 0. Thus $E(\lambda_1, T) + \cdots + E(\lambda_m, T)$ is a direct sum (by 1.45), as desired.

Now suppose V is finite-dimensional. Then

$$\begin{aligned}\dim E(\lambda_1, T) + \cdots + \dim E(\lambda_m, T) &= \dim(E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T)) \\ &\leq \dim V,\end{aligned}$$

where the first line follows from 3.94 and the second line follows from 2.37. ∎

עבור $\lambda \in \mathbf{F}$, יהיה לנו נוח לתת שם וסימון לקבוצת הווקטורים שאותם האופרטור T מעתיק אל λ כפול הווקטור.

5.52 הגדרה: מרחב עצמי, $E(\lambda, T)$

נניח ש־ $T \in \mathcal{L}(V)$ ו־ $\lambda \in \mathbf{F}$. המרחב העצמי של T השייך ל־ λ הוא התת־מרחב $E(\lambda, T)$ של V המוגדר על ידי

$$E(\lambda, T) = \operatorname{null}(T - \lambda I) = \{v \in V : Tv = \lambda v\}.$$

לכן $E(\lambda, T)$ הוא קבוצת כל הווקטורים העצמיים של T השייכים ל־ λ, יחד עם וקטור ה־ 0.

עבור $T \in \mathcal{L}(V)$ ו־ $\lambda \in \mathbf{F}$, הקבוצה $E(\lambda, T)$ היא תת־מרחב של V מאחר שמרחב האפס של כל העתקה ליניארית מעל V הוא תת־מרחב של V. מההגדרות נובע ש־ λ ערך עצמי של T אם־ורק־אם $E(\lambda, T) \neq \{0\}$.

5.53 דוגמה: מרחבים עצמיים של אופרטור

נניח שמטריצת האופרטור $T \in \mathcal{L}(V)$ ביחס לבסיס v_1, v_2, v_3 של V היא המטריצה בדוגמה 5.49. אזי

$$E(8, T) = \operatorname{span}(v_1), \quad E(5, T) = \operatorname{span}(v_2, v_3).$$

אם λ ערך עצמי של אופרטור $T \in \mathcal{L}(V)$, אזי T מצומצם ל־ $E(\lambda, T)$ הוא פשוט אופרטור הכפלה ב־ λ.

5.54 סכום מרחבים עצמיים הוא סכום ישר

נניח ש־ $T \in \mathcal{L}(V)$ ו־ $\lambda_1, \dots, \lambda_m$ ערכים עצמיים שונים של T. אזי

$$E(\lambda_1, T) + \cdots + E(\lambda_m, T)$$

הוא סכום ישר. בנוסף, אם V מממד סופי, אזי

$$\dim E(\lambda_1, T) + \cdots + \dim E(\lambda_m, T) \leq \dim V.$$

הוכחה כדי להראות כי $E(\lambda_1, T) + \cdots + E(\lambda_m, T)$ סכום ישר, נניח ש־

$$v_1 + \cdots + v_m = 0,$$

כאשר כל v_k ב־ $E(\lambda_k, T)$. מאחר שווקטורים עצמיים השייכים לערכים עצמיים שונים בלתי תלויים ליניארית (לפי 5.11), מכאן נובע שכל v_k שווה 0. לפיכך סכום המרחבים העצמיים $E(\lambda_1, T) + \cdots + E(\lambda_m, T)$ הוא סכום ישר (לפי 1.45), כפי שרצינו.

כעת נניח ש־ V מממד סופי ונקבל

$$\dim E(\lambda_1, T) + \cdots + \dim E(\lambda_m, T) = \dim(E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T))$$

$$\leq \dim V,$$

כאשר השורה הראשונה נובעת מ־ 3.94 והשורה השנייה מ־ 2.37. ■

Conditions for Diagonalizability

The following characterizations of diagonalizable operators will be useful.

5.55 *conditions equivalent to diagonalizability*

Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Let $\lambda_1, \dots, \lambda_m$ denote the distinct eigenvalues of T. Then the following are equivalent.

(a) T is diagonalizable.

(b) V has a basis consisting of eigenvectors of T.

(c) $V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T)$.

(d) $\dim V = \dim E(\lambda_1, T) + \cdots + \dim E(\lambda_m, T)$.

Proof An operator $T \in \mathcal{L}(V)$ has a diagonal matrix

$$\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

with respect to a basis $v_1, \dots, v_n$ of V if and only if $Tv_k = \lambda_k v_k$ for each k. Thus (a) and (b) are equivalent.

Suppose (b) holds; thus V has a basis consisting of eigenvectors of T. Hence every vector in V is a linear combination of eigenvectors of T, which implies that

$$V = E(\lambda_1, T) + \cdots + E(\lambda_m, T).$$

Now 5.54 shows that (c) holds, proving that (b) implies (c).

That (c) implies (d) follows immediately from 3.94.

Finally, suppose (d) holds; thus

$$\dim V = \dim E(\lambda_1, T) + \cdots + \dim E(\lambda_m, T). \tag{5.56}$$

Choose a basis of each $E(\lambda_k, T)$; put all these bases together to form a list $v_1, \dots, v_n$ of eigenvectors of T, where $n = \dim V$ (by 5.56). To show that this list is linearly independent, suppose

$$a_1 v_1 + \cdots + a_n v_n = 0,$$

where $a_1, \dots, a_n \in \mathbf{F}$. For each $k = 1, \dots, m$, let u_k denote the sum of all the terms $a_j v_j$ such that $v_j \in E(\lambda_k, T)$. Thus each u_k is in $E(\lambda_k, T)$, and

$$u_1 + \cdots + u_m = 0.$$

Because eigenvectors corresponding to distinct eigenvalues are linearly independent (see 5.11), this implies that each u_k equals 0. Because each u_k is a sum of terms $a_j v_j$, where the v_j's were chosen to be a basis of $E(\lambda_k, T)$, this implies that all a_j's equal 0. Thus $v_1, \dots, v_n$ is linearly independent and hence is a basis of V (by 2.38). Thus (d) implies (b), completing the proof. ∎

For additional conditions equivalent to diagonalizability, see 5.62, Exercises 5 and 15 in this section, Exercise 24 in Section 7B, and Exercise 15 in Section 8A.

תנאים ללכסינות

המאפיינים הבאים של אופרטור לכסין יהיו שימושיים.

5.55 תנאים שקולים ללכסינות

נניח ש־ V מממד סופי ו־ $T \in \mathcal{L}(V)$. יהיו $\lambda_1, \dots, \lambda_m$ הערכים העצמיים השונים של T. אזי התנאים הבאים שקולים.

א. T לכסין.

ב. יש ל־ V בסיס המורכב מווקטורים עצמיים של T.

ג. $V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T)$.

ד. $\dim V = \dim E(\lambda_1, T) + \cdots + \dim E(\lambda_m, T)$.

הוכחה לאופרטור $T \in \mathcal{L}(V)$ יש מטריצה אלכסונית

$$\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

ביחס לבסיס $v_1, \dots, v_n$ של V אם־ורק־אם $Tv_k = \lambda_k v_k$ לכל k. לפיכך (א) ו־ (ב) שקולים.

נניח ש־ (ב) מתקיים; לפיכך יש ל־ V בסיס המורכב מווקטורים עצמיים של T. לכן כל וקטור ב־ V הוא צירוף ליניארי של וקטורים עצמיים של T, מכאן נובע כי

$$V = E(\lambda_1, T) + \cdots + E(\lambda_m, T).$$

כעת 5.54 מראה ש־ (ג) מתקיים, וזו ההוכחה לכך ש־(ג) נגרר מ־ (ב).

את זה ש־ (ג) גורר את (ד) אנו מסיקים ישירות מ־ 3.94.

לבסוף, נניח ש־ (ד) מתקיים; לפיכך

$$\dim V = \dim E(\lambda_1, T) + \cdots + \dim E(\lambda_m, T). \tag{5.56}$$

נבחר בסיס של כל $E(\lambda_k, T)$; נקבץ את כל הבסיסים הללו יחד ליצירת רשימת הווקטורים העצמיים $v_1, \dots, v_n$ של T, כאשר $n = \dim V$ (לפי 5.56). כדי להראות שרשימה זו בלתי תלויה ליניארית, נניח ש־

$$a_1 v_1 + \cdots + a_n v_n = 0,$$

כאשר $a_1, \dots, a_n \in \mathbf{F}$. לכל $k = 1, \dots, m$, יהי u_k סכום כל המחוברים $a_j v_j$ כך ש־ $v_j \in E(\lambda_k, T)$. לפיכך, כל u_k הוא ב־ $E(\lambda_k, T)$, ובנוסף מתקיים

$$u_1 + \cdots + u_m = 0.$$

מאחר שווקטורים עצמיים השייכים לערכים עצמיים שונים בלתי תלויים ליניארית (ראו 5.11), מכאן נובע שכל u_k שווה 0. מאחר שכל u_k הוא סכום המחוברים $a_j v_j$, כאשר ה־ v_j־ים נבחרו כבסיס של $E(\lambda_k, T)$, מכאן נובע שכל ה־ a_j־ים שווים 0. לפיכך $v_1, \dots, v_n$ בלתי תלויה ליניארית ולכן היא בסיס של V (לפי 2.38). לפיכך (ד) גורר את (ב), ובכך השלמנו את ההוכחה.

לתנאים נוספים השקולים ללכסינות, ראו 5.62, תרגילים 5 ו־ 15 ביחידה זו, תרגיל 24 ביחידה זג, ותרגיל 15 ביחידה חו1.

As we know, every operator on a nonzero finite-dimensional complex vector space has an eigenvalue. However, not every operator on a nonzero finite-dimensional complex vector space has enough eigenvectors to be diagonalizable, as shown by the next example.

5.57 example: *an operator that is not diagonalizable*

Define an operator $T \in \mathcal{L}(\mathbf{F}^3)$ by $T(a, b, c) = (b, c, 0)$. The matrix of T with respect to the standard basis of $\mathbf{F}^3$ is

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix},$$

which is an upper-triangular matrix but is not a diagonal matrix.

As you should verify, 0 is the only eigenvalue of T and furthermore

$$E(0, T) = \{(a, 0, 0) \in \mathbf{F}^3 : a \in \mathbf{F}\}.$$

Hence conditions (b), (c), and (d) of 5.55 fail (of course, because these conditions are equivalent, it is sufficient to check that only one of them fails). Thus condition (a) of 5.55 also fails. Hence T is not diagonalizable, regardless of whether $\mathbf{F} = \mathbf{R}$ or $\mathbf{F} = \mathbf{C}$.

The next result shows that if an operator has as many distinct eigenvalues as the dimension of its domain, then the operator is diagonalizable.

5.58 *enough eigenvalues implies diagonalizability*

Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$ has $\dim V$ distinct eigenvalues. Then T is diagonalizable.

Proof Suppose T has distinct eigenvalues $\lambda_1, \dots, \lambda_{\dim V}$. For each k, let $v_k \in V$ be an eigenvector corresponding to the eigenvalue λ_k. Because eigenvectors corresponding to distinct eigenvalues are linearly independent (see 5.11), $v_1, \dots, v_{\dim V}$ is linearly independent.

A linearly independent list of $\dim V$ vectors in V is a basis of V (see 2.38); thus $v_1, \dots, v_{\dim V}$ is a basis of V. With respect to this basis consisting of eigenvectors, T has a diagonal matrix. ∎

In later chapters we will find additional conditions that imply that certain operators are diagonalizable. For example, see the real spectral theorem (7.29) and the complex spectral theorem (7.31).

The result above gives a sufficient condition for an operator to be diagonalizable. However, this condition is not necessary. For example, the operator T on $\mathbf{F}^3$ defined by $T(x, y, z) = (6x, 6y, 7z)$ has only two eigenvalues (6 and 7) and $\dim \mathbf{F}^3 = 3$, but T is diagonalizable (by the standard basis of $\mathbf{F}^3$).

כפי שאנו כבר יודעים, לכל אופרטור מעל מרחב וקטורי מרוכב מממד סופי השונה מאפס יש ערך עצמי. אך לא לכל אופרטור מעל מרחב וקטורי מרוכב מממד סופי השונה מאפס יש מספיק וקטורים עצמיים כדי שיהיה לכסין, כפי שנראה בדוגמה הבאה.

5.57 דוגמה: אופרטור שאינו לכסין

נגדיר אופרטור $T \in \mathcal{L}(\mathbf{F}^3)$ על ידי $T(a,b,c) = (b,c,0)$. המטריצה של T ביחס לבסיס הסטנדרטי של $\mathbf{F}^3$ היא

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix},$$

שהיא מטריצה משולשית עליונה אך אינה מטריצה אלכסונית.

כפי שעליכם לוודא, 0 הוא הערך העצמי היחיד של T ובנוסף

$$E(0,T) = \{(a,0,0) \in \mathbf{F}^3 : a \in \mathbf{F}\}.$$

לכן התנאים (ב), (ג), ו־ (ד) של 5.55 אינם מתקיימים (כמובן שמאחר שהתנאים הם שקולים, די היה בבדיקת אי התקיימותו של תנאי אחד בלבד). לפיכך תנאי (א) של 5.55 גם אינו מתקיים. לכן T אינו לכסין, בין אם $\mathbf{F} = \mathbf{R}$ ובין אם $\mathbf{F} = \mathbf{C}$.

המסקנה הבאה מראה שאם לאופרטור יש מספר ערכים עצמיים שונים השווה לממד התחום, אזי האופרטור לכסין.

5.58 כמות מספקת של ערכים עצמיים גוררת לכסינות

נניח ש־ V מממד סופי ומספר הערכים העצמיים השונים של $T \in \mathcal{L}(V)$ הוא $\dim V$. אזי T לכסין.

הוכחה נניח כי ל־ T ערכים עצמיים שונים $\lambda_1, \dots, \lambda_{\dim V}$. עבור כל k, יהי $v_k \in V$ וקטור עצמי השייך לערך העצמי λ_k. מאחר שווקטורים עצמיים השייכים לערכים עצמיים שונים הם בלתי תלויים ליניארית (ראו 5.11), $v_1, \dots, v_{\dim V}$ בלתי תלויה ליניארית.

רשימה בלתי תלויה ליניארית של $\dim V$ וקטורים ב־ V היא בסיס של V (ראו 2.38); לפיכך $v_1, \dots, v_{\dim V}$ בסיס של V. ביחס לבסיס זה המורכב מווקטורים עצמיים, יש ל־ T מטריצה אלכסונית. ■

בפרקי ההמשך נכיר תנאים נוספים הגוררים כי אופרטורים מסוימים לכסינים. לדוגמה ראו משפט הפירוק הספקטרלי הממשי (7.29) ומשפט הפירוק הספקטרלי המרוכב (7.31).

המסקנה לעיל מציגה תנאי מספיק לכך שאופרטור יהיה לכסין. אולם תנאי זה אינו הכרחי. לדוגמה, לאופרטור T מעל $\mathbf{F}^3$ המוגדר על ידי $T(x,y,z) = (6x, 6y, 7z)$ יש רק שני ערכים עצמיים (6 ו־ 7) ו־ $\dim \mathbf{F}^3 = 3$, ובכל זאת T לכסין (לפי הבסיס הסטנדרטי של $\mathbf{F}^3$).

The next example illustrates the importance of diagonalization, which can be used to compute high powers of an operator, taking advantage of the equation $T^k v = \lambda^k v$ if v is an eigenvector of T with eigenvalue λ.

For a spectacular application of these techniques, see Exercise 21, which shows how to use diagonalization to find an exact formula for the nth term of the Fibonacci sequence.

5.59 example: *using diagonalization to compute T^{100}*

Define $T \in \mathcal{L}(\mathbf{F}^3)$ by $T(x, y, z) = (2x + y, 5y + 3z, 8z)$. With respect to the standard basis, the matrix of T is

$$\begin{pmatrix} 2 & 1 & 0 \\ 0 & 5 & 3 \\ 0 & 0 & 8 \end{pmatrix}.$$

The matrix above is an upper-triangular matrix but it is not a diagonal matrix. By 5.41, the eigenvalues of T are 2, 5, and 8. Because T is an operator on a vector space of dimension three and T has three distinct eigenvalues, 5.58 assures us that there exists a basis of $\mathbf{F}^3$ with respect to which T has a diagonal matrix.

To find this basis, we only have to find an eigenvector for each eigenvalue. In other words, we have to find a nonzero solution to the equation

$$T(x, y, z) = \lambda(x, y, z)$$

for $\lambda = 2$, then for $\lambda = 5$, and then for $\lambda = 8$. Solving these simple equations shows that for $\lambda = 2$ we have an eigenvector $(1, 0, 0)$, for $\lambda = 5$ we have an eigenvector $(1, 3, 0)$, and for $\lambda = 8$ we have an eigenvector $(1, 6, 6)$.

Thus $(1, 0, 0), (1, 3, 0), (1, 6, 6)$ is a basis of $\mathbf{F}^3$ consisting of eigenvectors of T, and with respect to this basis the matrix of T is the diagonal matrix

$$\begin{pmatrix} 2 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 8 \end{pmatrix}.$$

To compute $T^{100}(0, 0, 1)$, for example, write $(0, 0, 1)$ as a linear combination of our basis of eigenvectors:

$$(0, 0, 1) = \tfrac{1}{6}(1, 0, 0) - \tfrac{1}{3}(1, 3, 0) + \tfrac{1}{6}(1, 6, 6).$$

Now apply T^{100} to both sides of the equation above, getting

$$\begin{aligned} T^{100}(0, 0, 1) &= \tfrac{1}{6}\big(T^{100}(1, 0, 0)\big) - \tfrac{1}{3}\big(T^{100}(1, 3, 0)\big) + \tfrac{1}{6}\big(T^{100}(1, 6, 6)\big) \\ &= \tfrac{1}{6}\big(2^{100}(1, 0, 0) - 2 \cdot 5^{100}(1, 3, 0) + 8^{100}(1, 6, 6)\big) \\ &= \tfrac{1}{6}\big(2^{100} - 2 \cdot 5^{100} + 8^{100},\ 6 \cdot 8^{100} - 6 \cdot 5^{100},\ 6 \cdot 8^{100}\big). \end{aligned}$$

הדוגמה הבאה ממחישה את חשיבותו של הלכסון, שבו ניתן להשתמש לחישוב חזקות מסדר גבוה של אופרטורים, תוך שימוש במשוואה $Tv = \lambda v$ אם v וקטור עצמי של T השייך לערך העצמי λ.

ליישום מרהיב של טכניקות אלו, ראו תרגיל 21, המראה איך ניתן להשתמש בלכסון כדי למצוא את הנוסחה המדויקת עבור האיבר ה־ n^{th} של סדרת פיבונאצ'י.

5.59 דוגמה: שימוש בלכסון לחישוב T^{100}

נגדיר $T \in \mathcal{L}(\mathbf{F}^3)$ על ידי $T(x, y, z) = (2x + y, 5y + 3z, 8z)$. ביחס לבסיס הסטנדרטי, המטריצה של T היא

$$\begin{pmatrix} 2 & 1 & 0 \\ 0 & 5 & 3 \\ 0 & 0 & 8 \end{pmatrix}.$$

מטריצה זו היא מטריצה משולשית עליונה אך אינה מטריצה אלכסונית. לפי 5.41, הערכים העצמיים של T הם 2, 5, ו־ 8. מאחר ש־ T אופרטור מעל מרחב וקטורי תלת־ממדי ויש ל־ T שלושה ערכים עצמיים שונים, 5.58 מבטיח לנו כי קיים בסיס של $\mathbf{F}^3$ אשר ביחס אליו יש ל־ T מטריצה אלכסונית.

כדי למצוא בסיס זה, כל שעלינו לעשות הוא למצוא וקטור עצמי לכל ערך עצמי. במילים אחרות, עלינו למצוא פתרון השונה מאפס למשוואה

$$T(x, y, z) = \lambda(x, y, z)$$

עבור $\lambda = 2$, אחר־כך עבור $\lambda = 5$, ואז עבור $\lambda = 8$. פתרון משוואות פשוטות אלו מראה כי עבור $\lambda = 2$ יש לנו וקטור עצמי $(1, 0, 0)$, עבור $\lambda = 5$ יש לנו וקטור עצמי $(1, 3, 0)$, ועבור $\lambda = 8$ יש לנו וקטור עצמי $(1, 6, 6)$.

לפיכך $(1, 0, 0), (1, 3, 0), (1, 6, 6)$ בסיס של $\mathbf{F}^3$ המורכב מהווקטורים העצמיים של T, וביחס לבסיס זה, המטריצה של T היא המטריצה האלכסונית

$$\begin{pmatrix} 2 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 8 \end{pmatrix}.$$

לחישוב $T^{100}(0, 0, 1)$, לדוגמה, נכתוב את $(0, 0, 1)$ כצירוף ליניארי של בסיס הווקטורים העצמיים שלנו:

$$(0, 0, 1) = \tfrac{1}{6}(1, 0, 0) - \tfrac{1}{3}(1, 3, 0) + \tfrac{1}{6}(1, 6, 6).$$

כעת נפעיל את T^{100} על שני אגפי המשוואה, ונקבל

$$\begin{aligned} T^{100}(0, 0, 1) &= \tfrac{1}{6}\big(T^{100}(1, 0, 0)\big) - \tfrac{1}{3}\big(T^{100}(1, 3, 0)\big) + \tfrac{1}{6}\big(T^{100}(1, 6, 6)\big) \\ &= \tfrac{1}{6}\big(2^{100}(1, 0, 0) - 2 \cdot 5^{100}(1, 3, 0) + 8^{100}(1, 6, 6)\big) \\ &= \tfrac{1}{6}\big(2^{100} - 2 \cdot 5^{100} + 8^{100},\ 6 \cdot 8^{100} - 6 \cdot 5^{100},\ 6 \cdot 8^{100}\big). \end{aligned}$$

We saw earlier that an operator T on a finite-dimensional vector space V has an upper-triangular matrix with respect to some basis of V if and only if the minimal polynomial of T equals $(z-\lambda_1)\cdots(z-\lambda_m)$ for some $\lambda_1,\dots,\lambda_m \in \mathbf{F}$ (see 5.44). As we previously noted (see 5.47), this condition is always satisfied if $\mathbf{F} = \mathbf{C}$.

Our next result 5.62 states that an operator $T \in \mathcal{L}(V)$ has a diagonal matrix with respect to some basis of V if and only if the minimal polynomial of T equals $(z-\lambda_1)\cdots(z-\lambda_m)$ for some ***distinct*** $\lambda_1,\dots,\lambda_m \in \mathbf{F}$. Before formally stating this result, we give two examples of using it.

5.60 example: *diagonalizable, but with no known exact eigenvalues*

Define $T \in \mathcal{L}(\mathbf{C}^5)$ by

$$T(z_1,z_2,z_3,z_4,z_5) = (-3z_5, z_1+6z_5, z_2, z_3, z_4).$$

The matrix of T is shown in Example 5.26, where we showed that the minimal polynomial of T is $3-6z+z^5$.

As mentioned in Example 5.28, no exact expression is known for any of the zeros of this polynomial, but numerical techniques show that the zeros of this polynomial are approximately $-1.67,\ 0.51,\ 1.40,\ -0.12+1.59i,\ -0.12-1.59i$.

The software that produces these approximations is accurate to more than three digits. Thus these approximations are good enough to show that the five numbers above are distinct. The minimal polynomial of T equals the fifth degree monic polynomial with these zeros. Now 5.62 shows that T is diagonalizable.

5.61 example: *showing that an operator is not diagonalizable*

Define $T \in \mathcal{L}(\mathbf{F}^3)$ by

$$T(z_1,z_2,z_3) = (6z_1+3z_2+4z_3, 6z_2+2z_3, 7z_3).$$

The matrix of T with respect to the standard basis of $\mathbf{F}^3$ is

$$\begin{pmatrix} 6 & 3 & 4 \\ 0 & 6 & 2 \\ 0 & 0 & 7 \end{pmatrix}.$$

The matrix above is an upper-triangular matrix but is not a diagonal matrix. Might T have a diagonal matrix with respect to some other basis of $\mathbf{F}^3$?

To answer this question, we will find the minimal polynomial of T. First note that the eigenvalues of T are the diagonal entries of the matrix above (by 5.41). Thus the zeros of the minimal polynomial of T are $6, 7$ [by 5.27(a)]. The diagonal of the matrix above tells us that $(T-6I)^2(T-7I) = 0$ (by 5.40). The minimal polynomial of T has degree at most 3 (by 5.22). Putting all this together, we see that the minimal polynomial of T is either $(z-6)(z-7)$ or $(z-6)^2(z-7)$.

A simple computation shows that $(T-6I)(T-7I) \neq 0$. Thus the minimal polynomial of T is $(z-6)^2(z-7)$.

Now 5.62 shows that T is not diagonalizable.

ראינו לעיל כי לאופרטור T מעל מרחב וקטורי מממד סופי V יש מטריצה משולשית עליונה ביחס לבסיס כלשהו של V אם־ורק־אם הפולינום המינימלי של T שווה ל־ $(z-\lambda_1)\cdots(z-\lambda_m)$ עבור $\lambda_1,\ldots,\lambda_m \in \mathbf{F}$ כלשהם (ראו 5.44). כפי שצוין קודם (ראו 5.47), תנאי זה תמיד מתקיים אם $\mathbf{F}=\mathbf{C}$.

מסקנתנו הבאה 5.62 קובעת כי לאופרטור $T \in \mathcal{L}(V)$ יש מטריצה אלכסונית ביחס לבסיס כלשהו של V אם־ורק־אם הפולינום המינימלי של T שווה ל־ $(z-\lambda_1)\cdots(z-\lambda_m)$ עבור $\lambda_1,\ldots,\lambda_m \in \mathbf{F}$ **שונים** כלשהם. בטרם ננסח מסקנה זו באופן רשמי, נביא שתי דוגמאות לשימוש במסקנה זו.

5.60 דוגמה: לכסין, אך הערכים העצמיים אינם יזועים במדויק

נגדיר $T \in \mathcal{L}(\mathbf{C}^5)$ על ידי

$$T(z_1,z_2,z_3,z_4,z_5) = (-3z_5, z_1+6z_5, z_2, z_3, z_4).$$

המטריצה של T היא זו המוצגת בדוגמה 5.26, שם הראינו כי הפולינום המינימלי של T הוא $3-6z+z^5$.

כפי שצוין בדוגמה 5.28, אין דרך לבטא במדויק אף אחד משורשי הפולינום, אך שיטות של אנליזה נומרית מראות לנו כי שורשי הפולינום הם המספרים הבאים בקירוב $-1.67,\ 0.51,\ 1.40,\ -0.12+1.59i,\ -0.12-1.59i$.

התוכנה שהפיקה קירובים אלו מדייקת מעבר לשלוש ספרות עשרוניות. לפיכך קירובים אלו טובים דיים להראות כי חמשת המספרים הללו שונים. הפולינום המינימלי של T שווה לפולינום מתוקן ממעלה חמישית עם שורשים אלו. כעת 5.62 מראה כי T לכסין.

5.61 דוגמה: הצגת מקרה שבו אופרטור אינו לכסין

נגדיר $T \in \mathcal{L}(\mathbf{F}^3)$ על ידי

$$T(z_1,z_2,z_3) = (6z_1+3z_2+4z_3, 6z_2+2z_3, 7z_3).$$

המטריצה של T ביחס לבסיס הסטנדרטי של $\mathbf{F}^3$ היא

$$\begin{pmatrix} 6 & 3 & 4 \\ 0 & 6 & 2 \\ 0 & 0 & 7 \end{pmatrix}.$$

מטריצה זו היא מטריצה משולשית עליונה אך אינה מטריצה אלכסונית. האם ל־ T יש מטריצה אלכסונית ביחס לבסיס אחר כלשהו של $\mathbf{F}^3$?

כדי לענות על שאלה זו, נמצא את הפולינום המינימלי של T. ראשית נשים לב כי הערכים העצמיים של T הם איברי אלכסון המטריצה (לפי 5.41). לפיכך שורשי הפולינום המינימלי של T הם $6, 7$ [לפי 5.27(א)]. אלכסון מטריצה זו מורה לנו כי $(T-6I)^2(T-7I)=0$ (לפי 5.40). מעלת הפולינום המינימלי של T היא לכל היותר 3 (לפי 5.22). בחיבור כל הנתונים הללו נקבל כי הפולינום המינימלי של T הוא או $(z-6)(z-7)$ או $(z-6)^2(z-7)$.

חישוב פשוט מראה כי $(T-6I)(T-7I) \neq 0$. לפיכך הפולינום המינימלי של T הוא $(z-6)^2(z-7)$.

כעת 5.62 מראה כי T אינו לכסין.

5.62 *necessary and sufficient condition for diagonalizability*

Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Then T is diagonalizable if and only if the minimal polynomial of T equals $(z - \lambda_1) \cdots (z - \lambda_m)$ for some list of distinct numbers $\lambda_1, \ldots, \lambda_m \in \mathbf{F}$.

Proof First suppose T is diagonalizable. Thus there is a basis $v_1, \ldots, v_n$ of V consisting of eigenvectors of T. Let $\lambda_1, \ldots, \lambda_m$ be the distinct eigenvalues of T. Then for each v_j, there exists λ_k with $(T - \lambda_k I)v_j = 0$. Thus

$$(T - \lambda_1 I) \cdots (T - \lambda_m I)v_j = 0,$$

which implies that the minimal polynomial of T equals $(z - \lambda_1) \cdots (z - \lambda_m)$.

To prove the implication in the other direction, now suppose the minimal polynomial of T equals $(z - \lambda_1) \cdots (z - \lambda_m)$ for some list of distinct numbers $\lambda_1, \ldots, \lambda_m \in \mathbf{F}$. Thus

5.63 $$(T - \lambda_1 I) \cdots (T - \lambda_m I) = 0.$$

We will prove that T is diagonalizable by induction on m. To get started, suppose $m = 1$. Then $T - \lambda_1 I = 0$, which means that T is a scalar multiple of the identity operator, which implies that T is diagonalizable.

Now suppose that $m > 1$ and the desired result holds for all smaller values of m. The subspace $\text{range}(T - \lambda_m I)$ is invariant under T [this is a special case of 5.18 with $p(z) = z - \lambda_m$]. Thus T restricted to $\text{range}(T - \lambda_m I)$ is an operator on $\text{range}(T - \lambda_m I)$.

If $u \in \text{range}(T - \lambda_m I)$, then $u = (T - \lambda_m I)v$ for some $v \in V$, and 5.63 implies

5.64 $$(T - \lambda_1 I) \cdots (T - \lambda_{m-1} I)u = (T - \lambda_1 I) \cdots (T - \lambda_m I)v = 0.$$

Hence $(z - \lambda_1) \cdots (z - \lambda_{m-1})$ is a polynomial multiple of the minimal polynomial of T restricted to $\text{range}(T - \lambda_m I)$ [by 5.29]. Thus by our induction hypothesis, there is a basis of $\text{range}(T - \lambda_m I)$ consisting of eigenvectors of T.

Suppose that $u \in \text{range}(T - \lambda_m I) \cap \text{null}(T - \lambda_m I)$. Then $Tu = \lambda_m u$. Now 5.64 implies that

$$\begin{aligned} 0 &= (T - \lambda_1 I) \cdots (T - \lambda_{m-1} I)u \\ &= (\lambda_m - \lambda_1) \cdots (\lambda_m - \lambda_{m-1})u. \end{aligned}$$

Because $\lambda_1, \ldots, \lambda_m$ are distinct, the equation above implies that $u = 0$. Hence $\text{range}(T - \lambda_m I) \cap \text{null}(T - \lambda_m I) = \{0\}$.

Thus $\text{range}(T - \lambda_m I) + \text{null}(T - \lambda_m I)$ is a direct sum (by 1.46) whose dimension is $\dim V$ (by 3.94 and 3.21). Hence $\text{range}(T - \lambda_m I) \oplus \text{null}(T - \lambda_m I) = V$. Every nonzero vector in $\text{null}(T - \lambda_m I)$ is an eigenvector of T with eigenvalue λ_m. Earlier in this proof we saw that there is a basis of $\text{range}(T - \lambda_m I)$ consisting of eigenvectors of T. Adjoining to that basis a basis of $\text{null}(T - \lambda_m I)$ gives a basis of V consisting of eigenvectors of T. The matrix of T with respect to this basis is a diagonal matrix, as desired. ∎

5.62 תנאי הכרחי ומספיק ללכסינות

נניח ש־ V מממד סופי ו־ $T \in \mathcal{L}(V)$. אזי T לכסין אם־ורק־אם הפולינום המינימלי של T שווה ל־ $(z-\lambda_1)\cdots(z-\lambda_m)$ עבור רשימת מספרים שונים $\lambda_1, \ldots, \lambda_m \in \mathbf{F}$.

הוכחה ראשית נניח כי T לכסין. לפיכך קיים בסיס $v_1, \ldots, v_n$ של V המורכב מהווקטורים העצמיים של T. יהיו $\lambda_1, \ldots, \lambda_m$ הערכים העצמיים השונים של T. אזי לכל v_j, קיים λ_k שעבורו $(T-\lambda_k I)v_j = 0$. לפיכך

$$(T-\lambda_1 I)\cdots(T-\lambda_m I)v_j = 0,$$

מכאן נובע כי הפולינום המינימלי של T שווה ל־ $(z-\lambda_1)\cdots(z-\lambda_m)$.

להוכחת הגרירה בכיוון ההפוך, נניח כעת כי הפולינום המינימלי של T שווה ל־ $(z-\lambda_1)\cdots(z-\lambda_m)$ עבור רשימת מספרים שונים $\lambda_1, \ldots, \lambda_m \in \mathbf{F}$. לפיכך

$$(T-\lambda_1 I)\cdots(T-\lambda_m I) = 0. \tag{5.63}$$

נוכיח כי T לכסין באינדוקציה על m. נבדוק תחילה עבור $m = 1$. נקבל $T - \lambda_1 I = 0$, משמעות הדבר היא ש־ T כפולת סקלר של אופרטור הזהות, ומכאן ש־ T לכסין.

כעת נניח ש־ $m > 1$ והתוצאה הרצויה מתקיימת לכל ערך הקטן מ־ m. התת־מרחב $\operatorname{range}(T-\lambda_m I)$ שמור תחת T [זהו מקרה פרטי של 5.18 עבור $p(z) = z - \lambda_m$]. לפיכך T מצומצם ל־ $\operatorname{range}(T-\lambda_m I)$ הוא אופרטור מעל $\operatorname{range}(T-\lambda_m I)$.

אם $u \in \operatorname{range}(T-\lambda_m I)$, אזי $u = (T-\lambda_m I)v$ עבור $v \in V$ כלשהו, ומ־ 5.63 נובע

$$(T-\lambda_1 I)\cdots(T-\lambda_{m-1} I)u = (T-\lambda_1 I)\cdots(T-\lambda_m I)v = 0. \tag{5.64}$$

לכן $(z-\lambda_1)\cdots(z-\lambda_{m-1})$ הוא כפולת פולינום של הפולינום המינימלי של T המצומצם ל־ $\operatorname{range}(T-\lambda_m I)$ [לפי 5.29]. לפיכך, לפי הנחת האינדוקציה, קיים בסיס של $\operatorname{range}(T-\lambda_m I)$ המורכב מהווקטורים העצמיים של T.

נניח ש־ $u \in \operatorname{range}(T-\lambda_m I) \cap \operatorname{null}(T-\lambda_m I)$. אזי $Tu = \lambda_m u$. בנוסף מ־ 5.64 נובע

$$\begin{aligned} 0 &= (T-\lambda_1 I)\cdots(T-\lambda_{m-1} I)u \\ &= (\lambda_m-\lambda_1)\cdots(\lambda_m-\lambda_{m-1})u. \end{aligned}$$

מאחר ש־ $\lambda_1, \ldots, \lambda_m$ ערכים עצמיים שונים, מהמשוואה לעיל נובע כי $u = 0$. לכן $\operatorname{range}(T-\lambda_m I) \cap \operatorname{null}(T-\lambda_m I) = \{0\}$.

לפיכך $\operatorname{range}(T-\lambda_m I) + \operatorname{null}(T-\lambda_m I)$ הוא סכום ישר (לפי 1.46) שממדו $\dim V$ (לפי 3.94 ו־ 3.21). לכן $\operatorname{range}(T-\lambda_m I) \oplus \operatorname{null}(T-\lambda_m I) = V$. כל וקטור השונה מאפס ב־ $\operatorname{null}(T-\lambda_m I)$ הוא וקטור עצמי של T עם ערך עצמי λ_m. קודם לכן בהוכחה זו ראינו שקיים בסיס של $\operatorname{range}(T-\lambda_m I)$ המורכב מהווקטורים העצמיים של T. צירוף בסיס זה לבסיס של $\operatorname{null}(T-\lambda_m I)$ ייתן בסיס של V המורכב מהווקטורים העצמיים של T. המטריצה של T ביחס לבסיס זה היא מטריצה אלכסונית, כפי שרצינו. ■

אלגברה ליניארית בדרך הישר, מהדורה רביעית, מחבר: שלדון אקסלר
Linear Algebra Done Right, fourth edition, by Sheldon Axler

No formula exists for the zeros of polynomials of degree 5 or greater. However, the previous result can be used to determine whether an operator on a complex vector space is diagonalizable without even finding approximations of the zeros of the minimal polynomial—see Exercise 15.

The next result will be a key tool when we prove a result about the simultaneous diagonalization of two operators; see 5.76. Note how the use of a characterization of diagonalizable operators in terms of the minimal polynomial (see 5.62) leads to a short proof of the next result.

5.65 *restriction of diagonalizable operator to invariant subspace*

Suppose $T \in \mathcal{L}(V)$ is diagonalizable and U is a subspace of V that is invariant under T. Then $T|_U$ is a diagonalizable operator on U.

Proof Because the operator T is diagonalizable, the minimal polynomial of T equals $(z - \lambda_1) \cdots (z - \lambda_m)$ for some list of distinct numbers $\lambda_1, \ldots, \lambda_m \in \mathbf{F}$ (by 5.62). The minimal polynomial of T is a polynomial multiple of the minimal polynomial of $T|_U$ (by 5.31). Hence the minimal polynomial of $T|_U$ has the form required by 5.62, which shows that $T|_U$ is diagonalizable. ∎

Gershgorin Disk Theorem

5.66 definition: *Gershgorin disks*

Suppose $T \in \mathcal{L}(V)$ and $v_1, \ldots, v_n$ is a basis of V. Let A denote the matrix of T with respect to this basis. A *Gershgorin disk* of T with respect to the basis $v_1, \ldots, v_n$ is a set of the form

$$\left\{ z \in \mathbf{F} : |z - A_{j,j}| \le \sum_{\substack{k=1 \\ k \ne j}}^{n} |A_{j,k}| \right\},$$

where $j \in \{1, \ldots, n\}$.

Because there are n choices for j in the definition above, T has n Gershgorin disks. If $\mathbf{F} = \mathbf{C}$, then for each $j \in \{1, \ldots, n\}$, the corresponding Gershgorin disk is a closed disk in $\mathbf{C}$ centered at $A_{j,j}$, which is the jth entry on the diagonal of A. The radius of this closed disk is the sum of the absolute values of the entries in row j of A, excluding the diagonal entry. If $\mathbf{F} = \mathbf{R}$, then the Gershgorin disks are closed intervals in $\mathbf{R}$.

In the special case that the square matrix A above is a diagonal matrix, each Gershgorin disk consists of a single point that is a diagonal entry of A (and each eigenvalue of T is one of those points, as required by the next result). One consequence of our next result is that if the nondiagonal entries of A are small, then each eigenvalue of T is near a diagonal entry of A.

לא קיימת נוסחת שורשים לפולינומים ממעלה 5 או יותר. בכל זאת, נוכל להשתמש במסקנה הקודמת כדי לקבוע האם אופרטור מעל מרחב וקטורי מרוכב לכסין, זאת אפילו מבלי למצוא קירובים לשורשי הפולינום המינימלי—ראו תרגיל 15.

המסקנה הבאה תשמש ככלי מרכזי להוכחת מסקנה הנוגעת ללכסון שני אופרטורים בו זמנית; ראו 5.76. שימו לב כיצד השימוש באפיון של אופרטורים לכסינים במונחים של פולינום מינימלי (ראו 5.62) מוביל להוכחה קצרה של המסקנה הבאה.

5.65 צמצום אופרטור לכסין לתת־מרחב שמור

נניח ש־ $T \in \mathcal{L}(V)$ לכסין ו־ U תת־מרחב של V השמור תחת T. אזי $T|_U$ אופרטור לכסין מעל U.

הוכחה מאחר ש־ T לכסין, הפולינום המינימלי של T שווה ל־ $(z-\lambda_1)\cdots(z-\lambda_m)$ עבור רשימה של מספרים שונים כלשהם $\lambda_1, \dots, \lambda_m \in \mathbf{F}$ (לפי 5.62). הפולינום המינימלי של T הוא כפולת פולינום של הפולינום המינימלי של $T|_U$ (לפי 5.31). לכן הפולינום המינימלי של $T|_U$ הוא מהצורה הנדרשת על ידי 5.62, מה שמראה כי $T|_U$ לכסין. ■

משפט עיגולי גרשגורן

5.66 הגדרה: עיגולי גרשגורן

נניח ש־ $T \in \mathcal{L}(V)$ ו־ $v_1, \dots, v_n$ בסיס של V. נסמן ב־ A את המטריצה של T ביחס לבסיס זה. עיגולי גרשגורן של T ביחס לבסיס $v_1, \dots, v_n$ הם הקבוצות מהצורה

$$\left\{ z \in \mathbf{F} : |z - A_{j,j}| \le \sum_{\substack{k=1 \\ k \ne j}}^{n} |A_{j,k}| \right\},$$

כאשר $j \in \{1, \dots, n\}$.

מאחר שישנן n אפשרויות עבור j בהגדרה לעיל, ל־ T יש n עיגולי גרשגורן. אם $\mathbf{F} = \mathbf{C}$, אזי לכל $j \in \{1, \dots, n\}$, עיגול גרשגורן המתאים הוא עיגול סגור ב־ $\mathbf{C}$ שמרכזו $A_{j,j}$, שהוא האיבר ה־ j^{th} באלכסון של A. הרדיוס של עיגול סגור זה הוא סכום הערכים המוחלטים של איברי שורה j של A, מלבד איבר האלכסון. אם $\mathbf{F} = \mathbf{R}$, אזי עיגולי גרשגורן הם קטעים סגורים ב־ $\mathbf{R}$.

במקרה המיוחד שבו המטריצה הריבועית לעיל A היא מטריצה אלכסונית, כל עיגול גרשגורן הוא נקודה יחידה שהיא איבר אלכסון של A (וכל ערך עצמי של T הוא אחת מהנקודות הללו, כפי שנדרש במסקנה הבאה). אחת התוצאות מהמסקנה הבאה היא שאם האיברים שאינם איברי אלכסון של A קטנים, אזי כל ערך עצמי של T קרוב לאיבר אלכסון כלשהו של A.

5.67 *Gershgorin disk theorem*

Suppose $T \in \mathcal{L}(V)$ and $v_1, \dots, v_n$ is a basis of V. Then each eigenvalue of T is contained in some Gershgorin disk of T with respect to the basis $v_1, \dots, v_n$.

Proof Suppose $\lambda \in \mathbf{F}$ is an eigenvalue of T. Let $w \in V$ be a corresponding eigenvector. There exist $c_1, \dots, c_n \in \mathbf{F}$ such that

5.68 $$w = c_1 v_1 + \cdots + c_n v_n.$$

Let A denote the matrix of T with respect to the basis $v_1, \dots, v_n$. Applying T to both sides of the equation above gives

5.69 $$\begin{aligned} \lambda w &= \sum_{k=1}^{n} c_k T v_k \\ &= \sum_{k=1}^{n} c_k \sum_{j=1}^{n} A_{j,k} v_j \end{aligned}$$

5.70 $$= \sum_{j=1}^{n} \Big(\sum_{k=1}^{n} A_{j,k} c_k \Big) v_j.$$

Let $j \in \{1, \dots, n\}$ be such that

$$|c_j| = \max\{|c_1|, \dots, |c_n|\}.$$

Using 5.68, we see that the coefficient of v_j on the left side of 5.69 equals λc_j, which must equal the coefficient of v_j on the right side of 5.70. In other words,

$$\lambda c_j = \sum_{k=1}^{n} A_{j,k}\, c_k.$$

Subtract $A_{j,j}\, c_j$ from each side of the equation above and then divide both sides by c_j to get

$$\begin{aligned} |\lambda - A_{j,j}| &= \left| \sum_{\substack{k=1 \\ k \neq j}}^{n} A_{j,k} \frac{c_k}{c_j} \right| \\ &\leq \sum_{\substack{k=1 \\ k \neq j}}^{n} |A_{j,k}|. \end{aligned}$$

Thus λ is in the jth Gershgorin disk with respect to the basis $v_1, \dots, v_n$. ∎

Exercise 22 gives a nice application of the Gershgorin disk theorem.

The Gershgorin disk theorem is named for Semyon Aronovich Gershgorin, who published this result in 1931.

Exercise 23 states that the radius of each Gershgorin disk could be changed to the sum of the absolute values of corresponding column entries (instead of row entries), excluding the diagonal entry, and the theorem above would still hold.

5.67 משפט עיגולי גרשגורן

נניח ש־ $T \in \mathcal{L}(V)$ ו־ $v_1, \dots, v_n$ בסיס של V. אזי כל ערך עצמי של T מוכל בעיגול גרשגורן כלשהו של T ביחס לבסיס $v_1, \dots, v_n$.

הוכחה נניח ש־ $\lambda \in \mathbf{F}$ ערך עצמי של T. יהי $w \in V$ וקטור עצמי השייך ל־ λ. קיימים $c_1, \dots, c_n \in \mathbf{F}$ כך ש־

$$w = c_1 v_1 + \cdots + c_n v_n. \tag{5.68}$$

נסמן ב־ A את המטריצה של T ביחס לבסיס $v_1, \dots, v_n$. הפעלת T על שני אגפי המשוואה נותנת

$$\lambda w = \sum_{k=1}^{n} c_k T v_k \tag{5.69}$$

$$= \sum_{k=1}^{n} c_k \sum_{j=1}^{n} A_{j,k} v_j$$

$$= \sum_{j=1}^{n} \Big(\sum_{k=1}^{n} A_{j,k} c_k \Big) v_j. \tag{5.70}$$

יהי $j \in \{1, \dots, n\}$ כך ש־

$$|c_j| = \max\{|c_1|, \dots, |c_n|\}.$$

לפי 5.68, אנו רואים שהמקדם של v_j באגף השמאלי של 5.69 שווה ל־ λc_j, שהוא חייב להיות שווה למקדם של v_j באגף הימני של 5.70. במילים אחרות,

$$\lambda c_j = \sum_{k=1}^{n} A_{j,k}\, c_k.$$

נחסיר $A_{j,j}\, c_j$ משני אגפי המשוואה לעיל ואז נחלק את שני אגפי המשוואה ב־ c_j ונקבל

$$|\lambda - A_{j,j}| = \left| \sum_{\substack{k=1 \\ k \neq j}}^{n} A_{j,k} \frac{c_k}{c_j} \right|$$

$$\leq \sum_{\substack{k=1 \\ k \neq j}}^{n} |A_{j,k}|.$$

לפיכך λ נמצא בעיגול גרשגורן ה־ j^{th} ביחס לבסיס $v_1, \dots, v_n$. ■

משפט עיגולי גרשגורן קרוי על שמו של סמיון ארוביץ' גרשגורן, שפרסם מסקנה זו בשנת 1931.

בתרגיל 22 ישנו יישום נחמד של משפט עיגולי גרשגורן.

תרגיל 23 קובע כי ניתן לשנות את הגדרת הרדיוס של כל עיגול גרשגורן כך שיהיה סכום הערכים המוחלטים של איברי העמודה המתאימה (במקום איברי השורה), מלבד איבר האלכסון, והמשפט לעיל עדיין מתקיים.

Exercises 5D

1 Suppose V is a finite-dimensional complex vector space and $T \in \mathcal{L}(V)$.

(a) Prove that if $T^4 = I$, then T is diagonalizable.

(b) Prove that if $T^4 = T$, then T is diagonalizable.

(c) Give an example of an operator $T \in \mathcal{L}(\mathbf{C}^2)$ such that $T^4 = T^2$ and T is not diagonalizable.

2 Suppose $T \in \mathcal{L}(V)$ has a diagonal matrix A with respect to some basis of V. Prove that if $\lambda \in \mathbf{F}$, then λ appears on the diagonal of A precisely $\dim E(\lambda, T)$ times.

3 Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Prove that if the operator T is diagonalizable, then $V = \operatorname{null} T \oplus \operatorname{range} T$.

4 Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Prove that the following are equivalent.

(a) $V = \operatorname{null} T \oplus \operatorname{range} T$.

(b) $V = \operatorname{null} T + \operatorname{range} T$.

(c) $\operatorname{null} T \cap \operatorname{range} T = \{0\}$.

5 Suppose V is a finite-dimensional complex vector space and $T \in \mathcal{L}(V)$. Prove that T is diagonalizable if and only if

$$V = \operatorname{null}(T - \lambda I) \oplus \operatorname{range}(T - \lambda I)$$

for every $\lambda \in \mathbf{C}$.

6 Suppose $T \in \mathcal{L}(\mathbf{F}^5)$ and $\dim E(8, T) = 4$. Prove that $T - 2I$ or $T - 6I$ is invertible.

7 Suppose $T \in \mathcal{L}(V)$ is invertible. Prove that

$$E(\lambda, T) = E\Big(\tfrac{1}{\lambda}, T^{-1}\Big)$$

for every $\lambda \in \mathbf{F}$ with $\lambda \neq 0$.

8 Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Let $\lambda_1, \ldots, \lambda_m$ denote the distinct nonzero eigenvalues of T. Prove that

$$\dim E(\lambda_1, T) + \cdots + \dim E(\lambda_m, T) \leq \dim \operatorname{range} T.$$

9 Suppose $R, T \in \mathcal{L}(\mathbf{F}^3)$ each have 2, 6, 7 as eigenvalues. Prove that there exists an invertible operator $S \in \mathcal{L}(\mathbf{F}^3)$ such that $R = S^{-1}TS$.

10 Find $R, T \in \mathcal{L}(\mathbf{F}^4)$ such that R and T each have 2, 6, 7 as eigenvalues, R and T have no other eigenvalues, and there does not exist an invertible operator $S \in \mathcal{L}(\mathbf{F}^4)$ such that $R = S^{-1}TS$.

תרגילים ה9

1 נניח ש־ V מרחב וקטורי מרוכב מממד סופי ו־ $T \in \mathcal{L}(V)$.

א. הוכיחו שאם $T^4 = I$, אזי T לכסין.

ב. הוכיחו שאם $T^4 = T$, אזי T לכסין.

ג. הביאו דוגמה לאופרטור $T \in \mathcal{L}(\mathbf{C}^2)$ כך ש־ $T^4 = T^2$ ו־ T אינו לכסין.

2 נניח שיש ל־ $T \in \mathcal{L}(V)$ מטריצה אלכסונית A ביחס לבסיס כלשהו של V. הוכיחו שאם $\lambda \in \mathbf{F}$, אזי λ מופיע באלכסון של A בדיוק $\dim E(\lambda, T)$ פעמים.

3 נניח ש־ V מממד סופי ו־ $T \in \mathcal{L}(V)$. הוכיחו שאם האופרטור T לכסין, אזי $V = \operatorname{null} T \oplus \operatorname{range} T$.

4 נניח ש־ V מממד סופי ו־ $T \in \mathcal{L}(V)$. הוכיחו שהתנאים הבאים שקולים.

א. $V = \operatorname{null} T \oplus \operatorname{range} T$.

ב. $V = \operatorname{null} T + \operatorname{range} T$.

ג. $\operatorname{null} T \cap \operatorname{range} T = \{0\}$.

5 נניח ש־ V מרחב וקטורי מרוכב מממד סופי ו־ $T \in \mathcal{L}(V)$. הוכיחו ש־ T לכסין אם־ורק־אם

$$V = \operatorname{null}(T - \lambda I) \oplus \operatorname{range}(T - \lambda I)$$

לכל $\lambda \in \mathbf{C}$.

6 נניח ש־ $T \in \mathcal{L}(\mathbf{F}^5)$ ו־ $\dim E(8, T) = 4$. הוכיחו ש־ $T - 2I$ או $T - 6I$ הפיך.

7 נניח ש־ $T \in \mathcal{L}(V)$ הפיך. הוכיחו ש־

$$E(\lambda, T) = E\left(\tfrac{1}{\lambda}, T^{-1}\right)$$

לכל $\lambda \in \mathbf{F}$ עם $\lambda \neq 0$.

8 נניח ש־ V מממד סופי ו־ $T \in \mathcal{L}(V)$. נסמן ב־ $\lambda_1, \ldots, \lambda_m$ את הערכים העצמיים השונים שאינם אפס של T. הוכיחו ש־

$$\dim E(\lambda_1, T) + \cdots + \dim E(\lambda_m, T) \leq \dim \operatorname{range} T.$$

9 נניח ש־ $R, T \in \mathcal{L}(\mathbf{F}^3)$ שלכל אחד מהם 2, 6, 7 הם ערכים עצמיים. הוכיחו שקיים אופרטור הפיך $S \in \mathcal{L}(\mathbf{F}^3)$ כך ש־ $R = S^{-1}TS$.

10 מצאו $R, T \in \mathcal{L}(\mathbf{F}^4)$ כך שלכל אחד מ־ R ו־ T המספרים 2, 6, 7 הם ערכים עצמיים, אין ל־ R ול־ T ערכים עצמיים נוספים, ולא קיים אופרטור הפיך $S \in \mathcal{L}(\mathbf{F}^4)$ כך ש־ $R = S^{-1}TS$.

11 Find $T \in \mathcal{L}(\mathbf{C}^3)$ such that 6 and 7 are eigenvalues of T and such that T does not have a diagonal matrix with respect to any basis of $\mathbf{C}^3$.

12 Suppose $T \in \mathcal{L}(\mathbf{C}^3)$ is such that 6 and 7 are eigenvalues of T. Furthermore, suppose T does not have a diagonal matrix with respect to any basis of $\mathbf{C}^3$. Prove that there exists $(z_1, z_2, z_3) \in \mathbf{C}^3$ such that

$$T(z_1, z_2, z_3) = (6 + 8z_1, 7 + 8z_2, 13 + 8z_3).$$

13 Suppose A is a diagonal matrix with distinct entries on the diagonal and B is a matrix of the same size as A. Show that $AB = BA$ if and only if B is a diagonal matrix.

14 (a) Give an example of a finite-dimensional complex vector space and an operator T on that vector space such that T^2 is diagonalizable but T is not diagonalizable.

(b) Suppose $\mathbf{F} = \mathbf{C}$, k is a positive integer, and $T \in \mathcal{L}(V)$ is invertible. Prove that T is diagonalizable if and only if T^k is diagonalizable.

15 Suppose V is a finite-dimensional complex vector space, $T \in \mathcal{L}(V)$, and p is the minimal polynomial of T. Prove that the following are equivalent.

(a) T is diagonalizable.

(b) There does not exist $\lambda \in \mathbf{C}$ such that p is a polynomial multiple of $(z - \lambda)^2$.

(c) p and its derivative p' have no zeros in common.

(d) The greatest common divisor of p and p' is the constant polynomial 1.

*The **greatest common divisor** of p and p' is the monic polynomial q of largest degree such that p and p' are both polynomial multiples of q. The Euclidean algorithm for polynomials (look it up) can quickly determine the greatest common divisor of two polynomials, without requiring any information about the zeros of the polynomials. Thus the equivalence of (a) and (d) above shows that we can determine whether T is diagonalizable without knowing anything about the zeros of p.*

16 Suppose that $T \in \mathcal{L}(V)$ is diagonalizable. Let $\lambda_1, \ldots, \lambda_m$ denote the distinct eigenvalues of T. Prove that a subspace U of V is invariant under T if and only if there exist subspaces $U_1, \ldots, U_m$ of V such that $U_k \subseteq E(\lambda_k, T)$ for each k and $U = U_1 \oplus \cdots \oplus U_m$.

17 Suppose V is finite-dimensional. Prove that $\mathcal{L}(V)$ has a basis consisting of diagonalizable operators.

18 Suppose that $T \in \mathcal{L}(V)$ is diagonalizable and U is a subspace of V that is invariant under T. Prove that the quotient operator T/U is a diagonalizable operator on V/U.

The quotient operator T/U was defined in Exercise 38 in Section 5A.

11 מצאו $T \in \mathcal{L}(\mathbf{C}^3)$ כך ש־ 6 ו־ 7 ערכים עצמיים של T וכך שאין ל־ T מטריצה אלכסונית ביחס לאף בסיס של $\mathbf{C}^3$.

12 נניח ש־ $T \in \mathcal{L}(\mathbf{C}^3)$ כך ש־ 6 ו־ 7 ערכים עצמיים של T. בנוסף, נניח שאין ל־ T מטריצה אלכסונית ביחס לאף בסיס של $\mathbf{C}^3$. הוכיחו שקיים $(z_1, z_2, z_3) \in \mathbf{C}^3$ כך ש־

$$T(z_1, z_2, z_3) = (6 + 8z_1, 7 + 8z_2, 13 + 8z_3).$$

13 נניח ש־ A מטריצה אלכסונית עם איברים שונים באלכסון ו־ B מטריצה בגודל זהה ל־ A. הראו ש־ $AB = BA$ אם־ורק־אם B מטריצה אלכסונית.

14 א. הביאו דוגמה למרחב וקטורי מרוכב מממד סופי ואופרטור T מעל אותו מרחב וקטורי כך ש־ T^2 לכסין אבל T אינו לכסין.

ב. נניח ש־ $\mathbf{F} = \mathbf{C}$, k מספר שלם חיובי, ו־ $T \in \mathcal{L}(V)$ הפיך. הוכיחו ש־ T לכסין אם־ורק־אם T^k לכסין.

15 נניח ש־ V מרחב וקטורי מרוכב מממד סופי, $T \in \mathcal{L}(V)$, ו־ p הפולינום המינימלי של T. הוכיחו שהתנאים הבאים שקולים.

א. T לכסין.

ב. לא קיים $\lambda \in \mathbf{C}$ כך ש־ p הוא כפולת פולינום של $(z - \lambda)^2$.

ג. ל־ p ולנגזרתו p' אין שורשים משותפים.

ד. המחלק המשותף המרבי של p ו־ p' הוא הפולינום הקבוע 1.

> **המחלק המשותף המרבי** של p ו־ p' הוא הפולינום המתוקן q מהמעלה הגבוהה ביותר כך ש־ p ו־ p' שניהם כפולת פולינום של q. בעזרת האלגוריתם האוקלידי לפולינומים (חפשו מידע לגביו) נוכל למצוא בזריזות את המחלק המשותף המרבי של שני פולינומים, ללא צורך במידע על שורשי הפולינומים. לפיכך השקילות בין (א) ו־ (ד) לעיל מלמדת אותנו כי נוכל לקבוע האם T לכסין בלי שיהיה בידינו מידע הנוגע לשורשיו של p.

16 נניח ש־ $T \in \mathcal{L}(V)$ לכסין. יהיו $\lambda_1, \dots, \lambda_m$ הערכים העצמיים השונים של T. הוכיחו שתת־מרחב U של V שמור תחת T אם־ורק־אם קיימים תתי־מרחב $U_1, \dots, U_m$ של V כך ש־ $U_k \subseteq E(\lambda_k, T)$ לכל k ו־ $U = U_1 \oplus \cdots \oplus U_m$.

17 נניח ש־ V מממד סופי. הוכיחו שיש ל־ $\mathcal{L}(V)$ בסיס המורכב מאופרטורים לכסינים.

18 נניח ש־ $T \in \mathcal{L}(V)$ לכסין ו־ U תת־מרחב של V השמור תחת T. הוכיחו שאופרטור המנה T/U אופרטור לכסין מעל V/U.

אופרטור המנה T/U הוגדר בתרגיל 38 ביחידה ה1.

19 Prove or give a counterexample: If $T \in \mathcal{L}(V)$ and there exists a subspace U of V that is invariant under T such that $T|_U$ and T/U are both diagonalizable, then T is diagonalizable.

See Exercise 13 in Section 5C for an analogous statement about upper-triangular matrices.

20 Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Prove that T is diagonalizable if and only if the dual operator T' is diagonalizable.

21 The *Fibonacci sequence* $F_0, F_1, F_2, \ldots$ is defined by

$$F_0 = 0, \ F_1 = 1, \text{ and } F_n = F_{n-2} + F_{n-1} \text{ for } n \geq 2.$$

Define $T \in \mathcal{L}(\mathbf{R}^2)$ by $T(x, y) = (y, x + y)$.

(a) Show that $T^n(0, 1) = (F_n, F_{n+1})$ for each nonnegative integer n.
(b) Find the eigenvalues of T.
(c) Find a basis of $\mathbf{R}^2$ consisting of eigenvectors of T.
(d) Use the solution to (c) to compute $T^n(0, 1)$. Conclude that

$$F_n = \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right]$$

for each nonnegative integer n.

(e) Use (d) to conclude that if n is a nonnegative integer, then the Fibonacci number F_n is the integer that is closest to

$$\frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^n.$$

Each F_n is a nonnegative integer, even though the right side of the formula in (d) does not look like an integer. The number

$$\frac{1+\sqrt{5}}{2}$$

is called the ***golden ratio****.*

22 Suppose $T \in \mathcal{L}(V)$ and A is an n-by-n matrix that is the matrix of T with respect to some basis of V. Prove that if

$$|A_{j,j}| > \sum_{\substack{k=1 \\ k \neq j}}^{n} |A_{j,k}|$$

for each $j \in \{1, \ldots, n\}$, then T is invertible.

This exercise states that if the diagonal entries of the matrix of T are large compared to the nondiagonal entries, then T is invertible.

23 Suppose the definition of the Gershgorin disks is changed so that the radius of the kth disk is the sum of the absolute values of the entries in column (instead of row) k of A, excluding the diagonal entry. Show that the Gershgorin disk theorem (5.67) still holds with this changed definition.

19 הוכיחו או הביאו דוגמה נגדית: אם $T \in \mathcal{L}(V)$ וקיים תת־מרחב U של V השמור תחת T כך ש־ $T|_U$ ו־ T/U שניהם לכסינים, אזי T לכסין.

ראו תרגיל 13 ביחידה הקודמת לקביעה דומה עבור מטריצות משולשיות עליונות.

20 נניח ש־ V מממד סופי ו־ $T \in \mathcal{L}(V)$. הוכיחו ש־ T לכסין אם־ורק־אם האופרטור הדואלי T' לכסין.

21 סדרת פיבונאצ'י $F_0, F_1, F_2, \ldots$ מוגדרת על ידי

$$F_0 = 0, \ F_1 = 1, \quad \text{ו־} \quad F_n = F_{n-2} + F_{n-1}$$

עבור $n \geq 2$.

הגדירו $T \in \mathcal{L}(\mathbf{R}^2)$ על ידי $T(x, y) = (y, x + y)$.

א. הראו ש־ $T^n(0, 1) = (F_n, F_{n+1})$ לכל מספר שלם אי־שלילי n.

ב. מצאו את הערכים העצמיים של T.

ג. מצאו בסיס של $\mathbf{R}^2$ המורכב מווקטורים עצמיים של T.

ד. השתמשו בפתרון (ג) לחישוב $T^n(0, 1)$. הגיעו למסקנה כי

$$F_n = \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right]$$

לכל מספר שלם אי־שלילי n.

ה. השתמשו ב־ (ד) כדי להסיק שאם n מספר שלם אי־שלילי, אזי מספר פיבונאצ'י F_n הוא המספר השלם הקרוב ביותר ל־

$$\frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^n.$$

כל F_n הוא מספר שלם אי־שלילי, למרות שהאגף הימני של הנוסחה ב־ (ד) אינו נראה כמו מספר שלם. המספר

$$\frac{1+\sqrt{5}}{2}$$

נקרא **יחס הזהב**.

22 נניח ש־ $T \in \mathcal{L}(V)$ ו־ A מטריצה $n \times n$ שהיא המטריצה של T ביחס לבסיס כלשהו של V. הוכיחו שאם

$$|A_{j,j}| > \sum_{\substack{k=1 \\ k \neq j}}^{n} |A_{j,k}|$$

לכל $j \in \{1, \ldots, n\}$, אזי T הפיך.

תרגיל זה קובע שאם איברי אלכסון המטריצה של T גדולים ביחס לאיברים שאינם באלכסון, אזי T הפיך.

23 נניח שאנו משנים את הגדרת עיגולי גרשגורן כך שהרדיוס של העיגול ה־ k^{th} הוא סכום הערכים המוחלטים של איברי העמודה (במקום איברי השורה) k של A, מלבד איבר האלכסון. הראו שלמרות שינוי זה משפט עיגולי גרשגורן (5.67) עדיין מתקיים.

5E Commuting Operators

5.71 definition: *commute*

- Two operators S and T on the same vector space *commute* if $ST = TS$.
- Two square matrices A and B of the same size *commute* if $AB = BA$.

For example, if I is the identity operator on V and $\lambda \in \mathbf{F}$, then λI commutes with every operator on V.

As another example, if T is an operator then T^2 and T^3 commute. More generally, if $p, q \in \mathcal{P}(\mathbf{F})$, then $p(T)$ and $q(T)$ commute [see 5.17(b)].

5.72 example: *partial differentiation operators commute*

Suppose m is a nonnegative integer. Let $\mathcal{P}_m(\mathbf{C}^2, \mathbf{C})$ denote the complex vector space of polynomials (with coefficients in $\mathbf{C}$) in two variables and of degree at most m, with the usual operations of addition and scalar multiplication of $\mathbf{C}$-valued functions. Thus the elements of $\mathcal{P}_m(\mathbf{C}^2, \mathbf{C})$ are functions p from $\mathbf{C}^2$ to $\mathbf{C}$ of the form

5.73
$$p(w, z) = \sum_{j+k \le m} a_{j,k} w^j z^k,$$

where the indices j and k take on all nonnegative integer values such that $j + k \le m$, each $a_{j,k}$ is in $\mathbf{C}$, and $w^j z^k$ denotes the function on $\mathbf{C}^2$ defined by $(w, z) \mapsto w^j z^k$.

Define operators $D_w, D_z \in \mathcal{L}(\mathcal{P}_m(\mathbf{C}^2, \mathbf{C}))$ by

$$D_w p = \frac{\partial p}{\partial w} = \sum_{j+k \le m} j a_{j,k} w^{j-1} z^k \quad \text{and} \quad D_z p = \frac{\partial p}{\partial z} = \sum_{j+k \le m} k a_{j,k} w^j z^{k-1},$$

where p is as in 5.73. The operators D_w and D_z are called partial differentiation operators because each of these operators differentiates with respect to one of the variables while pretending that the other variable is a constant.

The operators D_w and D_z commute because if p is as in 5.73, then

$$(D_w D_z) p = \sum_{j+k \le m} jk a_{j,k} w^{j-1} z^{k-1} = (D_z D_w) p.$$

The equation $D_w D_z = D_z D_w$ on $\mathcal{P}_m(\mathbf{C}^2, \mathbf{C})$ illustrates a more general result that the order of partial differentiation does not matter for nice functions.

Commuting matrices are unusual. For example, there are 214,358,881 ordered pairs of 2-by-2 matrices all of whose entries are integers in the interval $[-5, 5]$. Only about 0.3% of these ordered pairs of matrices commute.

All 214,358,881 (which equals 11^8) ordered pairs of the 2-by-2 matrices under consideration were checked by a computer to discover that only 674,609 of these ordered pairs of matrices commute.

5ה אופרטורים מתחלפים

5.71 הגדרה: חילוף

- שני אופרטורים S ו־ T מעל אותו מרחב וקטורי מתחלפים אם $ST = TS$.
- שתי מטריצות ריבועיות A ו־ B מאותו הגודל מתחלפות אם $AB = BA$.

לדוגמה, אם I אופרטור הזהות מעל V ו־ $\lambda \in \mathbf{F}$, אזי λI מתחלף עם כל אופרטור מעל V.

כדוגמה נוספת, אם T אופרטור אזי T^2 ו־ T^3 מתחלפים. בהכללה רחבה יותר, אם $p, q \in \mathcal{P}(\mathbf{F})$, אזי $p(T)$ ו־ $q(T)$ מתחלפים [ראו 5.17(ב)].

5.72 דוגמה: אופרטורי נגזרת חלקית מתחלפים

נניח ש־ m מספר שלם אי־שלילי. נסמן ב־ $\mathcal{P}_m(\mathbf{C}^2, \mathbf{C})$ את המרחב הווקטורי המרוכב של פולינומים (עם מקדמים ב־ $\mathbf{C}$) בשני משתנים שמעלתם m לכל היותר, יחד עם הפעולות הרגילות של חיבור וכפל בסקלר של פונקציות מרוכבות. לפיכך, האיברים ב־ $\mathcal{P}_m(\mathbf{C}^2, \mathbf{C})$ הם פונקציות p מ־ $\mathbf{C}^2$ אל $\mathbf{C}$ מהצורה

$$p(w, z) = \sum_{j+k \le m} a_{j,k} w^j z^k, \tag{5.73}$$

כאשר האינדקסים j ו־ k מקבלים את כל הערכים השלמים האי־שליליים כך ש־ $j + k \le m$, כל $a_{j,k}$ הוא ב־ $\mathbf{C}$, ו־ $w^j z^k$ מסמנת את הפונקציה מעל $\mathbf{C}^2$ המוגדרת על ידי $(w, z) \mapsto w^j z^k$.

נגדיר אופרטורים $D_w, D_z \in \mathcal{L}(\mathcal{P}_m(\mathbf{C}^2, \mathbf{C}))$ באופן הבא

$$D_w p = \frac{\partial p}{\partial w} = \sum_{j+k \le m} j a_{j,k} w^{j-1} z^k \quad \text{וגם} \quad D_z p = \frac{\partial p}{\partial z} = \sum_{j+k \le m} k a_{j,k} w^j z^{k-1},$$

כאשר p הוא כמוגדר ב־ 5.73. האופרטורים D_w ו־ D_z נקראים אופרטורי נגזרת חלקית מאחר שכל אחד מהאופרטורים הללו גוזר ביחס לאחד המשתנים תוך התייחסות למשתנה השני כקבוע.

האופרטורים D_w ו־ D_z מתחלפים שכן אם p הוא כמוגדר ב־ 5.73, אזי

$$(D_w D_z)p = \sum_{j+k \le m} jk a_{j,k} w^{j-1} z^{k-1} = (D_z D_w)p.$$

המשוואה $D_w D_z = D_z D_w$ מעל $\mathcal{P}_m(\mathbf{C}^2, \mathbf{C})$ ממחישה מסקנה כללית יותר שסדר הפעולות בגזירה חלקית אינו משפיע על התוצאה עבור פונקציות נאותות.

מטריצות מתחלפות הן תופעה חריגה. להמחשה, ישנם 214,358,881 זוגות סדורים של מטריצות 2×2 שכל איבריהן מספרים שלמים בקטע $[-5, 5]$. רק כ־0.3% מהזוגות הסדורים של המטריצות הללו מתחלפים.

כל 214,358,881 (מספר השווה ל־ 11^8) הזוגות הסדורים של המטריצות 2×2 בטווח האמור נבדקו באמצעות מחשב והתוצאה הראתה כי רק 674,609 מהזוגות הסדורים הללו מתחלפים.

The next result shows that two operators commute if and only if their matrices (with respect to the same basis) commute.

5.74 *commuting operators correspond to commuting matrices*

Suppose $S, T \in \mathcal{L}(V)$ and $v_1, \dots, v_n$ is a basis of V. Then S and T commute if and only if $\mathcal{M}(S, (v_1, \dots, v_n))$ and $\mathcal{M}(T, (v_1, \dots, v_n))$ commute.

Proof We have

$$\begin{aligned} S \text{ and } T \text{ commute} &\iff ST = TS \\ &\iff \mathcal{M}(ST) = \mathcal{M}(TS) \\ &\iff \mathcal{M}(S)\mathcal{M}(T) = \mathcal{M}(T)\mathcal{M}(S) \\ &\iff \mathcal{M}(S) \text{ and } \mathcal{M}(T) \text{ commute,} \end{aligned}$$

as desired. ∎

The next result shows that if two operators commute, then every eigenspace for one operator is invariant under the other operator. This result, which we will use several times, is one of the main reasons why a pair of commuting operators behaves better than a pair of operators that does not commute.

5.75 *eigenspace is invariant under commuting operator*

Suppose $S, T \in \mathcal{L}(V)$ commute and $\lambda \in \mathbf{F}$. Then $E(\lambda, S)$ is invariant under T.

Proof Suppose $v \in E(\lambda, S)$. Then

$$S(Tv) = (ST)v = (TS)v = T(Sv) = T(\lambda v) = \lambda Tv.$$

The equation above shows that $Tv \in E(\lambda, S)$. Thus $E(\lambda, S)$ is invariant under T. ∎

Suppose we have two operators, each of which is diagonalizable. If we want to do computations involving both operators (for example, involving their sum), then we want the two operators to be diagonalizable by the same basis, which according to the next result is possible when the two operators commute.

5.76 *simultaneous diagonalizability $\iff$ commutativity*

Two diagonalizable operators on the same vector space have diagonal matrices with respect to the same basis if and only if the two operators commute.

Proof First suppose $S, T \in \mathcal{L}(V)$ have diagonal matrices with respect to the same basis. The product of two diagonal matrices of the same size is the diagonal matrix obtained by multiplying the corresponding elements of the two diagonals. Thus any two diagonal matrices of the same size commute. Thus S and T commute, by 5.74.

המסקנה הבאה מראה ששני אופרטורים
מתחלפים אם־ורק־אם המטריצות שלהם (ביחס לאותו הבסיס) מתחלפות.

5.74 אופרטורים מתחלפים מותאמים למטריצות מתחלפות

נניח ש־ $S, T \in \mathcal{L}(V)$ ו־ $v_1, \dots, v_n$ בסיס של V. אזי S ו־ T מתחלפים אם־ורק־אם $\mathcal{M}(S, (v_1, \dots, v_n))$ ו־ $\mathcal{M}(T, (v_1, \dots, v_n))$ מתחלפות.

הוכחה מתקיים

$$\begin{aligned} S \text{ ו־ } T \text{ מתחלפים} &\iff ST = TS \\ &\iff \mathcal{M}(ST) = \mathcal{M}(TS) \\ &\iff \mathcal{M}(S)\mathcal{M}(T) = \mathcal{M}(T)\mathcal{M}(S) \\ &\iff \mathcal{M}(S) \text{ ו־ } \mathcal{M}(T) \text{ מתחלפות,} \end{aligned}$$

כפי שרצינו. □

המסקנה הבאה מראה שאם שני אופרטורים מתחלפים, אזי כל מרחב עצמי של אופרטור אחד שמור תחת האופרטור השני. מסקנה זו, שבה נשתמש מספר פעמים, היא אחת הסיבות המרכזיות לכך שזוג אופרטורים מתחלפים נוח יותר לעבודה מאשר זוג אופרטורים שאינו מתחלף.

5.75 מרחב עצמי שמור תחת אופרטור מתחלף

נניח ש־ $S, T \in \mathcal{L}(V)$ מתחלפים ו־ $\lambda \in \mathbf{F}$. אזי $E(\lambda, S)$ שמור תחת T.

הוכחה נניח ש־ $v \in E(\lambda, S)$. אזי

$$S(Tv) = (ST)v = (TS)v = T(Sv) = T(\lambda v) = \lambda Tv.$$

המשוואה לעיל מראה ש־ $Tv \in E(\lambda, S)$. לפיכך $E(\lambda, S)$ שמור תחת T. □

נניח שישנם שני אופרטורים, שכל אחד מהם לכסין. אם ברצוננו לבצע חישובים המערבים את שני האופרטורים (לדוגמה, חישוב המערב את סכומם), אזי נרצה ששני האופרטורים יהיו לכסינים ביחס לאותו הבסיס, מה שלפי המסקנה הבאה אפשרי כאשר שני האופרטורים מתחלפים.

5.76 לכסינות בו־זמנית $\iff$ קומוטטיביות

לשני אופרטורים לכסינים מעל אותו מרחב וקטורי יש מטריצות אלכסוניות ביחס לאותו בסיס אם־ורק־אם שני האופרטורים מתחלפים.

הוכחה ראשית נניח כי ל־ $S, T \in \mathcal{L}(V)$ יש מטריצות אלכסוניות ביחס לאותו בסיס. מכפלת שתי מטריצות אלכסוניות באותו הגודל היא מטריצה אלכסונית הנוצרת כתוצאה מהכפלת איברי האלכסונים של המטריצות בהתאמה. לפיכך כל שתי מטריצות אלכסוניות מאותו הגודל מתחלפות. לפיכך S ו־ T מתחלפים, לפי 5.74.

To prove the implication in the other direction, now suppose that $S, T \in \mathcal{L}(V)$ are diagonalizable operators that commute. Let $\lambda_1, \dots, \lambda_m$ denote the distinct eigenvalues of S. Because S is diagonalizable, 5.55(c) shows that

5.77 $$V = E(\lambda_1, S) \oplus \cdots \oplus E(\lambda_m, S).$$

For each $k = 1, \dots, m$, the subspace $E(\lambda_k, S)$ is invariant under T (by 5.75). Because T is diagonalizable, 5.65 implies that $T|_{E(\lambda_k, S)}$ is diagonalizable for each k. Hence for each $k = 1, \dots, m$, there is a basis of $E(\lambda_k, S)$ consisting of eigenvectors of T. Putting these bases together gives a basis of V (because of 5.77), with each vector in this basis being an eigenvector of both S and T. Thus S and T both have diagonal matrices with respect to this basis, as desired. ∎

See Exercise 2 for an extension of the result above to more than two operators.

Suppose V is a finite-dimensional nonzero complex vector space. Then every operator on V has an eigenvector (see 5.19). The next result shows that if two operators on V commute, then there is a vector in V that is an eigenvector for both operators (but the two commuting operators might not have a common eigenvalue). For an extension of the next result to more than two operators, see Exercise 9(a).

5.78 *common eigenvector for commuting operators*

Every pair of commuting operators on a finite-dimensional nonzero complex vector space has a common eigenvector.

Proof Suppose V is a finite-dimensional nonzero complex vector space and $S, T \in \mathcal{L}(V)$ commute. Let λ be an eigenvalue of S (5.19 tells us that S does indeed have an eigenvalue). Thus $E(\lambda, S) \neq \{0\}$. Also, $E(\lambda, S)$ is invariant under T (by 5.75).

Thus $T|_{E(\lambda, S)}$ has an eigenvector (again using 5.19), which is an eigenvector for both S and T, completing the proof. ∎

5.79 example: *common eigenvector for partial differentiation operators*

Let $\mathcal{P}_m(\mathbf{C}^2, \mathbf{C})$ be as in Example 5.72 and let $D_w, D_z \in \mathcal{L}(\mathcal{P}_m(\mathbf{C}^2, \mathbf{C}))$ be the commuting partial differentiation operators in that example. As you can verify, 0 is the only eigenvalue of each of these operators. Also

$$E(0, D_w) = \left\{ \sum_{k=0}^{m} a_k z^k : a_0, \dots, a_m \in \mathbf{C} \right\},$$

$$E(0, D_z) = \left\{ \sum_{j=0}^{m} c_j w^j : c_0, \dots, c_m \in \mathbf{C} \right\}.$$

The intersection of these two eigenspaces is the set of common eigenvectors of the two operators. Because $E(0, D_w) \cap E(0, D_z)$ is the set of constant functions, we see that D_w and D_z indeed have a common eigenvector, as promised by 5.78.

להוכחה בכיוון השני, נניח כעת ש־ $S, T \in \mathcal{L}(V)$ אופרטורים לכסינים מתחלפים. נסמן ב־ $\lambda_1, \dots, \lambda_m$ את הערכים העצמיים השונים של S. מאחר ש־ S לכסין, 5.55(ג) מראה ש־

$$V = E(\lambda_1, S) \oplus \cdots \oplus E(\lambda_m, S). \tag{5.77}$$

לכל $k = 1, \dots, m$, התת־מרחב $E(\lambda_k, S)$ שמור תחת T (לפי 5.75). מאחר ש־ T לכסין, 5.65 גורר ש־ $T|_{E(\lambda_k, S)}$ לכסין לכל k. לכן לכל $k = 1, \dots, m$, קיים בסיס של $E(\lambda_k, S)$ המורכב מהווקטורים העצמיים של T. איסוף כל הבסיסים הללו ייתן לנו בסיס של V (לפי 5.77), כאשר כל וקטור בבסיס זה הוא וקטור עצמי של שני האופרטורים S ו־ T. לפיכך ל־ S ו־ T יש מטריצות אלכסוניות ביחס לבסיס זה, כפי שרצינו. ■

ראו תרגיל 2 להרחבת מסקנה זו מעבר לשני אופרטורים.

נניח ש־ V מרחב וקטורי מרוכב מממד סופי השונה מאפס. אזי לכל אופרטור מעל V יש וקטור עצמי (ראו 5.19). המסקנה הבאה מראה כי אם שני אופרטורים מעל V מתחלפים, אזי קיים וקטור ב־ V שהוא וקטור עצמי המשותף לשני האופרטורים (אך ייתכן כי לשני האופרטורים המתחלפים לא יהיה ערך עצמי משותף). להרחבת המסקנה הבאה מעבר לשני אופרטורים, ראו תרגיל 9(א).

5.78 וקטור עצמי משותף לאופרטורים מתחלפים

לכל זוג אופרטורים מתחלפים מעל מרחב וקטורי מרוכב מממד סופי השונה מאפס יש וקטור עצמי משותף.

הוכחה נניח ש־ V מרחב וקטורי מרוכב מממד סופי השונה מאפס ו־ $S, T \in \mathcal{L}(V)$ מתחלפים. יהי λ ערך עצמי של S (לפי 5.19 ל־ S בהחלט יש ערך עצמי). לפיכך $E(\lambda, S) \neq \{0\}$. בנוסף, $E(\lambda, S)$ שמור תחת T (לפי 5.75).

לפיכך, ל־ $T|_{E(\lambda, S)}$ יש וקטור עצמי (שימוש נוסף ב־ 5.19), שהוא וקטור עצמי של שני האופרטורים S ו־ T, ובכך השלמנו את ההוכחה. ■

5.79 דוגמה: וקטור עצמי משותף לאופרטורי נגזרת חלקית

יהי $\mathcal{P}_m(\mathbf{C}^2, \mathbf{C})$ כבדוגמה 5.72 ויהיו $D_w, D_z \in \mathcal{L}(\mathcal{P}_m(\mathbf{C}^2, \mathbf{C}))$ האופרטורים של נגזרות חלקיות המתחלפים בדוגמה ההיא. כפי שתוכלו לוודא, 0 הוא הערך העצמי היחיד של כל אחד מהאופרטורים הללו. בנוסף

$$E(0, D_w) = \left\{ \sum_{k=0}^{m} a_k z^k : a_0, \dots, a_m \in \mathbf{C} \right\},$$

$$E(0, D_z) = \left\{ \sum_{j=0}^{m} c_j w^j : c_0, \dots, c_m \in \mathbf{C} \right\}.$$

החיתוך של שני מרחבים עצמיים אלו הוא קבוצת הווקטורים העצמיים המשותפים לשני האופרטורים. מאחר ש־ $E(0, D_w) \cap E(0, D_z)$ הוא קבוצת הפונקציות הקבועות, אנו רואים כי ל־ D_w ו־ D_z יש אכן וקטור עצמי משותף, כמובטח על ידי 5.78.

The next result extends 5.47 (the existence of a basis that gives an upper-triangular matrix) to two commuting operators.

5.80 *commuting operators are simultaneously upper triangularizable*

Suppose V is a finite-dimensional complex vector space and S, T are commuting operators on V. Then there is a basis of V with respect to which both S and T have upper-triangular matrices.

Proof Let $n = \dim V$. We will use induction on n. The desired result holds if $n = 1$ because all 1-by-1 matrices are upper triangular. Now suppose $n > 1$ and the desired result holds for all complex vector spaces whose dimension is $n - 1$.

Let v_1 be any common eigenvector of S and T (using 5.78). Hence $Sv_1 \in \operatorname{span}(v_1)$ and $Tv_1 \in \operatorname{span}(v_1)$. Let W be a subspace of V such that

$$V = \operatorname{span}(v_1) \oplus W;$$

see 2.33 for the existence of W. Define a linear map $P \colon V \to W$ by

$$P(av_1 + w) = w$$

for each $a \in \mathbf{C}$ and each $w \in W$. Define $\hat{S}, \hat{T} \in \mathcal{L}(W)$ by

$$\hat{S}w = P(Sw) \quad \text{and} \quad \hat{T}w = P(Tw)$$

for each $w \in W$. To apply our induction hypothesis to $\hat{S}$ and $\hat{T}$, we must first show that these two operators on W commute. To do this, suppose $w \in W$. Then there exists $a \in \mathbf{C}$ such that

$$(\hat{S}\hat{T})w = \hat{S}(P(Tw)) = \hat{S}(Tw - av_1) = P(S(Tw - av_1)) = P((ST)w),$$

where the last equality holds because v_1 is an eigenvector of S and $Pv_1 = 0$. Similarly,

$$(\hat{T}\hat{S})w = P((TS)w).$$

Because the operators S and T commute, the last two displayed equations show that $(\hat{S}\hat{T})w = (\hat{T}\hat{S})w$. Hence $\hat{S}$ and $\hat{T}$ commute.

Thus we can use our induction hypothesis to state that there exists a basis $v_2, \dots, v_n$ of W such that $\hat{S}$ and $\hat{T}$ both have upper-triangular matrices with respect to this basis. The list $v_1, \dots, v_n$ is a basis of V.

If $k \in \{2, \dots, n\}$, then there exist $a_k, b_k \in \mathbf{C}$ such that

$$Sv_k = a_k v_1 + \hat{S}v_k \quad \text{and} \quad Tv_k = b_k v_1 + \hat{T}v_k.$$

Because $\hat{S}$ and $\hat{T}$ have upper-triangular matrices with respect to $v_2, \dots, v_n$, we know that $\hat{S}v_k \in \operatorname{span}(v_2, \dots, v_k)$ and $\hat{T}v_k \in \operatorname{span}(v_2, \dots, v_k)$. Hence the equations above imply that

$$Sv_k \in \operatorname{span}(v_1, \dots, v_k) \quad \text{and} \quad Tv_k \in \operatorname{span}(v_1, \dots, v_k).$$

Thus S and T have upper-triangular matrices with respect to $v_1, \dots, v_n$, as desired. ∎

Exercise 9(b) extends the result above to more than two operators.

המסקנה הבאה מרחיבה את 5.47 (קיום בסיס היוצר מטריצה משולשית עליונה) לשני אופרטורים מתחלפים.

5.80 לאופרטורים מתחלפים יש מטריצות משולשיות עליונות בו־זמנית

נניח ש־ V מרחב וקטורי מרוכב מממד סופי ו־ S, T אופרטורים מתחלפים מעל V. אזי קיים בסיס של V אשר ביחס אליו לשני האופרטורים S ו־ T יש מטריצות משולשיות עליונות.

הוכחה יהי $n = \dim V$. נוכיח באמצעות אינדוקציה על n. אם $n = 1$ התוצאה הרצויה מתקיימת, שכן כל המטריצות 1×1 משולשיות עליונות. כעת נניח ש־ $n > 1$ והתוצאה הרצויה מתקיימת עבור כל המרחבים הווקטוריים המרוכבים מממד $n - 1$.

יהי v_1 וקטור עצמי כלשהו המשותף ל־ S ו־ T (לפי 5.78). לכן $Sv_1 \in \operatorname{span}(v_1)$ ו־ $Tv_1 \in \operatorname{span}(v_1)$. יהי W תת־מרחב של V כך ש־

$$V = \operatorname{span}(v_1) \oplus W;$$

ראו 2.33 לקיומו של W. נגדיר העתקה ליניארית $P : V \to W$ על ידי

$$P(av_1 + w) = w$$

לכל $a \in \mathbf{C}$ ולכל $w \in W$. נגדיר $\hat{S}, \hat{T} \in \mathcal{L}(W)$ על ידי

$$\hat{S}w = P(Sw) \quad \text{וגם} \quad \hat{T}w = P(Tw)$$

לכל $w \in W$. להחלת הנחת האינדוקציה שלנו על $\hat{S}$ ו־ $\hat{T}$, עלינו להראות תחילה כי שני אופרטורים אלו מעל W מתחלפים. לצורך כך, נניח ש־ $w \in W$. אזי קיים $a \in \mathbf{C}$ כך ש־

$$(\hat{S}\hat{T})w = \hat{S}(P(Tw)) = \hat{S}(Tw - av_1) = P(S(Tw - av_1)) = P((ST)w),$$

כאשר השוויון האחרון מתקיים מאחר ש־ v_1 הוא וקטור עצמי של S ו־ $Pv_1 = 0$. באופן דומה,

$$(\hat{T}\hat{S})w = P((TS)w).$$

מאחר שהאופרטורים S ו־ T מתחלפים, משתי המשוואות האחרונות אנו רואים כי $(\hat{S}\hat{T})w = (\hat{T}\hat{S})w$. לכן $\hat{S}$ ו־ $\hat{T}$ מתחלפים.

לפיכך אנו יכולים להשתמש בהנחת האינדוקציה כדי לקבוע כי קיים בסיס $v_2, \dots, v_n$ של W כך שיש ל־ $\hat{S}$ ו־ $\hat{T}$ מטריצות משולשיות עליונות ביחס לבסיס זה. הרשימה $v_1, \dots, v_n$ בסיס של V.

אם $k \in \{2, \dots, n\}$, אזי קיימים $a_k, b_k \in \mathbf{C}$ כך ש־

$$Sv_k = a_k v_1 + \hat{S}v_k \quad \text{וגם} \quad Tv_k = b_k v_1 + \hat{T}v_k.$$

מאחר שיש ל־ $\hat{S}$ ו־ $\hat{T}$ מטריצות משולשיות עליונות ביחס לבסיס $v_2, \dots, v_n$, אנו יודעים כי $\hat{S}v_k \in \operatorname{span}(v_2, \dots, v_k)$ ו־ $\hat{T}v_k \in \operatorname{span}(v_2, \dots, v_k)$. לכן מהמשוואות לעיל נובע כי

$$Sv_k \in \operatorname{span}(v_1, \dots, v_k) \quad \text{וגם} \quad Tv_k \in \operatorname{span}(v_1, \dots, v_k).$$

לפיכך יש ל־ S ו־ T מטריצות משולשיות עליונות ביחס ל־ $v_1, \dots, v_n$, כפי שרצינו. ■

תרגיל 9(ב) מרחיב מסקנה זו מעבר לשני אופרטורים.

In general, it is not possible to determine the eigenvalues of the sum or product of two operators from the eigenvalues of the two operators. However, the next result shows that something nice happens when the two operators commute.

5.81 *eigenvalues of sum and product of commuting operators*

Suppose V is a finite-dimensional complex vector space and S, T are commuting operators on V. Then

- every eigenvalue of $S + T$ is an eigenvalue of S plus an eigenvalue of T,
- every eigenvalue of ST is an eigenvalue of S times an eigenvalue of T.

Proof There is a basis of V with respect to which both S and T have upper-triangular matrices (by 5.80). With respect to that basis,

$$\mathcal{M}(S + T) = \mathcal{M}(S) + \mathcal{M}(T) \quad \text{and} \quad \mathcal{M}(ST) = \mathcal{M}(S)\mathcal{M}(T),$$

as stated in 3.35 and 3.43.

The definition of matrix addition shows that each entry on the diagonal of $\mathcal{M}(S + T)$ equals the sum of the corresponding entries on the diagonals of $\mathcal{M}(S)$ and $\mathcal{M}(T)$. Similarly, because $\mathcal{M}(S)$ and $\mathcal{M}(T)$ are upper-triangular matrices, the definition of matrix multiplication shows that each entry on the diagonal of $\mathcal{M}(ST)$ equals the product of the corresponding entries on the diagonals of $\mathcal{M}(S)$ and $\mathcal{M}(T)$. Furthermore, $\mathcal{M}(S + T)$ and $\mathcal{M}(ST)$ are upper-triangular matrices (see Exercise 2 in Section 5C).

Every entry on the diagonal of $\mathcal{M}(S)$ is an eigenvalue of S, and every entry on the diagonal of $\mathcal{M}(T)$ is an eigenvalue of T (by 5.41). Every eigenvalue of $S + T$ is on the diagonal of $\mathcal{M}(S + T)$, and every eigenvalue of ST is on the diagonal of $\mathcal{M}(ST)$ (these assertions follow from 5.41). Putting all this together, we conclude that every eigenvalue of $S + T$ is an eigenvalue of S plus an eigenvalue of T, and every eigenvalue of ST is an eigenvalue of S times an eigenvalue of T. ∎

Exercises 5E

1 Give an example of two commuting operators S, T on $\mathbf{F}^4$ such that there is a subspace of $\mathbf{F}^4$ that is invariant under S but not under T and there is a subspace of $\mathbf{F}^4$ that is invariant under T but not under S.

2 Suppose $\mathcal{E}$ is a subset of $\mathcal{L}(V)$ and every element of $\mathcal{E}$ is diagonalizable. Prove that there exists a basis of V with respect to which every element of $\mathcal{E}$ has a diagonal matrix if and only if every pair of elements of $\mathcal{E}$ commutes.

This exercise extends 5.76, which considers the case in which $\mathcal{E}$ contains only two elements. For this exercise, $\mathcal{E}$ may contain any number of elements, and $\mathcal{E}$ may even be an infinite set.

באופן כללי, לא ניתן לקבוע את הערכים העצמיים של סכום או מכפלת שני אופרטורים באמצעות הערכים העצמיים של שני האופרטורים. בכל זאת, המסקנה להלן מראה שקורה משהו נחמד כאשר שני אופרטורים מתחלפים.

5.81 ערכים עצמיים של סכום ומכפלת אופרטורים מתחלפים

נניח ש־ V מרחב וקטורי מרוכב מממד סופי ו־ S, T אופרטורים מתחלפים מעל V. אזי

- כל ערך עצמי של $S + T$ הוא ערך עצמי של S ועוד ערך עצמי של T,
- כל ערך עצמי של ST הוא ערך עצמי של S כפול ערך עצמי של T.

הוכחה קיים בסיס של V אשר ביחס אליו לשני האופרטורים S ו־ T יש מטריצות משולשיות עליונות (לפי 5.80). ביחס לבסיס זה,

$$\mathcal{M}(ST) = \mathcal{M}(S)\mathcal{M}(T), \quad \text{וגם} \quad \mathcal{M}(S + T) = \mathcal{M}(S) + \mathcal{M}(T)$$

כפי שנקבע ב־ 3.35 ו־ 3.43.

מהגדרת פעולת החיבור בין מטריצות אנו רואים שכל איבר באלכסון $\mathcal{M}(S + T)$ שווה לסכום האיברים המתאימים באלכסוני $\mathcal{M}(S)$ ו־ $\mathcal{M}(T)$. באופן דומה, מאחר ש־ $\mathcal{M}(S)$ ו־ $\mathcal{M}(T)$ מטריצות משולשיות עליונות, מהגדרת פעולת כפל מטריצות אנו רואים שכל איבר באלכסון $\mathcal{M}(ST)$ שווה למכפלת האיברים המתאימים באלכסוני $\mathcal{M}(S)$ ו־ $\mathcal{M}(T)$. יתר על כן, $\mathcal{M}(S + T)$ ו־ $\mathcal{M}(ST)$ מטריצות משולשיות עליונות (ראו תרגיל 2 ביחידה ה3).

כל איבר באלכסונה של $\mathcal{M}(S)$ הוא ערך עצמי של S, וכל איבר באלכסונה של $\mathcal{M}(T)$ הוא ערך עצמי של T (לפי 5.41). כל ערך עצמי של $S + T$ הוא על האלכסון של $\mathcal{M}(S + T)$, וכל ערך עצמי של ST הוא על האלכסון של $\mathcal{M}(ST)$ (קביעות אלו נובעות מ־ 5.41). מחיבור כל הנתונים הללו, אנו מסיקים כי כל ערך עצמי של $S + T$ הוא ערך עצמי של S ועוד ערך עצמי של T, וכל ערך עצמי של ST הוא ערך עצמי של S כפול ערך עצמי של T.

תרגילים ה5

1 הביאו דוגמה של שני אופרטורים מתחלפים S, T מעל $\mathbf{F}^4$ כך שקיים תת־מרחב של $\mathbf{F}^4$ השמור תחת S אבל לא תחת T וקיים תת־מרחב של $\mathbf{F}^4$ השמור תחת T אבל לא תחת S.

2 נניח ש־ $\mathcal{E}$ תת־קבוצה של $\mathcal{L}(V)$ וכל איבר של $\mathcal{E}$ לכסין. הוכיחו שקיים בסיס של V אשר ביחס אליו לכל איבר של $\mathcal{E}$ יש מטריצה אלכסונית אם־ורק־אם כל זוג איברים של $\mathcal{E}$ מתחלף.

תרגיל זה מרחיב את 5.76, העוסק במקרה שבו $\mathcal{E}$ מכיל רק שני איברים. לצורך תרגיל זה, $\mathcal{E}$ עשוי להכיל כל מספר של איברים, וייתכן אף ש־ $\mathcal{E}$ קבוצה אינסופית.

3 Suppose $S, T \in \mathcal{L}(V)$ are such that $ST = TS$. Suppose $p \in \mathcal{P}(\mathbf{F})$.

(a) Prove that $\operatorname{null} p(S)$ is invariant under T.

(b) Prove that $\operatorname{range} p(S)$ is invariant under T.

See 5.18 for the special case $S = T$.

4 Prove or give a counterexample: If A is a diagonal matrix and B is an upper-triangular matrix of the same size as A, then A and B commute.

5 Prove that a pair of operators on a finite-dimensional vector space commute if and only if their dual operators commute.

See 3.118 for the definition of the dual of an operator.

6 Suppose that V is a nonzero finite-dimensional complex vector space and $S, T \in \mathcal{L}(V)$ commute. Prove that there exist $\alpha, \lambda \in \mathbf{C}$ such that

$$\operatorname{range}(S - \alpha I) + \operatorname{range}(T - \lambda I) \neq V.$$

7 Suppose V is a complex vector space, $S \in \mathcal{L}(V)$ is diagonalizable, and $T \in \mathcal{L}(V)$ commutes with S. Prove that there is a basis of V such that S has a diagonal matrix with respect to this basis and T has an upper-triangular matrix with respect to this basis.

8 Suppose $m = 3$ in Example 5.72 and D_w, D_z are the commuting partial differentiation operators on $\mathcal{P}_3(\mathbf{C}^2, \mathbf{C})$ from that example. Find a basis of $\mathcal{P}_3(\mathbf{C}^2, \mathbf{C})$ with respect to which D_w and D_z each have an upper-triangular matrix.

9 Suppose V is a finite-dimensional nonzero complex vector space. Suppose that $\mathcal{E} \subseteq \mathcal{L}(V)$ is such that S and T commute for all $S, T \in \mathcal{E}$.

(a) Prove that there is a vector in V that is an eigenvector for every element of $\mathcal{E}$.

(b) Prove that there is a basis of V with respect to which every element of $\mathcal{E}$ has an upper-triangular matrix.

This exercise extends 5.78 and 5.80, which consider the case in which $\mathcal{E}$ contains only two elements. For this exercise, $\mathcal{E}$ may contain any number of elements, and $\mathcal{E}$ may even be an infinite set.

10 Give an example of two commuting operators S, T on a finite-dimensional real vector space such that $S + T$ has an eigenvalue that does not equal an eigenvalue of S plus an eigenvalue of T and ST has an eigenvalue that does not equal an eigenvalue of S times an eigenvalue of T.

This exercise shows that 5.81 does not hold on real vector spaces.

3 נניח ש־ $S, T \in \mathcal{L}(V)$ כך ש־ $ST = TS$. נניח ש־ $p \in \mathcal{P}(\mathbf{F})$.

א. הוכיחו ש־ $\operatorname{null} p(S)$ שמור תחת T.

ב. הוכיחו ש־ $\operatorname{range} p(S)$ שמור תחת T.

ראו 5.18 למקרה הפיוחד $S = T$.

4 הוכיחו או הביאו דוגמה נגדית: אם A מטריצה אלכסונית ו־ B מטריצה משולשית עליונה באותו הגודל כשל A, אזי A ו־ B מתחלפות.

5 הוכיחו שזוג אופרטורים מעל מרחב וקטורי מממד סופי מתחלף אם־ורק־אם האופרטורים הדואליים שלהם מתחלפים.

ראו 3.118 להגדרת אופרטור דואלי.

6 נניח ש־ V מרחב וקטורי מרוכב מממד סופי השונה מאפס ו־ $S, T \in \mathcal{L}(V)$ מתחלפים. הוכיחו שקיימים $\alpha, \lambda \in \mathbf{C}$ כך ש־

$$\operatorname{range}(S - \alpha I) + \operatorname{range}(T - \lambda I) \neq V.$$

7 נניח ש־ V מרחב וקטורי מרוכב, $S \in \mathcal{L}(V)$ לכסין, ו־ $T \in \mathcal{L}(V)$ מתחלף עם S. הוכיחו שקיים בסיס של V כך שיש ל־ S מטריצה אלכסונית ביחס לבסיס זה ויש ל־ T מטריצה משולשית עליונה ביחס לבסיס זה.

8 נניח ש־ $m = 3$ בדוגמה 5.72 ו־ D_w, D_z אופרטורי הנגזרת החלקית המתחלפים מעל $\mathcal{P}_3(\mathbf{C}^2, \mathbf{C})$ מאותה הדוגמה. מצאו בסיס של $\mathcal{P}_3(\mathbf{C}^2, \mathbf{C})$ אשר ביחס אליו יש לכל אחד מ־ D_w ו־ D_z מטריצה משולשית עליונה.

9 נניח ש־ V מרחב וקטורי מרוכב מממד סופי השונה מאפס. נניח ש־ $\mathcal{E} \subseteq \mathcal{L}(V)$ כך ש־ S ו־ T מתחלפים לכל $S, T \in \mathcal{E}$.

א. הוכיחו שיש וקטור ב־ V שהוא וקטור עצמי של כל איבר ב־ $\mathcal{E}$.

ב. הוכיחו שקיים בסיס של V אשר ביחס אליו לכל איבר ב־ $\mathcal{E}$ יש מטריצה משולשית עליונה.

תרגיל זה מרחיב את 5.78 ו־ 5.80, העוסקים במקרה שבו $\mathcal{E}$ מכיל רק שני איברים. לצורך תרגיל זה, $\mathcal{E}$ עשוי להכיל כל מספר של איברים, וייתכן אף ש־ $\mathcal{E}$ קבוצה אינסופית.

10 הביאו דוגמה של שני אופרטורים מתחלפים S, T מעל מרחב וקטורי ממשי מממד סופי כך שיש ל־ $S + T$ ערך עצמי שאינו שווה לערך עצמי של S ועוד ערך עצמי של T ושיש ל־ ST ערך עצמי שאינו שווה לערך עצמי של S כפול ערך עצמי של T.

תרגיל זה מראה כי 5.81 לא מתקיים מעל מרחב וקטורי ממשי.

Chapter 6

Inner Product Spaces

In making the definition of a vector space, we generalized the linear structure (addition and scalar multiplication) of $\mathbf{R}^2$ and $\mathbf{R}^3$. We ignored geometric features such as the notions of length and angle. These ideas are embedded in the concept of inner products, which we will investigate in this chapter.

Every inner product induces a norm, which you can think of as a length. This norm satisfies key properties such as the Pythagorean theorem, the triangle inequality, the parallelogram equality, and the Cauchy–Schwarz inequality.

The notion of perpendicular vectors in Euclidean geometry gets renamed to orthogonal vectors in the context of an inner product space. We will see that orthonormal bases are tremendously useful in inner product spaces. The Gram–Schmidt procedure constructs such bases. This chapter will conclude by putting together these tools to solve minimization problems.

standing assumptions for this chapter

- $\mathbf{F}$ denotes $\mathbf{R}$ or $\mathbf{C}$.
- V and W denote vector spaces over $\mathbf{F}$.

Matthew Petroff CC BY-SA

The George Peabody Library, now part of Johns Hopkins University, opened while James Sylvester (1814–1897) was the university's first mathematics professor. Sylvester's publications include the first use of the word ***matrix*** *in mathematics.*

פרק ו

מרחבי מכפלה פנימית

בתהליך יצירת ההגדרה של מרחב וקטורי, הכללנו את המבנה הליניארי (פעולות החיבור והכפל בסקלר) של $\mathbf{R}^2$ ו־ $\mathbf{R}^3$. התעלמנו מתכונות גאומטריות דוגמת המושגים אורך וזווית. רעיונות אלו מוטמעים במושג מכפלות פנימיות, שאותו נחקור בפרק זה.

כל מכפלה פנימית משרה נורמה, שעליה אפשר לחשוב כאורך. הנורמה מקיימת תכונות מפתח דוגמת משפט פיתגורס, אי־שוויון המשולש, כלל המקבילית, ואי־שוויון קושי־שוורץ.

המושג וקטורים מאונכים בגאומטריה האוקלידית מקבל שם חדש, וקטורים אורתוגונליים, בהקשר של מרחב מכפלה פנימית. אנו נראה כי בסיסים אורתונורמליים שימושיים להפליא במרחבי מכפלה פנימית. תהליך גרם־שמידט בונה בסיסים כאלו. הפרק מסתיים בשילוב כל הכלים הללו יחד לפתרון בעיות מזעור.

הנחות המוצא לפרק זה

- $\mathbf{F}$ מסמן $\mathbf{R}$ או $\mathbf{C}$.
- V ו־ W מסמנים מרחבים וקטוריים מעל $\mathbf{F}$.

Matthew Petroff CC BY-SA

הספרייה על שם ג'ורג' פיבודי, כעת חלק מאוניברסיטת ג'ון הופקינס, נפתחה בזמן שג'יימס סילבסטר (1814–1897) היה הפרופסור הראשון למתמטיקה באוניברסיטה. בפרסומיו של סילבסטר מופיע השימוש הראשון במילה **מטריצה** במתמטיקה.

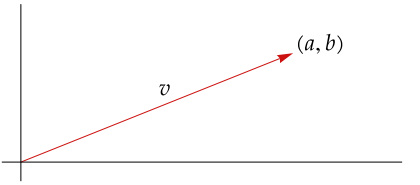
אלגברה ליניארית בדרך הישר, מהדורה רביעית, מחבר: שלדון אקסלר
Linear Algebra Done Right, fourth edition, by Sheldon Axler

6A Inner Products and Norms

Inner Products

To motivate the concept of inner product, think of vectors in $\mathbf{R}^2$ and $\mathbf{R}^3$ as arrows with initial point at the origin. The length of a vector v in $\mathbf{R}^2$ or $\mathbf{R}^3$ is called the *norm* of v and is denoted by $\|v\|$. Thus for $v = (a, b) \in \mathbf{R}^2$, we have

$$\|v\| = \sqrt{a^2 + b^2}.$$



This vector v has norm $\sqrt{a^2 + b^2}$.

Similarly, if $v = (a, b, c) \in \mathbf{R}^3$, then $\|v\| = \sqrt{a^2 + b^2 + c^2}$.

Even though we cannot draw pictures in higher dimensions, the generalization to $\mathbf{R}^n$ is easy: we define the norm of $x = (x_1, \dots, x_n) \in \mathbf{R}^n$ by

$$\|x\| = \sqrt{x_1^2 + \cdots + x_n^2}.$$

The norm is not linear on $\mathbf{R}^n$. To inject linearity into the discussion, we introduce the dot product.

6.1 definition: *dot product*

For $x, y \in \mathbf{R}^n$, the *dot product* of x and y, denoted by $x \cdot y$, is defined by

$$x \cdot y = x_1 y_1 + \cdots + x_n y_n,$$

where $x = (x_1, \dots, x_n)$ and $y = (y_1, \dots, y_n)$.

The dot product of two vectors in $\mathbf{R}^n$ is a number, not a vector. Notice that $x \cdot x = \|x\|^2$ for all $x \in \mathbf{R}^n$. Furthermore, the dot product on $\mathbf{R}^n$ has the following properties.

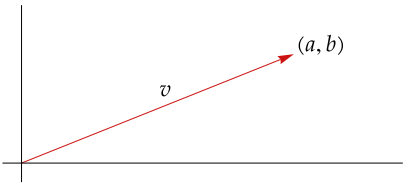
If we think of a vector as a point instead of as an arrow, then $\|x\|$ should be interpreted to mean the distance from the origin to the point x.

- $x \cdot x \geq 0$ for all $x \in \mathbf{R}^n$.
- $x \cdot x = 0$ if and only if $x = 0$.
- For $y \in \mathbf{R}^n$ fixed, the map from $\mathbf{R}^n$ to $\mathbf{R}$ that sends $x \in \mathbf{R}^n$ to $x \cdot y$ is linear.
- $x \cdot y = y \cdot x$ for all $x, y \in \mathbf{R}^n$.

An inner product is a generalization of the dot product. At this point you may be tempted to guess that an inner product is defined by abstracting the properties of the dot product discussed in the last paragraph. For real vector spaces, that guess is correct. However, so that we can make a definition that will be useful for both real and complex vector spaces, we need to examine the complex case before making the definition.

ו מכפלות פנימיות ונורמות

מכפלות פנימיות



הנורמה של וקטור v זה היא $\sqrt{a^2+b^2}$.

כמוטיבציה למושג המכפלה הפנימית, חשבו על וקטורים ב־ $\mathbf{R}^2$ ו־ $\mathbf{R}^3$ כחיצים שנקודת מוצאם היא בראשית הצירים. לאורך של וקטור v ב־ $\mathbf{R}^2$ או ב־ $\mathbf{R}^3$ אנו קוראים הנורמה של v שאותה אנו מסמנים ב־ $\|v\|$. לפיכך עבור $v = (a,b) \in \mathbf{R}^2$, נקבל

$$\|v\| = \sqrt{a^2+b^2}.$$

באופן דומה, אם $v = (a,b,c) \in \mathbf{R}^3$, אזי $\|v\| = \sqrt{a^2+b^2+c^2}$. אף שאין באפשרותנו לצייר תמונות בממדים גבוהים יותר, ההכללה ל־ $\mathbf{R}^n$ היא פשוטה: אנו מגדירים את הנורמה של $x = (x_1, \dots, x_n) \in \mathbf{R}^n$ על ידי

$$\|x\| = \sqrt{x_1^2 + \cdots + x_n^2}.$$

הנורמה אינה ליניארית מעל $\mathbf{R}^n$. כדי להזריק ליניאריות לתוך הדיון נציג את המכפלה הסקלרית.

6.1 הגדרה: מכפלה סקלרית

עבור $x, y \in \mathbf{R}^n$, המכפלה הסקלרית של x ו־ y, מסומנת על ידי $x \cdot y$, מוגדרת על ידי

$$x \cdot y = x_1 y_1 + \cdots + x_n y_n,$$

כאשר $x = (x_1, \dots, x_n)$ ו־ $y = (y_1, \dots, y_n)$.

אם אנו חושבים על וקטור כנקודה במקום כלל חץ, אזי הפרשנות ל־ $\|x\|$ צריכה להיות המרחק מראשית הצירים אל הנקודה x.

המכפלה הסקלרית של שני וקטורים ב־ $\mathbf{R}^n$ היא מספר, לא וקטור. נשים לב כי $x \cdot x = \|x\|^2$ לכל $x \in \mathbf{R}^n$. בנוסף, למכפלה הסקלרית מעל $\mathbf{R}^n$ התכונות הבאות.

- $x \cdot x \geq 0$ לכל $x \in \mathbf{R}^n$.
- $x \cdot x = 0$ אם־ורק־אם $x = 0$.
- עבור $y \in \mathbf{R}^n$ מקובע, ההעתקה מ־ $\mathbf{R}^n$ אל $\mathbf{R}$ השולחת $x \in \mathbf{R}^n$ אל $x \cdot y$ היא ליניארית.
- $x \cdot y = y \cdot x$ לכל $x, y \in \mathbf{R}^n$.

מכפלה פנימית היא הכללה של מכפלה סקלרית. בנקודה זו ייתכן שאתם מתפתים לנחש כי מכפלה פנימית מוגדרת באמצעות הפשטה של מאפייני המכפלה הסקלרית שתיארנו בפסקה הקודמת. עבור מרחבים וקטוריים ממשיים, ניחוש זה נכון. אולם, כדי שנוכל לנסח הגדרה שתהיה שימושית עבור שני המרחבים הווקטוריים, הממשי והמרוכב, עלינו לבחון את המקרה המרוכב טרם ניגש לניסוח ההגדרה.

Recall that if $\lambda = a + bi$, where $a, b \in \mathbf{R}$, then

- the absolute value of λ, denoted by $|\lambda|$, is defined by $|\lambda| = \sqrt{a^2 + b^2}$;
- the complex conjugate of λ, denoted by $\overline{\lambda}$, is defined by $\overline{\lambda} = a - bi$;
- $|\lambda|^2 = \lambda\overline{\lambda}$.

See Chapter 4 for the definitions and the basic properties of the absolute value and complex conjugate.

For $z = (z_1, \dots, z_n) \in \mathbf{C}^n$, we define the norm of z by

$$\|z\| = \sqrt{|z_1|^2 + \cdots + |z_n|^2}.$$

The absolute values are needed because we want $\|z\|$ to be a nonnegative number. Note that

$$\|z\|^2 = z_1\overline{z_1} + \cdots + z_n\overline{z_n}.$$

We want to think of $\|z\|^2$ as the inner product of z with itself, as we did in $\mathbf{R}^n$. The equation above thus suggests that the inner product of the vector $w = (w_1, \dots, w_n) \in \mathbf{C}^n$ with z should equal

$$w_1\overline{z_1} + \cdots + w_n\overline{z_n}.$$

If the roles of the w and z were interchanged, the expression above would be replaced with its complex conjugate. Thus we should expect that the inner product of w with z equals the complex conjugate of the inner product of z with w. With that motivation, we are now ready to define an inner product on V, which may be a real or a complex vector space.

One comment about the notation used in the next definition:

- For $\lambda \in \mathbf{C}$, the notation $\lambda \geq 0$ means λ is real and nonnegative.

6.2 definition: *inner product*

An *inner product* on V is a function that takes each ordered pair (u, v) of elements of V to a number $\langle u, v\rangle \in \mathbf{F}$ and has the following properties.

positivity
$\langle v, v\rangle \geq 0$ for all $v \in V$.

definiteness
$\langle v, v\rangle = 0$ if and only if $v = 0$.

additivity in first slot
$\langle u + v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$ for all $u, v, w \in V$.

homogeneity in first slot
$\langle \lambda u, v\rangle = \lambda\langle u, v\rangle$ for all $\lambda \in \mathbf{F}$ and all $u, v \in V$.

conjugate symmetry
$\langle u, v\rangle = \overline{\langle v, u\rangle}$ for all $u, v \in V$.

נזכיר כי אם $\lambda = a + bi$, כאשר $a, b \in \mathbf{R}$, אזי

- הערך המוחלט של λ, מסומן ב־ $|\lambda|$, מוגדר על ידי $|\lambda| = \sqrt{a^2 + b^2}$;
- הצמוד המרוכב של λ, מסומן ב־ $\overline{\lambda}$, מוגדר על ידי $\overline{\lambda} = a - bi$;
- $|\lambda|^2 = \lambda\overline{\lambda}$.

ראו פרק ד עבור ההגדרות והמאפיינים הבסיסיים של ערך מוחלט וצמוד מרוכב.

עבור $z = (z_1, \dots, z_n) \in \mathbf{C}^n$, אנו מגדירים את הנורמה של z על ידי

$$\|z\| = \sqrt{|z_1|^2 + \cdots + |z_n|^2}.$$

הערכים המוחלטים נדרשים מאחר שאנו מעוניינים ש־ $\|z\|$ יהיה מספר אי־שלילי. נשים לב כי

$$\|z\|^2 = z_1\overline{z_1} + \cdots + z_n\overline{z_n}.$$

אנו רוצים לחשוב על $\|z\|^2$ כמכפלה פנימית של z עם עצמו, בדומה למה שעשינו ב־ $\mathbf{R}^n$. המשוואה לעיל מרמזת לפיכך שהמכפלה הפנימית של הווקטור $w = (w_1, \dots, w_n) \in \mathbf{C}^n$ עם z צריכה להיות שווה ל־

$$w_1\overline{z_1} + \cdots + w_n\overline{z_n}.$$

אם תפקידיהם של w ו־ z היו מתחלפים, הביטוי לעיל היה מוחלף בצמוד המרוכב שלו. לפיכך עלינו לצפות שהמכפלה הפנימית של w עם z שווה לצמוד המרוכב של המכפלה הפנימית של z עם w. עם מוטיבציה זו, אנו מוכנים כעת לגשת ולהגדיר מכפלה פנימית מעל V, מרחב וקטורי ממשי או מרוכב.

הערה אחת בנוגע לסימון שבו אנו עושים שימוש בהגדרה הבאה:

- עבור $\lambda \in \mathbf{C}$, הסימון $\lambda \geq 0$ משמעותו ש־ λ מספר ממשי ואי־שלילי.

6.2 הגדרה: מכפלה פנימית

מכפלה פנימית מעל V היא פונקציה המעתיקה כל זוג סדור (u, v) של איברים ב־ V למספר $\langle u, v\rangle \in \mathbf{F}$ ומקיימת את התכונות הבאות.

חיוביות
$\langle v, v\rangle \geq 0$ לכל $v \in V$.

ודאות
$\langle v, v\rangle = 0$ אם־ורק־אם $v = 0$.

אדיטיביות ברכיב הראשון
$\langle u + v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$ לכל $u, v, w \in V$.

הומוגניות ברכיב הראשון
$\langle \lambda u, v\rangle = \lambda\langle u, v\rangle$ לכל $\lambda \in \mathbf{F}$ ולכל $u, v \in V$.

סימטריית צמוד (הרמיטיות)
$\langle u, v\rangle = \overline{\langle v, u\rangle}$ לכל $u, v \in V$.

Every real number equals its complex conjugate. Thus if we are dealing with a real vector space, then in the last condition above we can dispense with the complex conjugate and simply state that $\langle u, v\rangle = \langle v, u\rangle$ for all $u, v \in V$.

Most mathematicians define inner products as above, but many physicists use a definition that requires homogeneity in the second slot instead of the first slot.

6.3 example: *inner products*

(a) The *Euclidean inner product* on $\mathbf{F}^n$ is defined by

$$\big\langle (w_1, \dots, w_n), (z_1, \dots, z_n)\big\rangle = w_1\overline{z_1} + \cdots + w_n\overline{z_n}$$

for all $(w_1, \dots, w_n), (z_1, \dots, z_n) \in \mathbf{F}^n$.

(b) If $c_1, \dots, c_n$ are positive numbers, then an inner product can be defined on $\mathbf{F}^n$ by

$$\big\langle (w_1, \dots, w_n), (z_1, \dots, z_n)\big\rangle = c_1 w_1\overline{z_1} + \cdots + c_n w_n\overline{z_n}$$

for all $(w_1, \dots, w_n), (z_1, \dots, z_n) \in \mathbf{F}^n$.

(c) An inner product can be defined on the vector space of continuous real-valued functions on the interval $[-1, 1]$ by

$$\langle f, g\rangle = \int_{-1}^{1} fg$$

for all f, g continuous real-valued functions on $[-1, 1]$.

(d) An inner product can be defined on $\mathcal{P}(\mathbf{R})$ by

$$\langle p, q\rangle = p(0)q(0) + \int_{-1}^{1} p'q'$$

for all $p, q \in \mathcal{P}(\mathbf{R})$.

(e) An inner product can be defined on $\mathcal{P}(\mathbf{R})$ by

$$\langle p, q\rangle = \int_{0}^{\infty} p(x)q(x)e^{-x}\,dx$$

for all $p, q \in \mathcal{P}(\mathbf{R})$.

6.4 definition: *inner product space*

An *inner product space* is a vector space V along with an inner product on V.

The most important example of an inner product space is $\mathbf{F}^n$ with the Euclidean inner product given by (a) in the example above. When $\mathbf{F}^n$ is referred to as an inner product space, you should assume that the inner product is the Euclidean inner product unless explicitly told otherwise.

כל מספר ממשי שווה לצמוד המרוכב שלו. לפיכך אם אנו עוסקים במרחב וקטורי ממשי, אזי בתנאי האחרון של הגדרת מכפלה פנימית לעיל, נוכל לוותר על הצמוד המרוכב ופשוט לקבוע ש־ $\langle u, v\rangle = \langle v, u\rangle$ לכל $u, v \in V$.

רוב המתמטיקאים מגדירים מכפלות פנימיות כמובא לעיל, אך פיזיקאים רבים משתמשים בהגדרה הזורשת הומוגניות ברכיב השני במקום הרכיב הראשון.

6.3 דוגמה: מכפלות פנימיות

א. המכפלה הפנימית האוקלידית מעל $\mathbf{F}^n$ מוגדרת על ידי
$$\langle (w_1, \dots, w_n), (z_1, \dots, z_n)\rangle = w_1\overline{z_1} + \cdots + w_n\overline{z_n}$$
לכל $(w_1, \dots, w_n), (z_1, \dots, z_n) \in \mathbf{F}^n$.

ב. אם $c_1, \dots, c_n$ מספרים חיוביים, אזי מכפלה פנימית יכולה להיות מוגדרת מעל $\mathbf{F}^n$ על ידי
$$\langle (w_1, \dots, w_n), (z_1, \dots, z_n)\rangle = c_1w_1\overline{z_1} + \cdots + c_nw_n\overline{z_n}$$
לכל $(w_1, \dots, w_n), (z_1, \dots, z_n) \in \mathbf{F}^n$.

ג. מכפלה פנימית יכולה להיות מוגדרת מעל המרחב הווקטורי של הפונקציות הרציפות הממשיות מעל הקטע $[-1, 1]$ על ידי
$$\langle f, g\rangle = \int_{-1}^{1} fg$$
לכל f, g פונקציות רציפות ממשיות מעל $[-1, 1]$.

ד. מכפלה פנימית יכולה להיות מוגדרת מעל $\mathcal{P}(\mathbf{R})$ על ידי
$$\langle p, q\rangle = p(0)q(0) + \int_{-1}^{1} p'q'$$
לכל $p, q \in \mathcal{P}(\mathbf{R})$.

ה. מכפלה פנימית יכולה להיות מוגדרת מעל $\mathcal{P}(\mathbf{R})$ על ידי
$$\langle p, q\rangle = \int_0^\infty p(x)q(x)e^{-x}\,dx$$
לכל $p, q \in \mathcal{P}(\mathbf{R})$.

6.4 הגדרה: מרחב מכפלה פנימית

מרחב מכפלה פנימית הוא מרחב וקטורי V בצירוף מכפלה פנימית מעל V.

הדוגמה החשובה ביותר למרחב מכפלה פנימית היא $\mathbf{F}^n$ בצירוף המכפלה הפנימית האוקלידית המובאת בסעיף (א) בדוגמה לעיל. כאשר מתייחסים ל־ $\mathbf{F}^n$ כמרחב מכפלה פנימית, עליכם להניח כי המכפלה הפנימית היא המכפלה הפנימית האוקלידית, אלא אם כן צוין במפורש אחרת.

So that we do not have to keep repeating the hypothesis that V and W are inner product spaces, we make the following assumption.

6.5 notation: V, W

For the rest of this chapter and the next chapter, V and W denote inner product spaces over $\mathbf{F}$.

Note the slight abuse of language here. An inner product space is a vector space along with an inner product on that vector space. When we say that a vector space V is an inner product space, we are also thinking that an inner product on V is lurking nearby or is clear from the context (or is the Euclidean inner product if the vector space is $\mathbf{F}^n$).

6.6 *basic properties of an inner product*

(a) For each fixed $v \in V$, the function that takes $u \in V$ to $\langle u, v\rangle$ is a linear map from V to $\mathbf{F}$.

(b) $\langle 0, v\rangle = 0$ for every $v \in V$.

(c) $\langle v, 0\rangle = 0$ for every $v \in V$.

(d) $\langle u, v + w\rangle = \langle u, v\rangle + \langle u, w\rangle$ for all $u, v, w \in V$.

(e) $\langle u, \lambda v\rangle = \overline{\lambda}\langle u, v\rangle$ for all $\lambda \in \mathbf{F}$ and all $u, v \in V$.

Proof

(a) For $v \in V$, the linearity of $u \mapsto \langle u, v\rangle$ follows from the conditions of additivity and homogeneity in the first slot in the definition of an inner product.

(b) Every linear map takes 0 to 0. Thus (b) follows from (a).

(c) If $v \in V$, then the conjugate symmetry property in the definition of an inner product and (b) show that $\langle v, 0\rangle = \overline{\langle 0, v\rangle} = \overline{0} = 0$.

(d) Suppose $u, v, w \in V$. Then

$$\begin{aligned}\langle u, v + w\rangle &= \overline{\langle v + w, u\rangle}\\ &= \overline{\langle v, u\rangle + \langle w, u\rangle}\\ &= \overline{\langle v, u\rangle} + \overline{\langle w, u\rangle}\\ &= \langle u, v\rangle + \langle u, w\rangle.\end{aligned}$$

(e) Suppose $\lambda \in \mathbf{F}$ and $u, v \in V$. Then

$$\begin{aligned}\langle u, \lambda v\rangle &= \overline{\langle \lambda v, u\rangle}\\ &= \overline{\lambda\langle v, u\rangle}\\ &= \overline{\lambda}\,\overline{\langle v, u\rangle}\\ &= \overline{\lambda}\langle u, v\rangle.\end{aligned}$$

כדי שלא נצטרך לחזור שוב ושוב על ההנחות ש־ V ו־ W מרחבי מכפלה פנימית, אנו מניחים כעת את ההנחה הבאה.

סימון: V, W 6.5

עד לסוף פרק זה, ולאורכו של הפרק הבא, V ו־ W מייצגים מרחבי מכפלה פנימית מעל $\mathbf{F}$.

נשים לב לעיוות הלשוני הקל שביצענו. מרחב מכפלה פנימית הוא מרחב וקטורי בצירוף מכפלה פנימית מעל אותו מרחב וקטורי. כאשר אנו אומרים שמרחב וקטורי V הוא מרחב מכפלה פנימית, אנו לוקחים בחשבון שמכפלה פנימית מעל V אורבת לה בקרבת מקום או שזהות המכפלה ברורה מהקשר (או שזוהי המכפלה הפנימית האוקלידית במקרה שהמרחב הווקטורי הוא $\mathbf{F}^n$).

מאפיינים בסיסיים של מכפלה פנימית 6.6

א. לכל $v \in V$ מקובע, הפונקציה המעתיקה $u \in V$ אל $\langle u, v\rangle$ היא העתקה ליניארית מ־ V אל $\mathbf{F}$.

ב. $\langle 0, v\rangle = 0$ לכל $v \in V$.

ג. $\langle v, 0\rangle = 0$ לכל $v \in V$.

ד. $\langle u, v + w\rangle = \langle u, v\rangle + \langle u, w\rangle$ לכל $u, v, w \in V$.

ה. $\langle u, \lambda v\rangle = \overline{\lambda}\langle u, v\rangle$ לכל $\lambda \in \mathbf{F}$ ולכל $u, v \in V$.

הוכחה

א. עבור $v \in V$, הליניאריות של $u \mapsto \langle u, v\rangle$ נובעת מהתנאים לאדיטיביות והומוגניות ברכיב הראשון בהגדרת מכפלה פנימית.

ב. כל העתקה ליניארית מעתיקה 0 אל 0. לפיכך (ב) נובע מ־ (א).

ג. אם $v \in V$, אזי מתכונת סימטריית הצמוד (הרמיטיות) בהגדרת מכפלה פנימית יחד עם סעיף (ב) נקבל ש־ $\langle v, 0\rangle = \overline{\langle 0, v\rangle} = \overline{0} = 0$.

ד. נניח ש־ $u, v, w \in V$. אזי

$$\begin{aligned}\langle u, v + w\rangle &= \overline{\langle v + w, u\rangle}\\ &= \overline{\langle v, u\rangle + \langle w, u\rangle}\\ &= \overline{\langle v, u\rangle} + \overline{\langle w, u\rangle}\\ &= \langle u, v\rangle + \langle u, w\rangle.\end{aligned}$$

ה. נניח ש־ $\lambda \in \mathbf{F}$ ו־ $u, v \in V$. אזי

$$\begin{aligned}\langle u, \lambda v\rangle &= \overline{\langle \lambda v, u\rangle}\\ &= \overline{\lambda\langle v, u\rangle}\\ &= \overline{\lambda}\ \overline{\langle v, u\rangle}\\ &= \overline{\lambda}\langle u, v\rangle.\end{aligned}$$

■

Norms

Our motivation for defining inner products came initially from the norms of vectors on $\mathbf{R}^2$ and $\mathbf{R}^3$. Now we see that each inner product determines a norm.

6.7 definition: *norm*, $\|v\|$

For $v \in V$, the *norm* of v, denoted by $\|v\|$, is defined by

$$\|v\| = \sqrt{\langle v, v\rangle}.$$

6.8 example: *norms*

(a) If $(z_1, \dots, z_n) \in \mathbf{F}^n$ (with the Euclidean inner product), then

$$\|(z_1, \dots, z_n)\| = \sqrt{|z_1|^2 + \cdots + |z_n|^2}.$$

(b) For f in the vector space of continuous real-valued functions on $[-1, 1]$ and with inner product given as in 6.3(c), we have

$$\|f\| = \sqrt{\int_{-1}^{1} f^2}.$$

6.9 *basic properties of the norm*

Suppose $v \in V$.

(a) $\|v\| = 0$ if and only if $v = 0$.

(b) $\|\lambda v\| = |\lambda| \, \|v\|$ for all $\lambda \in \mathbf{F}$.

Proof

(a) The desired result holds because $\langle v, v\rangle = 0$ if and only if $v = 0$.

(b) Suppose $\lambda \in \mathbf{F}$. Then

$$\begin{aligned} \|\lambda v\|^2 &= \langle \lambda v, \lambda v\rangle \\ &= \lambda \langle v, \lambda v\rangle \\ &= \lambda \overline{\lambda} \langle v, v\rangle \\ &= |\lambda|^2 \, \|v\|^2. \end{aligned}$$

Taking square roots now gives the desired equality. ∎

The proof of (b) in the result above illustrates a general principle: working with norms squared is usually easier than working directly with norms.

נורמות

המוטיבציה שלנו להגדרת מכפלות פנימיות נבעה במקורה מהנורמות של הווקטורים מעל $\mathbf{R}^2$ ו־ $\mathbf{R}^3$. כעת נראה כי כל מכפלה פנימית קובעת נורמה.

6.7 הגדרה: נורמה, $\|v\|$

עבור $v \in V$, הנורמה של v, מסומנת ב־ $\|v\|$, מוגדרת על ידי
$$\|v\| = \sqrt{\langle v, v\rangle}.$$

6.8 דוגמה: נורמות

א. אם $(z_1, \dots, z_n) \in \mathbf{F}^n$ (בצירוף המכפלה הפנימית האוקלידית), אזי
$$\|(z_1, \dots, z_n)\| = \sqrt{|z_1|^2 + \cdots + |z_n|^2}.$$

ב. עבור f במרחב הווקטורי של פונקציות רציפות ממשיות מעל $[-1, 1]$ ובצירוף מכפלה פנימית הנתונה ב־ 6.3(ג), נקבל
$$\|f\| = \sqrt{\int_{-1}^{1} f^2}.$$

6.9 מאפיינים בסיסיים של הנורמה

נניח ש־ $v \in V$.

א. $\|v\| = 0$ אם־ורק־אם $v = 0$.

ב. $\|\lambda v\| = |\lambda| \, \|v\|$ לכל $\lambda \in \mathbf{F}$.

הוכחה

א. התוצאה הרצויה מתקיימת מאחר ש־ $\langle v, v\rangle = 0$ אם־ורק־אם $v = 0$.

ב. נניח ש־ $\lambda \in \mathbf{F}$. אזי
$$\begin{aligned}
\|\lambda v\|^2 &= \langle \lambda v, \lambda v\rangle \\
&= \lambda \langle v, \lambda v\rangle \\
&= \lambda \overline{\lambda} \langle v, v\rangle \\
&= |\lambda|^2 \, \|v\|^2.
\end{aligned}$$
הוצאת שורש כעת נותנת לנו את השוויון המבוקש.

ההוכחה של סעיף (ב) במסקנה לעיל ממחישה עיקרון כללי: עבודה עם ריבועי נורמות בדרך כלל קלה יותר מעבודה ישירה עם הנורמות.

Now we come to a crucial definition.

6.10 definition: *orthogonal*

Two vectors $u, v \in V$ are called *orthogonal* if $\langle u, v \rangle = 0$.

In the definition above, the order of the two vectors does not matter, because $\langle u, v \rangle = 0$ if and only if $\langle v, u \rangle = 0$. Instead of saying u and v are orthogonal, sometimes we say u is orthogonal to v.

*The word **orthogonal** comes from the Greek word **orthogonios**, which means right-angled.*

Exercise 15 asks you to prove that if u, v are nonzero vectors in $\mathbf{R}^2$, then

$$\langle u, v \rangle = \|u\| \, \|v\| \cos \theta,$$

where θ is the angle between u and v (thinking of u and v as arrows with initial point at the origin). Thus two nonzero vectors in $\mathbf{R}^2$ are orthogonal (with respect to the Euclidean inner product) if and only if the cosine of the angle between them is 0, which happens if and only if the vectors are perpendicular in the usual sense of plane geometry. Thus you can think of the word *orthogonal* as a fancy word meaning *perpendicular*.

We begin our study of orthogonality with an easy result.

6.11 *orthogonality and* 0

(a) 0 is orthogonal to every vector in V.

(b) 0 is the only vector in V that is orthogonal to itself.

Proof

(a) Recall that 6.6(b) states that $\langle 0, v \rangle = 0$ for every $v \in V$.

(b) If $v \in V$ and $\langle v, v \rangle = 0$, then $v = 0$ (by definition of inner product). ∎

For the special case $V = \mathbf{R}^2$, the next theorem was known over 3,500 years ago in Babylonia and then rediscovered and proved over 2,500 years ago in Greece. Of course, the proof below is not the original proof.

6.12 *Pythagorean theorem*

Suppose $u, v \in V$. If u and v are orthogonal, then

$$\|u + v\|^2 = \|u\|^2 + \|v\|^2.$$

Proof Suppose $\langle u, v \rangle = 0$. Then

$$\begin{aligned} \|u + v\|^2 &= \langle u + v, u + v \rangle \\ &= \langle u, u \rangle + \langle u, v \rangle + \langle v, u \rangle + \langle v, v \rangle \\ &= \|u\|^2 + \|v\|^2. \end{aligned}$$

∎

כעת אנו מגיעים להגדרה מכריעה.

6.10 הגדרה: אורתוגונלי

שני וקטורים $u, v \in V$ נקראים **אורתוגונליים** אם $\langle u, v\rangle = 0$.

הפילה **אורתוגונלי** מגיעה מהפילה היוונית **אוֹרתוֹגוֹניוֹס**, שמשמעותה ישר־זווית.

בהגדרה לעיל, אין חשיבות לסדר שני הווקטורים, מאחר ש־ $\langle u, v\rangle = 0$ אם־ורק־אם $\langle v, u\rangle = 0$. במקום לומר ש־ u ו־ v אורתוגונליים, לעיתים נאמר ש־ u אורתוגונלי ל־ v.

תרגיל 15 מבקש מכם להוכיח כי אם u, v וקטורים השונים מאפס ב־ $\mathbf{R}^2$, אזי

$$\langle u, v\rangle = \|u\| \, \|v\| \cos\theta,$$

כאשר θ היא הזווית בין u ל־ v (כשחושבים על u ו־ v כחיצים שנקודת מוצאם בראשית הצירים). לפיכך שני וקטורים השונים מאפס ב־ $\mathbf{R}^2$ אורתוגונליים (ביחס למכפלה הפנימית האוקלידית) אם־ורק־אם קוסינוס הזווית ביניהם שווה 0, תוצאה המתקבלת אם־ורק־אם הווקטורים אנכיים במובן הרגיל של גאומטריית המישור. לפיכך אפשר לחשוב על המילה אורתוגונלי כמילה מהודרת שמשמעותה אנכי.

אנו פותחים את לימוד האורתוגונליות עם מסקנה קלה.

6.11 אורתוגונליות ו־ 0

א. 0 אורתוגונלי לכל וקטור ב־ V.

ב. 0 הוא הווקטור היחיד ב־ V האורתוגונלי לעצמו.

הוכחה

א. נזכיר כי ב־ 6.6(ב) נקבע כי $\langle 0, v\rangle = 0$ לכל $v \in V$.

ב. אם $v \in V$ ו־ $\langle v, v\rangle = 0$, אזי $v = 0$ (לפי הגדרת מכפלה פנימית). ■

עבור המקרה הפרטי $V = \mathbf{R}^2$, המסקנה הבאה הייתה ידועה כבר לפני למעלה מ־ 3,500 שנה בבבל העתיקה ולאחר מכן התגלתה מחדש והוכחה לפני למעלה מ־ 2,500 שנה ביוון העתיקה. כמובן שההוכחה להלן אינה ההוכחה המקורית.

6.12 משפט פיתגורס

נניח ש־ $u, v \in V$. אם u ו־ v אורתוגונליים, אזי

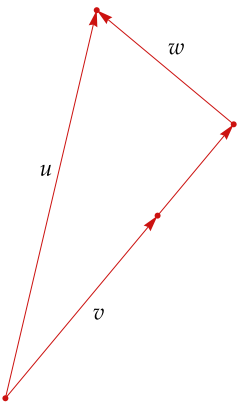
$$\|u + v\|^2 = \|u\|^2 + \|v\|^2.$$

הוכחה נניח ש־ $\langle u, v\rangle = 0$. אזי

$$\begin{aligned}\|u + v\|^2 &= \langle u + v, u + v\rangle \\ &= \langle u, u\rangle + \langle u, v\rangle + \langle v, u\rangle + \langle v, v\rangle \\ &= \|u\|^2 + \|v\|^2.\end{aligned}$$

■

Suppose $u, v \in V$, with $v \neq 0$. We would like to write u as a scalar multiple of v plus a vector w orthogonal to v, as suggested in the picture here.



An orthogonal decomposition:
u expressed as a scalar multiple of v plus a vector orthogonal to v.

To discover how to write u as a scalar multiple of v plus a vector orthogonal to v, let $c \in \mathbf{F}$ denote a scalar. Then

$$u = cv + (u - cv).$$

Thus we need to choose c so that v is orthogonal to $(u - cv)$. Hence we want

$$0 = \langle u - cv, v \rangle = \langle u, v \rangle - c\|v\|^2.$$

The equation above shows that we should choose c to be $\langle u, v \rangle / \|v\|^2$. Making this choice of c, we can write

$$u = \frac{\langle u, v \rangle}{\|v\|^2} v + \left(u - \frac{\langle u, v \rangle}{\|v\|^2} v \right).$$

As you should verify, the equation displayed above explicitly writes u as a scalar multiple of v plus a vector orthogonal to v. Thus we have proved the following key result.

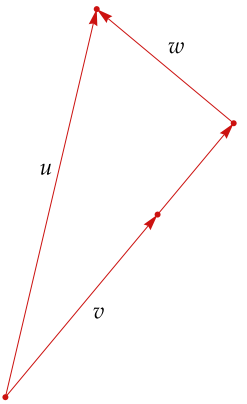
6.13 *an orthogonal decomposition*

Suppose $u, v \in V$, with $v \neq 0$. Set $c = \dfrac{\langle u, v \rangle}{\|v\|^2}$ and $w = u - \dfrac{\langle u, v \rangle}{\|v\|^2} v$. Then

$$u = cv + w \quad \text{and} \quad \langle w, v \rangle = 0.$$

The orthogonal decomposition 6.13 will be used in the proof of the Cauchy–Schwarz inequality, which is our next result and is one of the most important inequalities in mathematics.

נניח ש־ $u, v \in V$, עם $v \neq 0$. ברצוננו לבטא את u ככפולת סקלר של v ועוד וקטור w אורתוגונלי ל־ v, כפי שמוצע בתמונה להלן.



פירוק אורתוגונלי:
u מבוטא ככפולת סקלר של v ועוד וקטור אורתוגונלי ל־ v.

לגילוי הנוסחה כיצד ניתן לבטא את u ככפולת סקלר של v ועוד וקטור אורתוגונלי ל־ v, יהי $c \in \mathbf{F}$ סקלר. אזי

$$u = cv + (u - cv).$$

לפיכך עלינו לבחור c כך ש־ v אורתוגונלי ל־ $(u - cv)$. לכן, אנו מעוניינים ב־

$$0 = \langle u - cv, v \rangle = \langle u, v \rangle - c\|v\|^2.$$

המשוואה לעיל מורה לנו כי עלינו לבחור את c כ־ $\langle u, v \rangle / \|v\|^2$. בחירה זו של c מאפשרת לנו לכתוב את המשוואה הבאה

$$u = \frac{\langle u, v \rangle}{\|v\|^2} v + \left(u - \frac{\langle u, v \rangle}{\|v\|^2} v \right).$$

כפי שעליכם לוודא, המשוואה המוצגת כאן מבטאת את u ככפולת סקלר של v ועוד וקטור אורתוגונלי ל־ v באופן מפורש. לפיכך הוכחנו את מסקנת המפתח הבאה.

6.13 פירוק אורתוגונלי

נניח ש־ $u, v \in V$, עם $v \neq 0$. נגדיר $c = \dfrac{\langle u, v \rangle}{\|v\|^2}$ ו־ $w = u - \dfrac{\langle u, v \rangle}{\|v\|^2} v$. אזי

$$u = cv + w \quad \text{וגם} \quad \langle w, v \rangle = 0.$$

הפירוק האורתוגונלי 6.13 ישמש אותנו בהוכחת אי־שוויון קושי־שוורץ, שהוא המסקנה הבאה שלנו והוא אחד מהאי־שוויונות החשובים ביותר במתמטיקה.

6.14 *Cauchy–Schwarz inequality*

Suppose $u, v \in V$. Then

$$|\langle u, v\rangle| \le \|u\| \, \|v\|.$$

This inequality is an equality if and only if one of u, v is a scalar multiple of the other.

Proof If $v = 0$, then both sides of the desired inequality equal 0. Thus we can assume that $v \ne 0$. Consider the orthogonal decomposition

$$u = \frac{\langle u, v\rangle}{\|v\|^2} v + w$$

given by 6.13, where w is orthogonal to v. By the Pythagorean theorem,

$$\begin{aligned} \|u\|^2 &= \left\| \frac{\langle u, v\rangle}{\|v\|^2} v \right\|^2 + \|w\|^2 \\ &= \frac{|\langle u, v\rangle|^2}{\|v\|^2} + \|w\|^2 \\ &\ge \frac{|\langle u, v\rangle|^2}{\|v\|^2}. \end{aligned} \tag{6.15}$$

Multiplying both sides of this inequality by $\|v\|^2$ and then taking square roots gives the desired inequality.

The proof in the paragraph above shows that the Cauchy–Schwarz inequality is an equality if and only if 6.15 is an equality. This happens if and only if $w = 0$. But $w = 0$ if and only if u is a multiple of v (see 6.13). Thus the Cauchy–Schwarz inequality is an equality if and only if u is a scalar multiple of v or v is a scalar multiple of u (or both; the phrasing has been chosen to cover cases in which either u or v equals 0).

Augustin-Louis Cauchy (1789–1857) proved 6.16(a) in 1821. In 1859, Cauchy's student Viktor Bunyakovsky (1804–1889) proved integral inequalities like the one in 6.16(b). A few decades later, similar discoveries by Hermann Schwarz (1843–1921) attracted more attention and led to the name of this inequality.

6.16 example: *Cauchy–Schwarz inequality*

(a) If $x_1, \dots, x_n, y_1, \dots, y_n \in \mathbf{R}$, then

$$(x_1 y_1 + \cdots + x_n y_n)^2 \le (x_1^2 + \cdots + x_n^2)(y_1^2 + \cdots + y_n^2),$$

as follows from applying the Cauchy–Schwarz inequality to the vectors $(x_1, \dots, x_n), (y_1, \dots, y_n) \in \mathbf{R}^n$, using the usual Euclidean inner product.

6.14 אי־שוויון קושי־שוורץ

נניח ש־ $u, v \in V$. אזי

$$|\langle u, v\rangle| \leq \|u\| \, \|v\|.$$

אי־שוויון זה הוא שוויון אם־ורק־אם אחד מבין u ו־ v הוא כפולת סקלר של האחר.

הוכחה אם $v = 0$, אזי שני אגפי האי־שוויון המבוקש שווים 0. לפיכך נוכל להניח ש־ $v \neq 0$. נתבונן בפירוק האורתוגונלי

$$u = \frac{\langle u, v\rangle}{\|v\|^2} v + w$$

שקיבלנו ב־ 6.13, כאשר w אורתוגונלי ל־ v. לפי משפט פיתגורס,

$$\begin{aligned} \|u\|^2 &= \left\| \frac{\langle u, v\rangle}{\|v\|^2} v \right\|^2 + \|w\|^2 \\ &= \frac{|\langle u, v\rangle|^2}{\|v\|^2} + \|w\|^2 \\ &\geq \frac{|\langle u, v\rangle|^2}{\|v\|^2}. \end{aligned} \tag{6.15}$$

כעת נכפיל את שני אגפי האי־שוויון ב־ $\|v\|^2$ ולאחר מכן נוציא שורש ריבועי ונקבל את האי־שוויון המבוקש.

ההוכחה בפסקה למעלה מראה לנו כי אי־שוויון קושי־שוורץ הוא שוויון אם־ורק־אם 6.15 שוויון. הדבר קורה אם־ורק־אם $w = 0$. אבל כאמור לעיל, $w = 0$ אם־ורק־אם u כפולת סקלר של v (ראו 6.13). לפיכך אי־שוויון קושי־שוורץ הוא שוויון אם־ורק־אם u כפולת סקלר של v או v כפולת סקלר של u (או שניהם; הניסוח נבחר כדי לכסות מקרים שבהם אחד משני הווקטורים u או v שווה 0). ■

אוגוסטן־לואי קושי (1789–1857) הוכיח את 6.16(א) בשנת 1821. בשנת 1859, תלמידו של קושי ויקטור בּוּנְיָקוֹבְסְקִי (1804–1889) הוכיח את הגרסה האינטגרלית של האי־שוויון בדומה למופיע ב־ 6.16(ב). מספר עשורים מאוחר יותר, תגליות דומות של הרמן שוורץ (1843–1921) זכו לפרסום רב יותר והובילו להתבססות השם של אי־שוויון זה.

6.16 דוגמה: אי־שוויון קושי־שוורץ

א. אם $x_1, \dots, x_n, y_1, \dots, y_n \in \mathbf{R}$, אזי

$$(x_1 y_1 + \cdots + x_n y_n)^2 \leq (x_1^2 + \cdots + x_n^2)(y_1^2 + \cdots + y_n^2),$$

כפי שנובע מיישום אי־שוויון קושי־שוורץ על הווקטורים $(x_1, \dots, x_n), (y_1, \dots, y_n) \in \mathbf{R}^n$, בשימוש במכפלה הפנימית האוקלידית הרגילה.

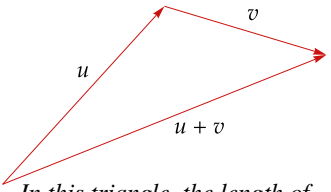
(b) If f, g are continuous real-valued functions on $[-1, 1]$, then

$$\left|\int_{-1}^{1} fg\right|^2 \le \left(\int_{-1}^{1} f^2\right)\left(\int_{-1}^{1} g^2\right),$$

as follows from applying the Cauchy–Schwarz inequality to Example 6.3(c).

The next result, called the triangle inequality, has the geometric interpretation that the length of each side of a triangle is less than the sum of the lengths of the other two sides.

Note that the triangle inequality implies that the shortest polygonal path between two points is a single line segment (a polygonal path consists of line segments).



In this triangle, the length of $u + v$ is less than the length of u plus the length of v.

6.17 ***triangle inequality***

Suppose $u, v \in V$. Then

$$\|u + v\| \le \|u\| + \|v\|.$$

This inequality is an equality if and only if one of u, v is a nonnegative real multiple of the other.

Proof We have

$$\begin{aligned}
\|u + v\|^2 &= \langle u + v, u + v\rangle \\
&= \langle u, u\rangle + \langle v, v\rangle + \langle u, v\rangle + \langle v, u\rangle \\
&= \langle u, u\rangle + \langle v, v\rangle + \langle u, v\rangle + \overline{\langle u, v\rangle} \\
&= \|u\|^2 + \|v\|^2 + 2\,\mathrm{Re}\langle u, v\rangle \\
&\le \|u\|^2 + \|v\|^2 + 2|\langle u, v\rangle| \qquad 6.18 \\
&\le \|u\|^2 + \|v\|^2 + 2\|u\|\,\|v\| \qquad 6.19 \\
&= \big(\|u\| + \|v\|\big)^2,
\end{aligned}$$

where 6.19 follows from the Cauchy–Schwarz inequality (6.14). Taking square roots of both sides of the inequality above gives the desired inequality.

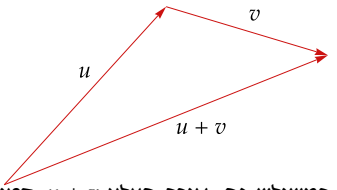
The proof above shows that the triangle inequality is an equality if and only if we have equality in 6.18 and 6.19. Thus we have equality in the triangle inequality if and only if

$$\langle u, v\rangle = \|u\|\,\|v\|. \qquad 6.20$$

If one of u, v is a nonnegative real multiple of the other, then 6.20 holds. Conversely, suppose 6.20 holds. Then the condition for equality in the Cauchy–Schwarz inequality (6.14) implies that one of u, v is a scalar multiple of the other. This scalar must be a nonnegative real number, by 6.20, completing the proof. ∎

For the reverse triangle inequality, see Exercise 20.

ב. אם f,g פונקציות רציפות ממשיות מעל $[-1,1]$, אזי
$$\left|\int_{-1}^{1} fg\right|^2 \le \left(\int_{-1}^{1} f^2\right)\left(\int_{-1}^{1} g^2\right),$$
כפי שנובע מיישום אי־שוויון קושי־שוורץ על דוגמה 6.3(ג).



במשולש זה, אורך הצלע $u+v$ קטן מאורך הצלע u ועוד אורך הצלע v.

המסקנה הבאה נקראת אי־שוויון המשולש. הפרשנות הגאומטרית של מסקנה זו היא שאורך כל צלע במשולש נתון קטן מסכום האורכים של שתי הצלעות האחרות.

נשים לב כי מאי־שוויון המשולש נובע כי הקו השבור הקצר ביותר בין שתי נקודות הוא קטע ישר יחיד המחבר ביניהן (קו שבור הוא קו הבנוי מסדרה של קטעים ישרים המחוברים זה לזה בנקודות הקצה).

6.17 אי־שוויון המשולש

נניח ש־ $u,v \in V$. אזי
$$\|u+v\| \le \|u\| + \|v\|.$$
האי־שוויון הזה הוא שוויון אם־ורק־אם אחד מבין u ו־ v הוא כפולת מספר ממשי אי־שלילי של האחר.

הוכחה מתקיים

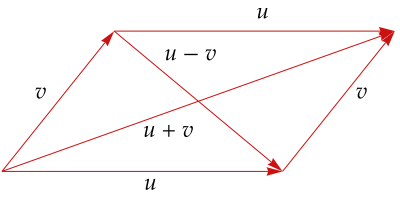
$$\begin{aligned}
\|u+v\|^2 &= \langle u+v, u+v\rangle \\
&= \langle u,u\rangle + \langle v,v\rangle + \langle u,v\rangle + \langle v,u\rangle \\
&= \langle u,u\rangle + \langle v,v\rangle + \langle u,v\rangle + \overline{\langle u,v\rangle} \\
&= \|u\|^2 + \|v\|^2 + 2\,\mathrm{Re}\langle u,v\rangle \\
&\le \|u\|^2 + \|v\|^2 + 2|\langle u,v\rangle| && 6.18 \\
&\le \|u\|^2 + \|v\|^2 + 2\|u\|\,\|v\| && 6.19 \\
&= (\|u\| + \|v\|)^2,
\end{aligned}$$

כאשר 6.19 נובע מאי־שוויון קושי־שוורץ (6.14). בהוצאת שורשים לשני אגפי האי־שוויון לעיל נקבל את האי־שוויון המבוקש.

ההוכחה לעיל מראה כי אי־שוויון המשולש הוא שוויון אם־ורק־אם יש לנו שוויון ב־ 6.18 וב־ 6.19. לפיכך באי־שוויון המשולש מתקיים שוויון אם־ורק־אם
$$\langle u,v\rangle = \|u\|\,\|v\|. \qquad 6.20$$
אם אחד מ־ u,v כפולת מספר ממשי אי־שלילי של האחר, אזי 6.20 מתקיים. בכיוון ההפוך, נניח ש־ 6.20 מתקיים. אזי מהתנאי לשוויון באי־שוויון קושי־שוורץ (6.14) נובע שאחד מבין u,v הוא כפולת סקלר של האחר. סקלר זה חייב להיות ממשי אי־שלילי לפי 6.20, ובזאת הושלמה ההוכחה. ∎

לאי־שוויון המשולש ההפוך, ראו תרגיל 20.

The next result is called the parallelogram equality because of its geometric interpretation: in every parallelogram, the sum of the squares of the lengths of the diagonals equals the sum of the squares of the lengths of the four sides. Note that the proof here is more straightforward than the usual proof in Euclidean geometry.



The diagonals of this parallelogram are $u + v$ and $u - v$.

6.21 *parallelogram equality*

Suppose $u, v \in V$. Then

$$\|u + v\|^2 + \|u - v\|^2 = 2(\|u\|^2 + \|v\|^2).$$

Proof We have

$$\begin{aligned}\|u + v\|^2 + \|u - v\|^2 &= \langle u + v, u + v\rangle + \langle u - v, u - v\rangle \\ &= \|u\|^2 + \|v\|^2 + \langle u, v\rangle + \langle v, u\rangle \\ &\quad + \|u\|^2 + \|v\|^2 - \langle u, v\rangle - \langle v, u\rangle \\ &= 2(\|u\|^2 + \|v\|^2),\end{aligned}$$

as desired. ∎

Exercises 6A

1 Prove or give a counterexample: If $v_1, \dots, v_m \in V$, then

$$\sum_{j=1}^{m} \sum_{k=1}^{m} \langle v_j, v_k\rangle \geq 0.$$

2 Suppose $S \in \mathcal{L}(V)$. Define $\langle \cdot, \cdot \rangle_1$ by

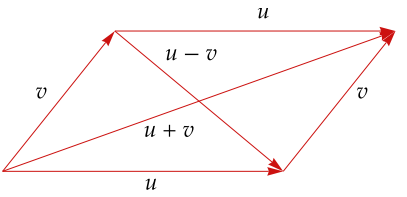
$$\langle u, v\rangle_1 = \langle Su, Sv\rangle$$

for all $u, v \in V$. Show that $\langle \cdot, \cdot \rangle_1$ is an inner product on V if and only if S is injective.

3 (a) Show that the function taking an ordered pair $((x_1, x_2), (y_1, y_2))$ of elements of $\mathbf{R}^2$ to $|x_1 y_1| + |x_2 y_2|$ is not an inner product on $\mathbf{R}^2$.

(b) Show that the function taking an ordered pair $((x_1, x_2, x_3), (y_1, y_2, y_3))$ of elements of $\mathbf{R}^3$ to $x_1 y_1 + x_3 y_3$ is not an inner product on $\mathbf{R}^3$.

4 Suppose $T \in \mathcal{L}(V)$ is such that $\|Tv\| \leq \|v\|$ for every $v \in V$. Prove that $T - \sqrt{2}I$ is injective.



אלכסוני מקבילית זו הם $u+v$ ו־ $u-v$.

המסקנה הבאה נקראת כלל המקבילית לאור הפרשנות הגאומטרית המוענקת לה: בכל מקבילית, סכום ריבועי אורכי האלכסונים שווה לסכום ריבועי אורכי ארבע הצלעות. נשים לב כי ההוכחה המובאת כאן פשוטה וישירה יותר מזו הנפוצה בגאומטריה אוקלידית.

6.21 כלל המקבילית

נניח ש־ $u, v \in V$. אזי

$$\|u+v\|^2 + \|u-v\|^2 = 2(\|u\|^2 + \|v\|^2).$$

הוכחה מתקיים

$$\begin{aligned}\|u+v\|^2 + \|u-v\|^2 &= \langle u+v, u+v\rangle + \langle u-v, u-v\rangle \\ &= \|u\|^2 + \|v\|^2 + \langle u, v\rangle + \langle v, u\rangle \\ &\quad + \|u\|^2 + \|v\|^2 - \langle u, v\rangle - \langle v, u\rangle \\ &= 2(\|u\|^2 + \|v\|^2),\end{aligned}$$

כפי שרצינו.

תרגילים 6ו

1 הוכיחו או הביאו דוגמה נגדית: אם $v_1, \ldots, v_m \in V$, אזי

$$\sum_{j=1}^{m}\sum_{k=1}^{m}\langle v_j, v_k\rangle \geq 0.$$

2 נניח ש־ $S \in \mathcal{L}(V)$. נגדיר $\langle \cdot, \cdot\rangle_1$ על ידי

$$\langle u, v\rangle_1 = \langle Su, Sv\rangle$$

לכל $u, v \in V$. הראו ש־ $\langle \cdot, \cdot\rangle_1$ מכפלה פנימית מעל V אם־ורק־אם S חד־חד־ערכי.

3 א. הראו שהפונקציה המעתיקה זוג סדור $((x_1, x_2), (y_1, y_2))$ של איברי $\mathbf{R}^2$ אל $|x_1y_1| + |x_2y_2|$ אינה מכפלה פנימית מעל $\mathbf{R}^2$.

ב. הראו שהפונקציה המעתיקה זוג סדור $((x_1, x_2, x_3), (y_1, y_2, y_3))$ של איברי $\mathbf{R}^3$ אל $x_1y_1 + x_3y_3$ אינה מכפלה פנימית מעל $\mathbf{R}^3$.

4 נניח ש־ $T \in \mathcal{L}(V)$ אופרטור כך ש־ $\|Tv\| \leq \|v\|$ לכל $v \in V$. הוכיחו ש־ $T - \sqrt{2}I$ חד־חד־ערכי.

5 Suppose V is a real inner product space.

(a) Show that $\langle u+v, u-v\rangle = \|u\|^2 - \|v\|^2$ for every $u, v \in V$.

(b) Show that if $u, v \in V$ have the same norm, then $u+v$ is orthogonal to $u-v$.

(c) Use (b) to show that the diagonals of a rhombus are perpendicular to each other.

6 Suppose $u, v \in V$. Prove that $\langle u, v\rangle = 0 \iff \|u\| \le \|u + av\|$ for all $a \in \mathbf{F}$.

7 Suppose $u, v \in V$. Prove that $\|au + bv\| = \|bu + av\|$ for all $a, b \in \mathbf{R}$ if and only if $\|u\| = \|v\|$.

8 Suppose $a, b, c, x, y \in \mathbf{R}$ and $a^2 + b^2 + c^2 + x^2 + y^2 \le 1$. Prove that $a + b + c + 4x + 9y \le 10$.

9 Suppose $u, v \in V$ and $\|u\| = \|v\| = 1$ and $\langle u, v\rangle = 1$. Prove that $u = v$.

10 Suppose $u, v \in V$ and $\|u\| \le 1$ and $\|v\| \le 1$. Prove that

$$\sqrt{1 - \|u\|^2}\sqrt{1 - \|v\|^2} \le 1 - |\langle u, v\rangle|.$$

11 Find vectors $u, v \in \mathbf{R}^2$ such that u is a scalar multiple of $(1, 3)$, v is orthogonal to $(1, 3)$, and $(1, 2) = u + v$.

12 Suppose a, b, c, d are positive numbers.

(a) Prove that $(a+b+c+d)\left(\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}\right) \ge 16$.

(b) For which positive numbers a, b, c, d is the inequality above an equality?

13 Show that the square of an average is less than or equal to the average of the squares. More precisely, show that if $a_1, \dots, a_n \in \mathbf{R}$, then the square of the average of $a_1, \dots, a_n$ is less than or equal to the average of $a_1^2, \dots, a_n^2$.

14 Suppose $v \in V$ and $v \ne 0$. Prove that $v/\|v\|$ is the unique closest element on the unit sphere of V to v. More precisely, prove that if $u \in V$ and $\|u\| = 1$, then

$$\left\| v - \frac{v}{\|v\|} \right\| \le \|v - u\|,$$

with equality only if $u = v/\|v\|$.

15 Suppose u, v are nonzero vectors in $\mathbf{R}^2$. Prove that

$$\langle u, v\rangle = \|u\|\,\|v\| \cos\theta,$$

where θ is the angle between u and v (thinking of u and v as arrows with initial point at the origin).

Hint: Use the law of cosines on the triangle formed by u, v, and u − v.

5 נניח ש־ V מרחב מכפלה פנימית ממשי.

א. הראו ש־ $\langle u+v, u-v\rangle = \|u\|^2 - \|v\|^2$ לכל $u, v \in V$.

ב. הראו שאם ל־ $u, v \in V$ אותה נורמה, אזי $u+v$ אורתוגונלי ל־ $u-v$.

ג. השתמשו בסעיף (ב) כדי להראות כי אלכסוני המעוין אנכיים זה לזה.

6 נניח ש־ $u, v \in V$. הוכיחו ש־ $\|u\| \le \|u+av\| \iff \langle u, v\rangle = 0$ לכל $a \in \mathbf{F}$.

7 נניח ש־ $u, v \in V$. הוכיחו ש־ $\|au+bv\| = \|bu+av\|$ לכל $a, b \in \mathbf{R}$ אם־ורק־אם $\|u\| = \|v\|$.

8 נניח ש־ $a, b, c, x, y \in \mathbf{R}$ ו־ $a^2+b^2+c^2+x^2+y^2 \le 1$. הוכיחו ש־ $a+b+c+4x+9y \le 10$.

9 נניח ש־ $u, v \in V$ ו־ $\|u\| = \|v\| = 1$ ו־ $\langle u, v\rangle = 1$. הוכיחו ש־ $u = v$.

10 נניח ש־ $u, v \in V$ ו־ $\|u\| \le 1$ ו־ $\|v\| \le 1$. הוכיחו ש־
$$\sqrt{1-\|u\|^2}\sqrt{1-\|v\|^2} \le 1 - |\langle u, v\rangle|.$$

11 מצאו וקטורים $u, v \in \mathbf{R}^2$ כך שהווקטור u כפולת סקלר של $(1,3)$, הווקטור v אורתוגונלי ל־ $(1,3)$, והווקטור $u+v = (1,2)$.

12 נניח ש־ a, b, c, d מספרים חיוביים.

א. הוכיחו ש־ $(a+b+c+d)\left(\frac{1}{a}+\frac{1}{b}+\frac{1}{c}+\frac{1}{d}\right) \ge 16$.

ב. עבור אילו מספרים חיוביים a, b, c, d האי־שוויון לעיל הוא שוויון?

13 הראו שריבוע של ממוצע הוא קטן או שווה לממוצע הריבועים. ביתר דיוק, הראו כי אם $a_1, \dots, a_n \in \mathbf{R}$, אזי ריבוע הממוצע של $a_1, \dots, a_n$ הוא קטן או שווה לממוצע של $a_1^2, \dots, a_n^2$.

14 נניח ש־ $v \in V$ ו־ $v \ne 0$. הוכיחו ש־ $v/\|v\|$ הוא האיבר היחיד של סְפֵירַת היחידה של V הקרוב ביותר ל־ v. ביתר דיוק, הוכיחו כי אם $u \in V$ ו־ $\|u\| = 1$, אזי
$$\left\| v - \frac{v}{\|v\|} \right\| \le \|v-u\|,$$
עם שוויון רק במקרה ש־ $u = v/\|v\|$.

15 נניח ש־ u, v וקטורים השונים מאפס ב־ $\mathbf{R}^2$. הוכיחו ש־
$$\langle u, v\rangle = \|u\|\,\|v\| \cos\theta,$$
כאשר θ היא הזווית בין u ל־ v (כשחושבים על u ו־ v כחיצים שנקודת מוצאם בראשית הצירים).

רמז: השתמשו במשפט הקוסינוסים על המשולש הנוצר על ידי u, v, ו־ $u-v$.

16 The angle between two vectors (thought of as arrows with initial point at the origin) in $\mathbf{R}^2$ or $\mathbf{R}^3$ can be defined geometrically. However, geometry is not as clear in $\mathbf{R}^n$ for $n > 3$. Thus the angle between two nonzero vectors $x, y \in \mathbf{R}^n$ is defined to be

$$\arccos \frac{\langle x, y \rangle}{\|x\| \, \|y\|},$$

where the motivation for this definition comes from Exercise 15. Explain why the Cauchy–Schwarz inequality is needed to show that this definition makes sense.

17 Prove that

$$\Big(\sum_{k=1}^{n} a_k b_k\Big)^2 \le \Big(\sum_{k=1}^{n} k a_k^2\Big)\Big(\sum_{k=1}^{n} \frac{b_k^2}{k}\Big)$$

for all real numbers $a_1, \dots, a_n$ and $b_1, \dots, b_n$.

18 (a) Suppose $f : [1, \infty) \to [0, \infty)$ is continuous. Show that

$$\Big(\int_1^\infty f\Big)^2 \le \int_1^\infty x^2 (f(x))^2 \, dx.$$

(b) For which continuous functions $f : [1, \infty) \to [0, \infty)$ is the inequality in (a) an equality with both sides finite?

19 Suppose $v_1, \dots, v_n$ is a basis of V and $T \in \mathcal{L}(V)$. Prove that if λ is an eigenvalue of T, then

$$|\lambda|^2 \le \sum_{j=1}^{n} \sum_{k=1}^{n} |\mathcal{M}(T)_{j,k}|^2,$$

where $\mathcal{M}(T)_{j,k}$ denotes the entry in row j, column k of the matrix of T with respect to the basis $v_1, \dots, v_n$.

20 Prove that if $u, v \in V$, then $\big|\, \|u\| - \|v\| \,\big| \le \|u - v\|$.

The inequality above is called the ***reverse triangle inequality****. For the reverse triangle inequality when $V = \mathbf{C}$, see Exercise 2 in Chapter 4.*

21 Suppose $u, v \in V$ are such that

$$\|u\| = 3, \quad \|u + v\| = 4, \quad \|u - v\| = 6.$$

What number does $\|v\|$ equal?

22 Show that if $u, v \in V$, then

$$\|u + v\| \, \|u - v\| \le \|u\|^2 + \|v\|^2.$$

23 Suppose $v_1, \dots, v_m \in V$ are such that $\|v_k\| \le 1$ for each $k = 1, \dots, m$. Show that there exist $a_1, \dots, a_m \in \{1, -1\}$ such that

$$\|a_1 v_1 + \cdots + a_m v_m\| \le \sqrt{m}.$$

16 הזווית בין שני וקטורים (כשחושבים עליהם כחיצים שמוצאם בראשית הצירים) ב־ $\mathbf{R}^2$ או $\mathbf{R}^3$ ניתנת להגדרה באופן גאומטרי. אך הגאומטריה אינה ברורה באותה מידה ב־ $\mathbf{R}^n$ כאשר $n > 3$. לפיכך הזווית בין שני וקטורים השונים מאפס $x, y \in \mathbf{R}^n$ מוגדרת כ־

$$\arccos \frac{\langle x, y \rangle}{\|x\| \, \|y\|},$$

כאשר המוטיבציה להגדרה זו נובעת מתרגיל 15. הסבירו מדוע אנו זקוקים לאי־שוויון קושי־שוורץ כדי להראות שהגדרה זו הגיונית.

17 הוכיחו כי

$$\left(\sum_{k=1}^{n} a_k b_k\right)^2 \le \left(\sum_{k=1}^{n} k a_k^2\right)\left(\sum_{k=1}^{n} \frac{b_k^2}{k}\right)$$

מתקיים לכל המספרים הממשיים $a_1, \dots, a_n$ ו־ $b_1, \dots, b_n$.

18 א. נניח ש־ $f: [1, \infty) \to [0, \infty)$ רציפה. הראו שמתקיים

$$\left(\int_1^\infty f\right)^2 \le \int_1^\infty x^2 (f(x))^2 \, dx.$$

ב. עבור אילו פונקציות רציפות $f: [1, \infty) \to [0, \infty)$ האי־שוויון ב־ (א) הוא שוויון ששני אגפיו סופיים?

19 נניח ש־ $v_1, \dots, v_n$ בסיס של V ו־ $T \in \mathcal{L}(V)$. הוכיחו כי אם λ ערך עצמי של T, אזי

$$|\lambda|^2 \le \sum_{j=1}^{n} \sum_{k=1}^{n} |\mathcal{M}(T)_{j,k}|^2,$$

כאשר $\mathcal{M}(T)_{j,k}$ מסמן את האיבר בשורה j, עמודה k של המטריצה של T ביחס לבסיס $v_1, \dots, v_n$.

20 הוכיחו שאם $u, v \in V$, אזי $\big| \|u\| - \|v\| \big| \le \|u - v\|$.

האי־שוויון לעיל נקרא **אי־שוויון המשולש ההפוך**. עבור אי־שוויון המשולש ההפוך כאשר $V = \mathbf{C}$, ראו תרגיל 2 בפרק ד.

21 נניח ש־ $u, v \in V$ וקטורים כך ש־

$$\|u\| = 3, \quad \|u + v\| = 4, \quad \|u - v\| = 6.$$

לאיזה מספר שווה $\|v\|$?

22 הראו שאם $u, v \in V$, אזי

$$\|u + v\| \, \|u - v\| \le \|u\|^2 + \|v\|^2.$$

23 נניח ש־ $v_1, \dots, v_m \in V$ וקטורים כך ש־ $\|v_k\| \le 1$ לכל $k = 1, \dots, m$. הראו שקיימים $a_1, \dots, a_m \in \{1, -1\}$ כך ש־

$$\|a_1 v_1 + \cdots + a_m v_m\| \le \sqrt{m}.$$

אלגברה ליניארית בדרך הישר, מהדורה רביעית, מחבר: שלדון אקסלר
Linear Algebra Done Right, fourth edition, by Sheldon Axler

24 Prove or give a counterexample: If $\|\cdot\|$ is the norm associated with an inner product on $\mathbf{R}^2$, then there exists $(x, y) \in \mathbf{R}^2$ such that $\|(x, y)\| \neq \max\{|x|, |y|\}$.

25 Suppose $p > 0$. Prove that there is an inner product on $\mathbf{R}^2$ such that the associated norm is given by

$$\|(x, y)\| = (|x|^p + |y|^p)^{1/p}$$

for all $(x, y) \in \mathbf{R}^2$ if and only if $p = 2$.

26 Suppose V is a real inner product space. Prove that

$$\langle u, v\rangle = \frac{\|u + v\|^2 - \|u - v\|^2}{4}$$

for all $u, v \in V$.

27 Suppose V is a complex inner product space. Prove that

$$\langle u, v\rangle = \frac{\|u + v\|^2 - \|u - v\|^2 + \|u + iv\|^2 i - \|u - iv\|^2 i}{4}$$

for all $u, v \in V$.

28 A norm on a vector space U is a function

$$\|\cdot\| \colon U \to [0, \infty)$$

such that $\|u\| = 0$ if and only if $u = 0$, $\|\alpha u\| = |\alpha|\|u\|$ for all $\alpha \in \mathbf{F}$ and all $u \in U$, and $\|u + v\| \leq \|u\| + \|v\|$ for all $u, v \in U$. Prove that a norm satisfying the parallelogram equality comes from an inner product (in other words, show that if $\|\cdot\|$ is a norm on U satisfying the parallelogram equality, then there is an inner product $\langle\cdot, \cdot\rangle$ on U such that $\|u\| = \langle u, u\rangle^{1/2}$ for all $u \in U$).

29 Suppose $V_1, \ldots, V_m$ are inner product spaces. Show that the equation

$$\big\langle (u_1, \ldots, u_m), (v_1, \ldots, v_m)\big\rangle = \langle u_1, v_1\rangle + \cdots + \langle u_m, v_m\rangle$$

defines an inner product on $V_1 \times \cdots \times V_m$.

In the expression above on the right, for each $k = 1, \ldots, m$, the inner product $\langle u_k, v_k\rangle$ denotes the inner product on V_k. Each of the spaces $V_1, \ldots, V_m$ may have a different inner product, even though the same notation is used here.

30 Suppose V is a real inner product space. For $u, v, w, x \in V$, define

$$\langle u + iv, w + ix\rangle_{\mathbf{C}} = \langle u, w\rangle + \langle v, x\rangle + (\langle v, w\rangle - \langle u, x\rangle)i.$$

(a) Show that $\langle\cdot, \cdot\rangle_{\mathbf{C}}$ makes $V_{\mathbf{C}}$ into a complex inner product space.
(b) Show that if $u, v \in V$, then

$$\langle u, v\rangle_{\mathbf{C}} = \langle u, v\rangle \quad \text{and} \quad \|u + iv\|_{\mathbf{C}}^2 = \|u\|^2 + \|v\|^2.$$

See Exercise 8 in Section 1B for the definition of the complexification $V_{\mathbf{C}}$.

24 הוכיחו או הביאו דוגמה נגדית: אם $\|\cdot\|$ היא הנורמה המשויכת למכפלה פנימית מעל $\mathbf{R}^2$, אזי קיים $(x,y) \in \mathbf{R}^2$ כך ש־ $\|(x,y)\| \neq \max\{|x|,|y|\}$.

25 נניח ש־ $p > 0$. הוכיחו שקיימת מכפלה פנימית מעל $\mathbf{R}^2$ כך שהנורמה המשויכת אליה נתונה על ידי

$$\|(x,y)\| = \left(|x|^p + |y|^p\right)^{1/p}$$

לכל $(x,y) \in \mathbf{R}^2$ אם־ורק־אם $p = 2$.

26 נניח ש־ V מרחב מכפלה פנימית ממשי. הוכיחו ש־

$$\langle u, v\rangle = \frac{\|u+v\|^2 - \|u-v\|^2}{4}$$

לכל $u, v \in V$.

27 נניח ש־ V מרחב מכפלה פנימית מרוכב. הוכיחו ש־

$$\langle u, v\rangle = \frac{\|u+v\|^2 - \|u-v\|^2 + \|u+iv\|^2 i - \|u-iv\|^2 i}{4}$$

לכל $u, v \in V$.

28 נורמה מעל מרחב וקטורי U היא פונקציה

$$\|\cdot\| \colon U \to [0,\infty)$$

כך ש־ $\|u\| = 0$ אם־ורק־אם $u = 0$, $\|\alpha u\| = |\alpha|\|u\|$ לכל $\alpha \in \mathbf{F}$ ולכל $u \in U$, ו־ $\|u\| + \|v\| \leq \|u+v\|$ לכל $u, v \in U$. הוכיחו שנורמה המקיימת את כלל המקבילית מוצאה במכפלה פנימית (במילים אחרות, הראו שאם $\|\cdot\|$ נורמה מעל U המקיימת את כלל המקבילית, אזי קיימת מכפלה פנימית $\langle\cdot,\cdot\rangle$ מעל U כך ש־ $\|u\| = \langle u, u\rangle^{1/2}$ לכל $u \in U$).

29 נניח ש־ $V_1, \ldots, V_m$ מרחבי מכפלה פנימית. הראו כי המשוואה

$$\big\langle (u_1, \ldots, u_m), (v_1, \ldots, v_m)\big\rangle = \langle u_1, v_1\rangle + \cdots + \langle u_m, v_m\rangle$$

מגדירה מכפלה פנימית מעל $V_1 \times \cdots \times V_m$.

בביטוי לעיל באגף הימני, לכל $k = 1, \ldots, m$, המכפלה הפנימית $\langle u_k, v_k\rangle$ מציינת את המכפלה הפנימית מעל V_k. ייתכן שלכל אחד מהמרחבים $V_1, \ldots, V_m$ יש מכפלה פנימית שונה, למרות השימוש בסימון זהה לכל המכפלות.

30 נניח ש־ V מרחב מכפלה פנימית ממשי. עבור $u, v, w, x \in V$, הגדירו

$$\langle u+iv, w+ix\rangle_{\mathbf{C}} = \langle u, w\rangle + \langle v, x\rangle + (\langle v, w\rangle - \langle u, x\rangle)i.$$

א. הראו ש־ $\langle\cdot,\cdot\rangle_{\mathbf{C}}$ אכן הופכת את $V_{\mathbf{C}}$ למרחב מכפלה פנימית מרוכב.

ב. הראו שאם $u, v \in V$, אזי

$$\langle u, v\rangle_{\mathbf{C}} = \langle u, v\rangle \quad \text{וגם} \quad \|u+iv\|_{\mathbf{C}}^2 = \|u\|^2 + \|v\|^2.$$

ראו תרגיל 8 ביחידה אב עבור הגדרת מירכוב $V_{\mathbf{C}}$.

31 Suppose $u, v, w \in V$. Prove that

$$\|w - \tfrac{1}{2}(u+v)\|^2 = \frac{\|w-u\|^2 + \|w-v\|^2}{2} - \frac{\|u-v\|^2}{4}.$$

32 Suppose that E is a subset of V with the property that $u, v \in E$ implies $\frac{1}{2}(u+v) \in E$. Let $w \in V$. Show that there is at most one point in E that is closest to w. In other words, show that there is at most one $u \in E$ such that

$$\|w - u\| \le \|w - x\|$$

for all $x \in E$.

33 Suppose f, g are differentiable functions from $\mathbf{R}$ to $\mathbf{R}^n$.

(a) Show that

$$\langle f(t), g(t) \rangle' = \langle f'(t), g(t) \rangle + \langle f(t), g'(t) \rangle.$$

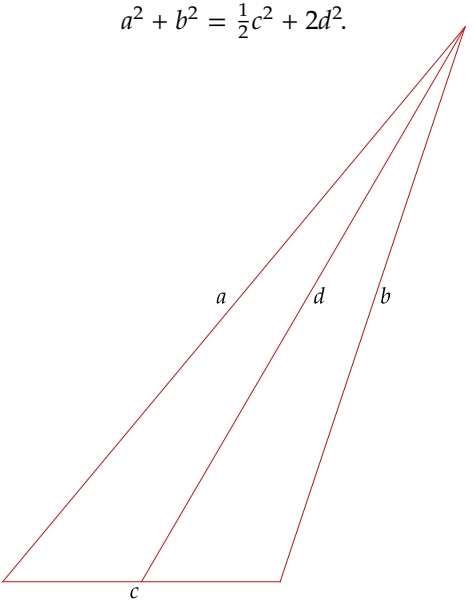
(b) Suppose c is a positive number and $\|f(t)\| = c$ for every $t \in \mathbf{R}$. Show that $\langle f'(t), f(t) \rangle = 0$ for every $t \in \mathbf{R}$.

(c) Interpret the result in (b) geometrically in terms of the tangent vector to a curve lying on a sphere in $\mathbf{R}^n$ centered at the origin.

A function $f : \mathbf{R} \to \mathbf{R}^n$ is called differentiable if there exist differentiable functions $f_1, \dots, f_n$ from $\mathbf{R}$ to $\mathbf{R}$ such that $f(t) = (f_1(t), \dots, f_n(t))$ for each $t \in \mathbf{R}$. Furthermore, for each $t \in \mathbf{R}$, the derivative $f'(t) \in \mathbf{R}^n$ is defined by $f'(t) = (f_1'(t), \dots, f_n'(t))$.

34 Use inner products to prove Apollonius's identity: In a triangle with sides of length a, b, and c, let d be the length of the line segment from the midpoint of the side of length c to the opposite vertex. Then

$$a^2 + b^2 = \tfrac{1}{2}c^2 + 2d^2.$$



31 נניח ש־ $u, v, w \in V$. הוכיחו ש־

$$\|w - \tfrac{1}{2}(u+v)\|^2 = \frac{\|w-u\|^2 + \|w-v\|^2}{2} - \frac{\|u-v\|^2}{4}.$$

32 נניח ש־ E תת־קבוצה של V עם התכונה ש־ $u, v \in E$ גורר ש־ $\frac{1}{2}(u+v) \in E$. יהי $w \in V$. הראו שקיימת לכל היותר נקודה אחת ב־ E הקרובה ביותר ל־ w. במילים אחרות, הראו שיש לכל היותר $u \in E$ אחד כך ש־

$$\|w - u\| \le \|w - x\|$$

לכל $x \in E$.

33 נניח ש־ f, g פונקציות גזירות מ־ $\mathbf{R}$ אל $\mathbf{R}^n$.

א. הראו ש־

$$\big\langle f(t), g(t) \big\rangle' = \big\langle f'(t), g(t) \big\rangle + \big\langle f(t), g'(t) \big\rangle.$$

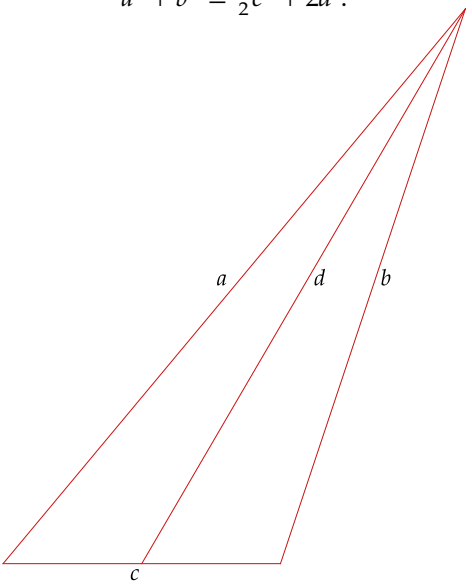
ב. נניח ש־ c מספר חיובי ו־ $\|f(t)\| = c$ לכל $t \in \mathbf{R}$. הראו ש־ $\big\langle f'(t), f(t) \big\rangle = 0$ לכל $t \in \mathbf{R}$.

ג. העניקו פרשנות גאומטרית לתוצאה בסעיף (ב) במונחים של הווקטור המשיק לעקום המונח על ספֵּירָה ב־ $\mathbf{R}^n$ שמרכזה בראשית הצירים.

פונקציה $f\colon \mathbf{R} \to \mathbf{R}^n$ נקראת גְּזִירָה אם קיימות פונקציות גזירות $f_1, \dots, f_n$ מ־ $\mathbf{R}$ אל $\mathbf{R}$ כך ש־ $f(t) = (f_1(t), \dots, f_n(t))$ לכל $t \in \mathbf{R}$. בנוסף, לכל $t \in \mathbf{R}$, הנגזרת $f'(t) \in \mathbf{R}^n$ מוגדרת על ידי $f'(t) = (f_1'(t), \dots, f_n'(t))$.

34 השתמשו במכפלה פנימית כדי להוכיח את משפט התיכון (נקרא גם משפט אפולוניוס): במשולש שאורך צלעותיו a, b, ו־ c, יהי d אורך הקטע מנקודת האמצע של הצלע שאורכה c אל הקודקוד שממול. אזי

$$a^2 + b^2 = \tfrac{1}{2}c^2 + 2d^2.$$



אלגברה ליניארית בדרך הישר, מהדורה רביעית, מחבר: שלדון אקסלר
Linear Algebra Done Right, fourth edition, by Sheldon Axler

35 Fix a positive integer n. The *Laplacian* Δp of a twice differentiable real-valued function p on $\mathbf{R}^n$ is the function on $\mathbf{R}^n$ defined by

$$\Delta p = \frac{\partial^2 p}{\partial x_1^2} + \cdots + \frac{\partial^2 p}{\partial x_n^2}.$$

The function p is called *harmonic* if $\Delta p = 0$.

A *polynomial* on $\mathbf{R}^n$ is a linear combination (with coefficients in $\mathbf{R}$) of functions of the form $x_1^{m_1} \cdots x_n^{m_n}$, where $m_1, \ldots, m_n$ are nonnegative integers.

Suppose q is a polynomial on $\mathbf{R}^n$. Prove that there exists a harmonic polynomial p on $\mathbf{R}^n$ such that $p(x) = q(x)$ for every $x \in \mathbf{R}^n$ with $\|x\| = 1$.

The only fact about harmonic functions that you need for this exercise is that if p is a harmonic function on $\mathbf{R}^n$ and $p(x) = 0$ for all $x \in \mathbf{R}^n$ with $\|x\| = 1$, then $p = 0$.

Hint: A reasonable guess is that the desired harmonic polynomial p is of the form $q + (1 - \|x\|^2)r$ for some polynomial r. Prove that there is a polynomial r on $\mathbf{R}^n$ such that $q + (1 - \|x\|^2)r$ is harmonic by defining an operator T on a suitable vector space by

$$Tr = \Delta((1 - \|x\|^2)r)$$

and then showing that T is injective and hence surjective.

In realms of numbers, where the secrets lie,
A noble truth emerges from the deep,
Cauchy and Schwarz, their wisdom they apply,
An inequality for all to keep.

Two vectors, by this bond, are intertwined,
As inner products weave a gilded thread,
Their magnitude, by providence, confined,
A bound to which their destiny is wed.

Though shadows fall, and twilight dims the day,
This inequality will stand the test,
To guide us in our quest, to light the way,
And in its truth, our understanding rest.

So sing, ye muses, of this noble feat,
Cauchy–Schwarz, the bound that none can beat.

—written by ChatGPT with input *Shakespearean sonnet on Cauchy–Schwarz inequality*

35 נְקַבֵּעַ מספר שלם חיובי n. הלפלסיאן Δp של פונקציה ממשית הגזירה פעמיים p מעל $\mathbf{R}^n$ הוא הפונקציה מעל $\mathbf{R}^n$ המוגדרת על ידי

$$\Delta p = \frac{\partial^2 p}{\partial x_1^2} + \cdots + \frac{\partial^2 p}{\partial x_n^2}.$$

הפונקציה p נקראת הרמונית אם $\Delta p = 0$. פולינום מעל $\mathbf{R}^n$ הוא צירוף ליניארי (עם מקדמים ב־ $\mathbf{R}$) של פונקציות מהצורה $x_1^{m_1} \cdots x_n^{m_n}$, כאשר $m_1, \ldots, m_n$ מספרים שלמים אי־שליליים. נניח ש־ q פולינום מעל $\mathbf{R}^n$. הוכיחו שקיים פולינום הרמוני p מעל $\mathbf{R}^n$ כך ש־ $p(x) = q(x)$ לכל $x \in \mathbf{R}^n$ עם $\|x\| = 1$.

העובדה היחידה הנוגעת לפונקציות הרמוניות שעליכם לדעת לצורך תרגיל זה היא שאם p פונקציה הרמונית מעל $\mathbf{R}^n$ ו־ $p(x) = 0$ לכל $x \in \mathbf{R}^n$ עם $\|x\| = 1$, אזי $p = 0$.

רמז: ניחוש סביר יהיה כי הפולינום ההרמוני הרצוי p הוא בצורה $q + (1 - \|x\|^2)r$ עבור פולינום r כלשהו. הוכיחו שקיים פולינום r מעל $\mathbf{R}^n$ כך ש־ $q + (1 - \|x\|^2)r$ הרמוני באמצעות הגדרת אופרטור T מעל מרחב וקטורי מתאים על ידי

$$Tr = \Delta((1 - \|x\|^2)r)$$

והראו ש־ T חד־חד־ערכי ולכן על.

בְּמֶרְחָב שֶׁל מַכְפֶּלָה פְּנִימִית
שְׁנֵי וֶקְטוֹרִים נִפְגָּשִׁים בִּדְמָמָה,
וְעֶרְכָּהּ הַמֻּחְלָט שֶׁל הַתַּבְנִית
לֹא יַעֲלֶה עַל כָּפָל נוֹרְמָה בְּנוֹרְמָה.

הַהוֹכָחָה: פֵּרוּק אוֹרְתוֹגוֹנָלִי —
רְכִיב מַקְבִּיל וּרְכִיב שֶׁנִּצָּב;
מִשְׁפַּט פִּיתָגוֹרָס, צַעַד פּוֹרְמָלִי,
וְאִי־הַשְּׁוִיוֹן כְּבָר מְעֻצָּב.

וּמָתַי הַשְּׁוִיוֹן מִתְקַיֵּם?
רַק אִם זֶה כָּפָל סְקָלָרִי שֶׁל זֶה:
הָרְכִיב הַנִּצָּב אָז מִתְעַלֵּם,
וְכָל הַסִּפּוּר עַל קַו אֶחָד נֶחֱזֶה.

מִכָּאן — קוֹסִינוּס שֶׁל זָוִית מֻגְדֶּרֶת,
כִּי הַיַּחַס בֵּין מִינוּס אַחַת לְאַחַת נִשְׁאָר.

—קלוד (Fable 5)

קלט: Shakespearean sonnet on Cauchy–Schwarz inequality in Hebrew

6B Orthonormal Bases

Orthonormal Lists and the Gram–Schmidt Procedure

6.22 definition: *orthonormal*

- A list of vectors is called *orthonormal* if each vector in the list has norm 1 and is orthogonal to all the other vectors in the list.
- In other words, a list $e_1, \dots, e_m$ of vectors in V is orthonormal if

$$\langle e_j, e_k \rangle = \begin{cases} 1 & \text{if } j = k, \\ 0 & \text{if } j \neq k \end{cases}$$

for all $j, k \in \{1, \dots, m\}$.

6.23 example: *orthonormal lists*

(a) The standard basis of $\mathbf{F}^n$ is an orthonormal list.

(b) $\left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right), \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right)$ is an orthonormal list in $\mathbf{F}^3$.

(c) $\left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right), \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right), \left(\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, -\frac{2}{\sqrt{6}}\right)$ is an orthonormal list in $\mathbf{F}^3$.

(d) Suppose n is a positive integer. Then, as Exercise 4 asks you to verify,

$$\frac{1}{\sqrt{2\pi}}, \frac{\cos x}{\sqrt{\pi}}, \frac{\cos 2x}{\sqrt{\pi}}, \dots, \frac{\cos nx}{\sqrt{\pi}}, \frac{\sin x}{\sqrt{\pi}}, \frac{\sin 2x}{\sqrt{\pi}}, \dots, \frac{\sin nx}{\sqrt{\pi}}$$

is an orthonormal list of vectors in $C[-\pi, \pi]$, the vector space of continuous real-valued functions on $[-\pi, \pi]$ with inner product

$$\langle f, g \rangle = \int_{-\pi}^{\pi} fg.$$

The orthonormal list above is often used for modeling periodic phenomena, such as tides.

(e) Suppose we make $\mathcal{P}_2(\mathbf{R})$ into an inner product space using the inner product given by

$$\langle p, q \rangle = \int_{-1}^{1} pq$$

for all $p, q \in \mathcal{P}_2(\mathbf{R})$. The standard basis $1, x, x^2$ of $\mathcal{P}_2(\mathbf{R})$ is not an orthonormal list because the vectors in that list do not have norm 1. Dividing each vector by its norm gives the list $1/\sqrt{2}, \sqrt{3/2}x, \sqrt{5/2}x^2$, in which each vector has norm 1, and the second vector is orthogonal to the first and third vectors. However, the first and third vectors are not orthogonal. Thus this is not an orthonormal list. Soon we will see how to construct an orthonormal list from the standard basis $1, x, x^2$ (see Example 6.34).

7ו בסיסים אורתונורמליים

רשימות אורתונורמליות ותהליך גרם־שמידט

6.22 הגדרה: אורתונורמלי

- רשימת וקטורים נקראת אורתוגונורמלית אם הנורמה של כל וקטור ברשימה היא 1 וכל וקטור אורתוגונלי לכל שאר הווקטורים ברשימה.
- במילים אחרות, רשימת וקטורים $e_1, \dots, e_m$ ב־ V היא אורתונורמלית אם

$$\langle e_j, e_k \rangle = \begin{cases} 1 & \text{if } j = k, \\ 0 & \text{if } j \neq k \end{cases}$$

לכל $j, k \in \{1, \dots, m\}$.

6.23 דוגמה: רשימות אורתונורמליות

א. הבסיס הסטנדרטי של $\mathbf{F}^n$ הוא רשימה אורתונורמלית.

ב. $\left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right), \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right)$ היא רשימה אורתונורמלית ב־ $\mathbf{F}^3$.

ג. $\left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right), \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right), \left(\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, -\frac{2}{\sqrt{6}}\right)$ היא רשימה אורתונורמלית ב־ $\mathbf{F}^3$.

ד. נניח ש־ n מספר שלם חיובי. אזי כפי שתרגיל 4 מבקש ממכם לוודא,

$$\frac{1}{\sqrt{2\pi}}, \frac{\cos x}{\sqrt{\pi}}, \frac{\cos 2x}{\sqrt{\pi}}, \dots, \frac{\cos nx}{\sqrt{\pi}}, \frac{\sin x}{\sqrt{\pi}}, \frac{\sin 2x}{\sqrt{\pi}}, \dots, \frac{\sin nx}{\sqrt{\pi}}$$

היא רשימה אורתונורמלית של וקטורים ב־ $C[-\pi, \pi]$, המרחב הווקטורי של פונקציות רציפות ממשיות מעל $[-\pi, \pi]$ עם המכפלה הפנימית

$$\langle f, g \rangle = \int_{-\pi}^{\pi} fg.$$

השימוש ברשימה אורתונורמלית זו נפוץ בבניית מודלים לתופעות מחזוריות דוגמת גאות ושפל.

ה. נניח שאנו הופכים את $\mathcal{P}_2(\mathbf{R})$ למרחב מכפלה פנימית בשימוש במכפלה הפנימית הנתונה על ידי

$$\langle p, q \rangle = \int_{-1}^{1} pq$$

לכל $p, q \in \mathcal{P}_2(\mathbf{R})$. הבסיס הסטנדרטי $1, x, x^2$ של $\mathcal{P}_2(\mathbf{R})$ אינו אורתונורמלי, שכן הנורמה של הווקטורים ברשימה זו שונה מ־ 1. כאשר נחלק כל וקטור בנורמה שלו נקבל את הרשימה $1/\sqrt{2}, \sqrt{3/2}x, \sqrt{5/2}x^2$, שבה הנורמה של כל וקטור היא 1, והווקטור השני אורתוגונלי לווקטורים הראשון והשלישי. אולם, הווקטורים הראשון והשלישי אינם אורתוגונליים. לפיכך רשימה זו אינה אורתונורמלית. בקרוב נראה כיצד לבנות רשימה אורתונורמלית מתוך הבסיס הסטנדרטי $1, x, x^2$ (ראו דוגמה 6.34).

Orthonormal lists are particularly easy to work with, as illustrated by the next result.

6.24 *norm of an orthonormal linear combination*

Suppose $e_1, \dots, e_m$ is an orthonormal list of vectors in V. Then

$$\|a_1 e_1 + \cdots + a_m e_m\|^2 = |a_1|^2 + \cdots + |a_m|^2$$

for all $a_1, \dots, a_m \in \mathbf{F}$.

Proof Because each e_k has norm 1, this follows from repeated applications of the Pythagorean theorem (6.12). ∎

The result above has the following important corollary.

6.25 *orthonormal lists are linearly independent*

Every orthonormal list of vectors is linearly independent.

Proof Suppose $e_1, \dots, e_m$ is an orthonormal list of vectors in V and $a_1, \dots, a_m \in \mathbf{F}$ are such that

$$a_1 e_1 + \cdots + a_m e_m = 0.$$

Then $|a_1|^2 + \cdots + |a_m|^2 = 0$ (by 6.24), which means that all the a_k's are 0. Thus $e_1, \dots, e_m$ is linearly independent. ∎

Now we come to an important inequality.

6.26 *Bessel's inequality*

Suppose $e_1, \dots, e_m$ is an orthonormal list of vectors in V. If $v \in V$ then

$$|\langle v, e_1 \rangle|^2 + \cdots + |\langle v, e_m \rangle|^2 \le \|v\|^2.$$

Proof Suppose $v \in V$. Then

$$v = \underbrace{\langle v, e_1 \rangle e_1 + \cdots + \langle v, e_m \rangle e_m}_{u} + \underbrace{v - \langle v, e_1 \rangle e_1 - \cdots - \langle v, e_m \rangle e_m}_{w}.$$

Let u and w be defined as in the equation above. If $k \in \{1, \dots, m\}$, then $\langle w, e_k \rangle = \langle v, e_k \rangle - \langle v, e_k \rangle \langle e_k, e_k \rangle = 0$. This implies that $\langle w, u \rangle = 0$. The Pythagorean theorem now implies that

$$\begin{aligned} \|v\|^2 &= \|u\|^2 + \|w\|^2 \\ &\ge \|u\|^2 \\ &= |\langle v, e_1 \rangle|^2 + \cdots + |\langle v, e_m \rangle|^2, \end{aligned}$$

where the last line comes from 6.24. ∎

רשימות אורתונורמליות הן נוחות לעבודה באופן מיוחד, כפי שממחישה המסקנה הבאה.

6.24 נורמה של צירוף ליניארי אורתונורמלי

נניח ש־ $e_1, \dots, e_m$ רשימת וקטורים אורתונורמלית ב־ V. אזי
$$\|a_1 e_1 + \cdots + a_m e_m\|^2 = |a_1|^2 + \cdots + |a_m|^2$$
לכל $a_1, \dots, a_m \in \mathbf{F}$.

הוכחה מאחר שהנורמה של כל וקטור e_k היא 1, מסקנה זו נובעת מיישום חוזר ונשנה של משפט פיתגורס (6.12). ■

ממסקנה זו נובעת התוצאה הישירה החשובה הבאה.

6.25 רשימות אורתונורמליות בלתי תלויות ליניארית

כל רשימה אורתונורמלית של וקטורים היא בלתי תלויה ליניארית.

הוכחה נניח ש־ $e_1, \dots, e_m$ רשימה אורתונורמלית של וקטורים ב־ V ו־ $a_1, \dots, a_m \in \mathbf{F}$ כך ש־
$$a_1 e_1 + \cdots + a_m e_m = 0.$$
אזי $|a_1|^2 + \cdots + |a_m|^2 = 0$ (לפי 6.24), משמעות הדבר היא שכל ה־ a_k־ים הם 0. לפיכך $e_1, \dots, e_m$ בלתי תלויה ליניארית. ■

כעת אנו מגיעים לאי־שוויון בעל חשיבות רבה.

6.26 אי־שוויון בֶּסֶל

נניח ש־ $e_1, \dots, e_m$ רשימה אורתונורמלית של וקטורים ב־ V. אם $v \in V$ אזי
$$|\langle v, e_1 \rangle|^2 + \cdots + |\langle v, e_m \rangle|^2 \le \|v\|^2.$$

הוכחה נניח ש־ $v \in V$. אזי
$$v = \underbrace{\langle v, e_1 \rangle e_1 + \cdots + \langle v, e_m \rangle e_m}_{u} + \underbrace{v - \langle v, e_1 \rangle e_1 - \cdots - \langle v, e_m \rangle e_m}_{w}.$$
יהיו u ו־ w מוגדרים באופן המוצג במשוואה לעיל. אם $k \in \{1, \dots, m\}$, אזי $\langle w, e_k \rangle = \langle v, e_k \rangle - \langle v, e_k \rangle \langle e_k, e_k \rangle = 0$. מכאן נובע ש־ $\langle w, u \rangle = 0$. כעת נסיק כי על פי משפט פיתגורס
$$\begin{aligned} \|v\|^2 &= \|u\|^2 + \|w\|^2 \\ &\ge \|u\|^2 \\ &= |\langle v, e_1 \rangle|^2 + \cdots + |\langle v, e_m \rangle|^2, \end{aligned}$$
כאשר השורה האחרונה נובעת מ־ 6.24. ■

The next definition introduces one of the most useful concepts in the study of inner product spaces.

6.27 definition: *orthonormal basis*

An *orthonormal basis* of V is an orthonormal list of vectors in V that is also a basis of V.

For example, the standard basis is an orthonormal basis of $\mathbf{F}^n$.

6.28 *orthonormal lists of the right length are orthonormal bases*

Suppose V is finite-dimensional. Then every orthonormal list of vectors in V of length $\dim V$ is an orthonormal basis of V.

Proof By 6.25, every orthonormal list of vectors in V is linearly independent. Thus every such list of the right length is a basis—see 2.38. ∎

6.29 example: *an orthonormal basis of* $\mathbf{F}^4$

As mentioned above, the standard basis is an orthonormal basis of $\mathbf{F}^4$. We now show that

$$\left(\tfrac{1}{2},\tfrac{1}{2},\tfrac{1}{2},\tfrac{1}{2}\right),\left(\tfrac{1}{2},\tfrac{1}{2},-\tfrac{1}{2},-\tfrac{1}{2}\right),\left(\tfrac{1}{2},-\tfrac{1}{2},-\tfrac{1}{2},\tfrac{1}{2}\right),\left(-\tfrac{1}{2},\tfrac{1}{2},-\tfrac{1}{2},\tfrac{1}{2}\right)$$

is also an orthonormal basis of $\mathbf{F}^4$.

We have

$$\left\|\left(\tfrac{1}{2},\tfrac{1}{2},\tfrac{1}{2},\tfrac{1}{2}\right)\right\| = \sqrt{\tfrac{1}{2^2}+\tfrac{1}{2^2}+\tfrac{1}{2^2}+\tfrac{1}{2^2}} = 1.$$

Similarly, the other three vectors in the list above also have norm 1.

Note that

$$\left\langle\left(\tfrac{1}{2},\tfrac{1}{2},\tfrac{1}{2},\tfrac{1}{2}\right),\left(\tfrac{1}{2},\tfrac{1}{2},-\tfrac{1}{2},-\tfrac{1}{2}\right)\right\rangle = \tfrac{1}{2}\cdot\tfrac{1}{2}+\tfrac{1}{2}\cdot\tfrac{1}{2}+\tfrac{1}{2}\cdot\left(-\tfrac{1}{2}\right)+\tfrac{1}{2}\cdot\left(-\tfrac{1}{2}\right) = 0.$$

Similarly, the inner product of any two distinct vectors in the list above also equals 0.

Thus the list above is orthonormal. Because we have an orthonormal list of length four in the four-dimensional vector space $\mathbf{F}^4$, this list is an orthonormal basis of $\mathbf{F}^4$ (by 6.28).

In general, given a basis $e_1,\dots,e_n$ of V and a vector $v \in V$, we know that there is some choice of scalars $a_1,\dots,a_n \in \mathbf{F}$ such that

$$v = a_1e_1 + \cdots + a_ne_n.$$

Computing the numbers $a_1,\dots,a_n$ that satisfy the equation above can be a long computation for an arbitrary basis of V. The next result shows, however, that this is easy for an orthonormal basis—just take $a_k = \langle v, e_k\rangle$.

ההגדרה הבאה מציגה את אחד המושגים השימושיים ביותר בחקר מרחבי מכפלה פנימית.

6.27 הגדרה: בסיס אורתונורמלי

בסיס אורתונורמלי של V הוא רשימה אורתונורמלית של וקטורים ב־ V שהיא גם בסיס של V.

לדוגמה, הבסיס הסטנדרטי הוא בסיס אורתונורמלי של $\mathbf{F}^n$.

6.28 רשימות אורתונורמליות באורך המתאים הן בסיסים אורתונורמליים

נניח ש־ V מממד סופי. אזי כל רשימה אורתונורמלית של וקטורים ב־ V שאורכה $\dim V$ היא בסיס אורתונורמלי של V.

הוכחה לפי 6.25, כל רשימה אורתונורמלית של וקטורים ב־ V היא בלתי תלויה ליניארית. לפיכך כל רשימה שכזו באורך המתאים היא בסיס—ראו 2.38. ■

6.29 דוגמה: בסיס אורתונורמלי של $\mathbf{F}^4$

כפי שצוין לעיל, הבסיס הסטנדרטי הוא בסיס אורתונורמלי של $\mathbf{F}^4$. כעת נראה ש־

$$\left(\tfrac{1}{2},\tfrac{1}{2},\tfrac{1}{2},\tfrac{1}{2}\right),\left(\tfrac{1}{2},\tfrac{1}{2},-\tfrac{1}{2},-\tfrac{1}{2}\right),\left(\tfrac{1}{2},-\tfrac{1}{2},-\tfrac{1}{2},\tfrac{1}{2}\right),\left(-\tfrac{1}{2},\tfrac{1}{2},-\tfrac{1}{2},\tfrac{1}{2}\right)$$

גם הוא בסיס אורתונורמלי של $\mathbf{F}^4$.

מתקיים

$$\left\|\left(\tfrac{1}{2},\tfrac{1}{2},\tfrac{1}{2},\tfrac{1}{2}\right)\right\| = \sqrt{\tfrac{1}{2^2}+\tfrac{1}{2^2}+\tfrac{1}{2^2}+\tfrac{1}{2^2}} = 1.$$

באופן דומה, הנורמה של שלושת הווקטורים האחרים ברשימה היא 1.

נשים לב לכך ש־

$$\left\langle\left(\tfrac{1}{2},\tfrac{1}{2},\tfrac{1}{2},\tfrac{1}{2}\right),\left(\tfrac{1}{2},\tfrac{1}{2},-\tfrac{1}{2},-\tfrac{1}{2}\right)\right\rangle = \tfrac{1}{2}\cdot\tfrac{1}{2}+\tfrac{1}{2}\cdot\tfrac{1}{2}+\tfrac{1}{2}\cdot\left(-\tfrac{1}{2}\right)+\tfrac{1}{2}\cdot\left(-\tfrac{1}{2}\right) = 0.$$

באופן דומה, המכפלה הפנימית של כל שני וקטורים שונים ברשימה לעיל שווה 0.

לפיכך הרשימה לעיל אורתונורמלית. מכיוון שיש בידינו רשימה אורתונורמלית בת ארבעה איברים במרחב וקטורי ארבע־ממדי $\mathbf{F}^4$, רשימה זו היא בסיס אורתונורמלי של $\mathbf{F}^4$ (לפי 6.28).

באופן כללי, בהינתן בסיס $e_1,\dots,e_n$ של V ווקטור $v \in V$, אנו יודעים כי קיימת בחירה כלשהי של סקלרים $a_1,\dots,a_n \in \mathbf{F}$ כך ש־

$$v = a_1e_1 + \cdots + a_ne_n.$$

חישוב הערכים $a_1,\dots,a_n$ כך שיקיימו את המשוואה לעיל יכול להיות תהליך חישוב ארוך עבור בסיס שרירותי של V. אולם, המסקנה הבאה מראה כי עבור בסיס אורתונורמלי מדובר בפעולה קלה—כל שעלינו לעשות הוא לחשב $a_k = \langle v, e_k\rangle$.

Notice how the next result makes each inner product space of dimension n behave like $\mathbf{F}^n$, with the role of the coordinates of a vector in $\mathbf{F}^n$ played by $\langle v, e_1\rangle, \dots, \langle v, e_n\rangle$.

The formula below for $\|v\|$ is called Parseval's identity. It was published in 1799 in the context of Fourier series.

6.30 *writing a vector as a linear combination of an orthonormal basis*

Suppose $e_1, \dots, e_n$ is an orthonormal basis of V and $u, v \in V$. Then

(a) $v = \langle v, e_1\rangle e_1 + \cdots + \langle v, e_n\rangle e_n$;

(b) $\|v\|^2 = |\langle v, e_1\rangle|^2 + \cdots + |\langle v, e_n\rangle|^2$;

(c) $\langle u, v\rangle = \langle u, e_1\rangle\overline{\langle v, e_1\rangle} + \cdots + \langle u, e_n\rangle\overline{\langle v, e_n\rangle}$.

Proof Because $e_1, \dots, e_n$ is a basis of V, there exist scalars $a_1, \dots, a_n$ such that

$$v = a_1 e_1 + \cdots + a_n e_n.$$

Because $e_1, \dots, e_n$ is orthonormal, taking the inner product of both sides of this equation with e_k gives $\langle v, e_k\rangle = a_k$. Thus (a) holds.

Now (b) follows immediately from (a) and 6.24.

Take the inner product of u with each side of (a) and then get (c) by using conjugate linearity [6.6(d) and 6.6(e)] in the second slot of the inner product. ∎

6.31 example: *finding coefficients for a linear combination*

Suppose we want to write the vector $(1, 2, 4, 7) \in \mathbf{F}^4$ as a linear combination of the orthonormal basis

$$\left(\tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2}\right), \left(\tfrac{1}{2}, \tfrac{1}{2}, -\tfrac{1}{2}, -\tfrac{1}{2}\right), \left(\tfrac{1}{2}, -\tfrac{1}{2}, -\tfrac{1}{2}, \tfrac{1}{2}\right), \left(-\tfrac{1}{2}, \tfrac{1}{2}, -\tfrac{1}{2}, \tfrac{1}{2}\right)$$

of $\mathbf{F}^4$ from Example 6.29. Instead of solving a system of four linear equations in four unknowns, as typically would be required if we were working with a nonorthonormal basis, we simply evaluate four inner products and use 6.30(a), getting that $(1, 2, 4, 7)$ equals

$$7\left(\tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2}\right) - 4\left(\tfrac{1}{2}, \tfrac{1}{2}, -\tfrac{1}{2}, -\tfrac{1}{2}\right) + \left(\tfrac{1}{2}, -\tfrac{1}{2}, -\tfrac{1}{2}, \tfrac{1}{2}\right) + 2\left(-\tfrac{1}{2}, \tfrac{1}{2}, -\tfrac{1}{2}, \tfrac{1}{2}\right).$$

Now that we understand the usefulness of orthonormal bases, how do we go about finding them? For example, does $\mathcal{P}_m(\mathbf{R})$ with inner product as in 6.3(c) have an orthonormal basis? The next result will lead to answers to these questions.

The algorithm used in the next proof is called the *Gram–Schmidt procedure*. It gives a method for turning a linearly independent list into an orthonormal list with the same span as the original list.

Jørgen Gram (1850–1916) and Erhard Schmidt (1876–1959) popularized this algorithm that constructs orthonormal lists.

נשים לב כיצד המסקנה הבאה הופכת את התנהגות כל מרחב מכפלה פנימית n־ממדי לדומה לזו של $\mathbf{F}^n$, כאשר את תפקיד הקואורדינטות של וקטור ב־ $\mathbf{F}^n$ ממלאים $\langle v, e_1\rangle, \dots, \langle v, e_n\rangle$.

הנוסחה להלן עבור $\|v\|$ נקראת שוויון פָּרְסֶבַל. היא פורסמה בשנת 1799 בהקשר של טור פורייה.

6.30 ביטוי וקטור כצירוף ליניארי של בסיס אורתונורמלי

נניח ש־ $e_1, \dots, e_n$ בסיס אורתונורמלי של V ו־ $u, v \in V$. אזי

א. $v = \langle v, e_1\rangle e_1 + \cdots + \langle v, e_n\rangle e_n$,

ב. $\|v\|^2 = |\langle v, e_1\rangle|^2 + \cdots + |\langle v, e_n\rangle|^2$,

ג. $\langle u, v\rangle = \langle u, e_1\rangle\overline{\langle v, e_1\rangle} + \cdots + \langle u, e_n\rangle\overline{\langle v, e_n\rangle}$.

הוכחה מאחר ש־ $e_1, \dots, e_n$ בסיס של V, קיימים סקלרים $a_1, \dots, a_n$ כך ש־

$$v = a_1 e_1 + \cdots + a_n e_n.$$

מאחר ש־ $e_1, \dots, e_n$ בסיס אורתונורמלי, בחישוב המכפלה הפנימית של שני אגפי המשוואה עם e_k נקבל $\langle v, e_k\rangle = a_k$. לפיכך (א) מתקיים.

כעת (ב) נובע ישירות מ־ (א) ו־ 6.24.

נחשב את המכפלה הפנימית של u עם כל אגף של (א) ואז נקבל את (ג) בהפעלת הליניאריות של הצמוד על הרכיב השני של המכפלה הפנימית [6.6(ד) ו־ 6.6(ה)]. ■

6.31 דוגמה: מציאת מקדמים לצירוף ליניארי

נניח שעלינו לבטא את הווקטור $(1, 2, 4, 7) \in \mathbf{F}^4$ כצירוף ליניארי של הבסיס האורתונורמלי

$$\left(\tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2}\right), \left(\tfrac{1}{2}, \tfrac{1}{2}, -\tfrac{1}{2}, -\tfrac{1}{2}\right), \left(\tfrac{1}{2}, -\tfrac{1}{2}, -\tfrac{1}{2}, \tfrac{1}{2}\right), \left(-\tfrac{1}{2}, \tfrac{1}{2}, -\tfrac{1}{2}, \tfrac{1}{2}\right)$$

של $\mathbf{F}^4$ מדוגמה 6.29. במקום לפתור מערכת של ארבע משוואות ליניאריות בארבעה נעלמים, כפי שהיה עלינו לעשות אלמלא היינו עובדים עם בסיס אורתונורמלי, אנו פשוט מחשבים ארבע מכפלות פנימיות ומשתמשים ב־ 6.30(א) ונקבל כי $(1, 2, 4, 7)$ שווה ל־

$$7\left(\tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2}\right) - 4\left(\tfrac{1}{2}, \tfrac{1}{2}, -\tfrac{1}{2}, -\tfrac{1}{2}\right) + \left(\tfrac{1}{2}, -\tfrac{1}{2}, -\tfrac{1}{2}, \tfrac{1}{2}\right) + 2\left(-\tfrac{1}{2}, \tfrac{1}{2}, -\tfrac{1}{2}, \tfrac{1}{2}\right).$$

כעת כשאנו מבינים את התועלת של בסיסים אורתונורמליים, איך אנו מוצאים אותם? לדוגמה, האם יש ל־ $\mathcal{P}_m(\mathbf{R})$ בצירוף המכפלה הפנימית המופיעה ב־ 6.3(ג) בסיס אורתונורמלי? המסקנה הבאה תוביל אותנו אל התשובות לשאלות אלו.

יוֹרְגֶן פֶּדֶרְסֶן גְּרָם (1850–1916) ו־ אֶרְהַרְד שְׁמִידְט (1876–1959) תרמו לפופולריות של אלגוריתם זה לבניית רשימות אורתונורמליות.

האלגוריתם שבו משתמשים בהוכחה להלן נקרא תהליך גרם־שמידט. הוא מעניק שיטה להפיכת רשימה בלתי תלויה ליניארית לרשימה אורתונורמלית עם פריסה זהה לרשימה המקורית.

6.32 *Gram–Schmidt procedure*

Suppose $v_1, \dots, v_m$ is a linearly independent list of vectors in V. Let $f_1 = v_1$. For $k = 2, \dots, m$, define f_k inductively by

$$f_k = v_k - \frac{\langle v_k, f_1 \rangle}{\|f_1\|^2} f_1 - \cdots - \frac{\langle v_k, f_{k-1} \rangle}{\|f_{k-1}\|^2} f_{k-1}.$$

For each $k = 1, \dots, m$, let $e_k = \dfrac{f_k}{\|f_k\|}$. Then $e_1, \dots, e_m$ is an orthonormal list of vectors in V such that

$$\operatorname{span}(v_1, \dots, v_k) = \operatorname{span}(e_1, \dots, e_k)$$

for each $k = 1, \dots, m$.

Proof We will show by induction on k that the desired conclusion holds. To get started with $k = 1$, note that because $e_1 = f_1/\|f_1\|$, we have $\|e_1\| = 1$; also, $\operatorname{span}(v_1) = \operatorname{span}(e_1)$ because e_1 is a nonzero multiple of v_1.

Suppose $1 < k \le m$ and the list $e_1, \dots, e_{k-1}$ generated by 6.32 is an orthonormal list such that

6.33 $$\operatorname{span}(v_1, \dots, v_{k-1}) = \operatorname{span}(e_1, \dots, e_{k-1}).$$

Because $v_1, \dots, v_m$ is linearly independent, we have $v_k \notin \operatorname{span}(v_1, \dots, v_{k-1})$. Thus $v_k \notin \operatorname{span}(e_1, \dots, e_{k-1}) = \operatorname{span}(f_1, \dots, f_{k-1})$, which implies that $f_k \ne 0$. Hence we are not dividing by 0 in the definition of e_k given in 6.32. Dividing a vector by its norm produces a new vector with norm 1; thus $\|e_k\| = 1$.

Let $j \in \{1, \dots, k-1\}$. Then

$$\begin{aligned}
\langle e_k, e_j \rangle &= \frac{1}{\|f_k\| \, \|f_j\|} \langle f_k, f_j \rangle \\
&= \frac{1}{\|f_k\| \, \|f_j\|} \left\langle v_k - \frac{\langle v_k, f_1 \rangle}{\|f_1\|^2} f_1 - \cdots - \frac{\langle v_k, f_{k-1} \rangle}{\|f_{k-1}\|^2} f_{k-1}, f_j \right\rangle \\
&= \frac{1}{\|f_k\| \, \|f_j\|} (\langle v_k, f_j \rangle - \langle v_k, f_j \rangle) \\
&= 0.
\end{aligned}$$

Thus $e_1, \dots, e_k$ is an orthonormal list.

From the definition of e_k given in 6.32, we see that $v_k \in \operatorname{span}(e_1, \dots, e_k)$. Combining this information with 6.33 shows that

$$\operatorname{span}(v_1, \dots, v_k) \subseteq \operatorname{span}(e_1, \dots, e_k).$$

Both lists above are linearly independent (the v's by hypothesis, and the e's by orthonormality and 6.25). Thus both subspaces above have dimension k, and hence they are equal, completing the induction step and thus completing the proof. ∎

6.32 תהליך גרם־שמידט

נניח ש־ $v_1, \dots, v_m$ רשימה בלתי תלויה ליניארית של וקטורים ב־ V. יהי $f_1 = v_1$. עבור $k = 2, \dots, m$, נגדיר f_k באופן אינדוקטיבי על ידי

$$f_k = v_k - \frac{\langle v_k, f_1 \rangle}{\|f_1\|^2} f_1 - \dots - \frac{\langle v_k, f_{k-1} \rangle}{\|f_{k-1}\|^2} f_{k-1}.$$

לכל $k = 1, \dots, m$, יהי $e_k = \dfrac{f_k}{\|f_k\|}$. אזי $e_1, \dots, e_m$ רשימה אורתונורמלית של וקטורים ב־ V כך ש־

$$\operatorname{span}(v_1, \dots, v_k) = \operatorname{span}(e_1, \dots, e_k)$$

לכל $k = 1, \dots, m$.

הוכחה נראה באינדוקציה על k כי התוצאה הרצויה מתקיימת. נתחיל בבדיקה עבור $k = 1$: נשים לב כי מאחר ש־ $f_1 = f_1/\|f_1\|$, קיבלנו ש־ $\|e_1\| = 1$; בנוסף, $\operatorname{span}(v_1) = \operatorname{span}(e_1)$ שכן e_1 כפולה השונה מאפס של v_1.

נניח ש־ $1 < k \le m$ והרשימה $e_1, \dots, e_{k-1}$ שנוצרה על ידי 6.32 היא רשימה אורתונורמלית כך ש־

$$\operatorname{span}(v_1, \dots, v_{k-1}) = \operatorname{span}(e_1, \dots, e_{k-1}). \tag{6.33}$$

מאחר ש־ $v_1, \dots, v_m$ בלתי תלויה ליניארית, הרי ש־ $v_k \notin \operatorname{span}(v_1, \dots, v_{k-1})$. לפיכך $v_k \notin \operatorname{span}(e_1, \dots, e_{k-1}) = \operatorname{span}(f_1, \dots, f_{k-1})$, מה שגורר ש־ $f_k \neq 0$. לכן איננו מבצעים חלוקה ב־ 0 בהגדרת e_k הנתונה ב־ 6.32. חלוקת וקטור בנורמה שלו מפיקה וקטור חדש עם נורמה 1; לפיכך $\|e_k\| = 1$.

יהי $j \in \{1, \dots, k-1\}$. אזי

$$\begin{aligned}
\langle e_k, e_j \rangle &= \frac{1}{\|f_k\| \, \|f_j\|} \langle f_k, f_j \rangle \\
&= \frac{1}{\|f_k\| \, \|f_j\|} \left\langle v_k - \frac{\langle v_k, f_1 \rangle}{\|f_1\|^2} f_1 - \dots - \frac{\langle v_k, f_{k-1} \rangle}{\|f_{k-1}\|^2} f_{k-1}, f_j \right\rangle \\
&= \frac{1}{\|f_k\| \, \|f_j\|} (\langle v_k, f_j \rangle - \langle v_k, f_j \rangle) \\
&= 0.
\end{aligned}$$

לפיכך $e_1, \dots, e_k$ היא רשימה אורתונורמלית.

מהגדרת e_k הנתונה ב־ 6.32, אנו רואים ש־ $v_k \in \operatorname{span}(e_1, \dots, e_k)$. שילוב נתון זה עם 6.33 מראה לנו ש־

$$\operatorname{span}(v_1, \dots, v_k) \subseteq \operatorname{span}(e_1, \dots, e_k).$$

שתי הרשימות לעיל בלתי תלויות ליניארית (רשימת ה־ v־ים לפי ההנחה, ורשימת ה־ e־ים בזכות האורתונורמליות ו־ 6.25). לפיכך ממד שני תתי־המרחב לעיל הוא k, ולכן הם שווים, ובכך השלמנו את צעד האינדוקציה ולפיכך את ההוכחה. ■

6.34 example: *an orthonormal basis of* $\mathcal{P}_2(\mathbf{R})$

Suppose we make $\mathcal{P}_2(\mathbf{R})$ into an inner product space using the inner product given by

$$\langle p, q\rangle = \int_{-1}^{1} pq$$

for all $p, q \in \mathcal{P}_2(\mathbf{R})$. We know that $1, x, x^2$ is a basis of $\mathcal{P}_2(\mathbf{R})$, but it is not an orthonormal basis. We will find an orthonormal basis of $\mathcal{P}_2(\mathbf{R})$ by applying the Gram–Schmidt procedure with $v_1 = 1$, $v_2 = x$, and $v_3 = x^2$.

To get started, take $f_1 = v_1 = 1$. Thus $\|f_1\|^2 = \int_{-1}^{1} 1 = 2$. Hence the formula in 6.32 tells us that

$$f_2 = v_2 - \frac{\langle v_2, f_1\rangle}{\|f_1\|^2} f_1 = x - \frac{\langle x, 1\rangle}{\|f_1\|^2} = x,$$

where the last equality holds because $\langle x, 1\rangle = \int_{-1}^{1} t\, dt = 0$.

The formula above for f_2 implies that $\|f_2\|^2 = \int_{-1}^{1} t^2\, dt = \frac{2}{3}$. Now the formula in 6.32 tells us that

$$f_3 = v_3 - \frac{\langle v_3, f_1\rangle}{\|f_1\|^2} f_1 - \frac{\langle v_3, f_2\rangle}{\|f_2\|^2} f_2 = x^2 - \tfrac{1}{2}\langle x^2, 1\rangle - \tfrac{3}{2}\langle x^2, x\rangle x = x^2 - \tfrac{1}{3}.$$

The formula above for f_3 implies that

$$\|f_3\|^2 = \int_{-1}^{1} \left(t^2 - \tfrac{1}{3}\right)^2 dt = \int_{-1}^{1} \left(t^4 - \tfrac{2}{3}t^2 + \tfrac{1}{9}\right) dt = \tfrac{8}{45}.$$

Now dividing each of f_1, f_2, f_3 by its norm gives us the orthonormal list

$$\sqrt{\tfrac{1}{2}}, \sqrt{\tfrac{3}{2}}x, \sqrt{\tfrac{45}{8}}\left(x^2 - \tfrac{1}{3}\right).$$

The orthonormal list above has length three, which is the dimension of $\mathcal{P}_2(\mathbf{R})$. Hence this orthonormal list is an orthonormal basis of $\mathcal{P}_2(\mathbf{R})$ [by 6.28].

Now we can answer the question about the existence of orthonormal bases.

6.35 *existence of orthonormal basis*

Every finite-dimensional inner product space has an orthonormal basis.

Proof Suppose V is finite-dimensional. Choose a basis of V. Apply the Gram–Schmidt procedure (6.32) to it, producing an orthonormal list of length $\dim V$. By 6.28, this orthonormal list is an orthonormal basis of V. ∎

Sometimes we need to know not only that an orthonormal basis exists, but also that every orthonormal list can be extended to an orthonormal basis. In the next corollary, the Gram–Schmidt procedure shows that such an extension is always possible.

6.34 דוגמה: בסיס אורתונורמלי של $\mathcal{P}_2(\mathbf{R})$

נניח שאנו הופכים את $\mathcal{P}_2(\mathbf{R})$ למרחב מכפלה פנימית בשימוש במכפלה הפנימית הבאה

$$\langle p, q\rangle = \int_{-1}^{1} pq$$

לכל $p, q \in \mathcal{P}_2(\mathbf{R})$. אנו יודעים כי $1, x, x^2$ בסיס של $\mathcal{P}_2(\mathbf{R})$, אבל אינו בסיס אורתונורמלי. אנו נמצא בסיס אורתונורמלי של $\mathcal{P}_2(\mathbf{R})$ באמצעות יישום תהליך גרם־שמידט, עם $v_1 = 1$, $v_2 = x$, ו־ $v_3 = x^2$.

נתחיל בכך שניקח $f_1 = v_1 = 1$. לפיכך $\|f_1\|^2 = \int_{-1}^{1} 1 = 2$. לכן הנוסחה ב־ 6.32 אומרת לנו ש־

$$f_2 = v_2 - \frac{\langle v_2, f_1\rangle}{\|f_1\|^2} f_1 = x - \frac{\langle x, 1\rangle}{\|f_1\|^2} = x,$$

כאשר השוויון האחרון מתקיים מאחר ש־ $\langle x, 1\rangle = \int_{-1}^{1} t\, dt = 0$.

הנוסחה לעיל עבור f_2 גוררת ש־ $\|f_2\|^2 = \int_{-1}^{1} t^2\, dt = \frac{2}{3}$. כעת הנוסחה ב־ 6.32 אומרת לנו ש־

$$f_3 = v_3 - \frac{\langle v_3, f_1\rangle}{\|f_1\|^2} f_1 - \frac{\langle v_3, f_2\rangle}{\|f_2\|^2} f_2 = x^2 - \tfrac{1}{2}\langle x^2, 1\rangle - \tfrac{3}{2}\langle x^2, x\rangle x = x^2 - \tfrac{1}{3}.$$

הנוסחה לעיל עבור f_3 גוררת ש־

$$\|f_3\|^2 = \int_{-1}^{1} \left(t^2 - \tfrac{1}{3}\right)^2 dt = \int_{-1}^{1} \left(t^4 - \tfrac{2}{3}t^2 + \tfrac{1}{9}\right) dt = \tfrac{8}{45}.$$

כעת, חלוקת כל אחד מהווקטורים f_1, f_2, f_3 בנורמה שלו תיתן את הרשימה האורתונורמלית

$$\sqrt{\tfrac{1}{2}}, \sqrt{\tfrac{3}{2}}x, \sqrt{\tfrac{45}{8}}\left(x^2 - \tfrac{1}{3}\right).$$

ברשימה האורתונורמלית לעיל יש שלושה איברים, שזהו הממד של $\mathcal{P}_2(\mathbf{R})$. לכן רשימה אורתונורמלית זו היא בסיס אורתונורמלי של $\mathcal{P}_2(\mathbf{R})$ [לפי 6.28].

כעת ביכולתנו לענות על שאלת קיומם של בסיסים אורתונורמליים.

6.35 קיומו של בסיס אורתונורמלי

לכל מרחב מכפלה פנימית מממד סופי יש בסיס אורתונורמלי.

הוכחה נניח ש־ V מממד סופי. נבחר בסיס של V. ניישם את תהליך גרם־שמידט (6.32) על הבסיס הנבחר, וכתוצאה מכך נקבל רשימה אורתונורמלית שאורכה הוא $\dim V$. לפי 6.28, רשימה אורתונורמלית זו היא בסיס אורתונורמלי של V. ■

לפעמים נצטרך לדעת שלא זו בלבד שקיים בסיס אורתונורמלי, אלא שכל רשימה אורתונורמלית ניתנת להשלמה לבסיס אורתונורמלי. התוצאה הישירה הבאה של תהליך גרם־שמידט מראה כי השלמה כזו תמיד אפשרית.

6.36 *every orthonormal list extends to an orthonormal basis*

Suppose V is finite-dimensional. Then every orthonormal list of vectors in V can be extended to an orthonormal basis of V.

Proof Suppose $e_1, \dots, e_m$ is an orthonormal list of vectors in V. Then $e_1, \dots, e_m$ is linearly independent (by 6.25). Hence this list can be extended to a basis $e_1, \dots, e_m, v_1, \dots, v_n$ of V (see 2.32). Now apply the Gram–Schmidt procedure (6.32) to $e_1, \dots, e_m, v_1, \dots, v_n$, producing an orthonormal list

$$e_1, \dots, e_m, f_1, \dots, f_n;$$

here the formula given by the Gram–Schmidt procedure leaves the first m vectors unchanged because they are already orthonormal. The list above is an orthonormal basis of V by 6.28. ∎

Recall that a matrix is called upper triangular if it looks like this:

$$\begin{pmatrix} * & & * \\ & \ddots & \\ 0 & & * \end{pmatrix},$$

where the 0 in the matrix above indicates that all entries below the diagonal equal 0, and asterisks are used to denote entries on and above the diagonal.

In the last chapter, we gave a necessary and sufficient condition for an operator to have an upper-triangular matrix with respect to some basis (see 5.44). Now that we are dealing with inner product spaces, we would like to know whether there exists an *orthonormal* basis with respect to which we have an upper-triangular matrix. The next result shows that the condition for an operator to have an upper-triangular matrix with respect to some orthonormal basis is the same as the condition to have an upper-triangular matrix with respect to an arbitrary basis.

6.37 *upper-triangular matrix with respect to some orthonormal basis*

Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Then T has an upper-triangular matrix with respect to some orthonormal basis of V if and only if the minimal polynomial of T equals $(z - \lambda_1) \cdots (z - \lambda_m)$ for some $\lambda_1, \dots, \lambda_m \in \mathbf{F}$.

Proof Suppose T has an upper-triangular matrix with respect to some basis $v_1, \dots, v_n$ of V. Thus $\operatorname{span}(v_1, \dots, v_k)$ is invariant under T for each $k = 1, \dots, n$ (see 5.39).

Apply the Gram–Schmidt procedure to $v_1, \dots, v_n$, producing an orthonormal basis $e_1, \dots, e_n$ of V. Because

$$\operatorname{span}(e_1, \dots, e_k) = \operatorname{span}(v_1, \dots, v_k)$$

for each k (see 6.32), we conclude that $\operatorname{span}(e_1, \dots, e_k)$ is invariant under T for each $k = 1, \dots, n$. Thus, by 5.39, T has an upper-triangular matrix with respect to the orthonormal basis $e_1, \dots, e_n$. Now use 5.44 to complete the proof. ∎

6.36 כל רשימה אורתונורמלית ניתנת להשלמה לבסיס אורתונורמלי

נניח ש־ V מממד סופי. אזי ניתן להשלים כל רשימה אורתונורמלית של וקטורים ב־ V לכדי בסיס אורתונורמלי של V.

הוכחה נניח ש־ $e_1, \dots, e_m$ היא רשימה אורתונורמלית של וקטורים ב־ V. אזי $e_1, \dots, e_m$ בלתי תלויה ליניארית (לפי 6.25). לכן רשימה זו ניתנת להשלמה לכדי בסיס $e_1, \dots, e_m, v_1, \dots, v_n$ של V (ראו 2.32). כעת ניישם את תהליך גרם־שמידט (6.32) על $e_1, \dots, e_m, v_1, \dots, v_n$, ונקבל כתוצאה את הרשימה האורתונורמלית

$$e_1, \dots, e_m, f_1, \dots, f_n;$$

במקרה זה, הנוסחה שניתנת על ידי תהליך גרם־שמידט תשאיר את m הווקטורים הראשונים ללא שינוי מאחר שהם כבר אורתונורמליים. הרשימה לעיל היא בסיס אורתונורמלי של V לפי 6.28. ■

נזכיר כי מטריצה נקראת משולשית עליונה אם היא נראית כך:

$$\begin{pmatrix} * & & * \\ & \ddots & \\ 0 & & * \end{pmatrix},$$

כאשר ה־ 0 במטריצה לעיל מציין שכל האיברים תחת האלכסון שווים 0, והכוכביות משמשות לסמן את האיברים על האלכסון ומעליו.

בפרק האחרון, סיפקנו תנאי הכרחי ומספיק לכך שלאופרטור תהיה מטריצה משולשית עליונה ביחס לבסיס כלשהו (ראו 5.44). כעת כשאנו עוסקים במרחבי מכפלה פנימית, נרצה לדעת האם קיים בסיס אורתונורמלי אשר ביחס אליו יש לנו מטריצה משולשית עליונה. המסקנה הבאה מראה כי התנאי הנדרש שתהיה לאופרטור מטריצה משולשית עליונה ביחס לבסיס אורתונורמלי כלשהו הוא תנאי זהה לזה הנדרש עבור מטריצה משולשית עליונה ביחס לבסיס שרירותי.

6.37 מטריצה משולשית עליונה ביחס לבסיס אורתונורמלי כלשהו

נניח ש־ V מממד סופי ו־ $T \in \mathcal{L}(V)$. אזי יש ל־ T מטריצה משולשית עליונה ביחס לבסיס אורתונורמלי כלשהו של V אם־ורק־אם הפולינום המינימלי של T שווה ל־ $(z - \lambda_1) \cdots (z - \lambda_m)$ עבור $\lambda_1, \dots, \lambda_m$ כלשהם ב־ $\mathbf{F}$.

הוכחה נניח שיש ל־ T מטריצה משולשית עליונה ביחס ל־ $v_1, \dots, v_n$ בסיס כלשהו של V. לפיכך $\operatorname{span}(v_1, \dots, v_k)$ שמור תחת T לכל $k = 1, \dots, n$ (ראו 5.39).

יישום תהליך גרם־שמידט על $v_1, \dots, v_n$ מניב בסיס אורתונורמלי $e_1, \dots, e_n$ של V. מאחר ש־

$$\operatorname{span}(e_1, \dots, e_k) = \operatorname{span}(v_1, \dots, v_k)$$

לכל k (ראו 6.32), אנו מסיקים כי $\operatorname{span}(e_1, \dots, e_k)$ שמור תחת T לכל $k = 1, \dots, n$. לפיכך, לפי 5.39, יש ל־ T מטריצה משולשית עליונה ביחס לבסיס האורתונורמלי $e_1, \dots, e_n$. כעת השתמשו ב־ 5.44 להשלמת ההוכחה. ■

Issai Schur (1875–1941) published a proof of the next result in 1909.

For complex vector spaces, the next result is an important application of the result above. See Exercise 20 for a version of Schur's theorem that applies simultaneously to more than one operator.

6.38 *Schur's theorem*

Every operator on a finite-dimensional complex inner product space has an upper-triangular matrix with respect to some orthonormal basis.

Proof The desired result follows from the second version of the fundamental theorem of algebra (4.13) and 6.37. ∎

Linear Functionals on Inner Product Spaces

Because linear maps into the scalar field $\mathbf{F}$ play a special role, we defined a special name for them and their vector space in Section 3F. Those definitions are repeated below in case you skipped Section 3F.

6.39 definition: *linear functional, dual space, V'*

- A *linear functional* on V is a linear map from V to $\mathbf{F}$.
- The *dual space* of V, denoted by V', is the vector space of all linear functionals on V. In other words, $V' = \mathcal{L}(V, \mathbf{F})$.

6.40 example: *linear functional on $\mathbf{F}^3$*

The function $\varphi : \mathbf{F}^3 \to \mathbf{F}$ defined by

$$\varphi(z_1, z_2, z_3) = 2z_1 - 5z_2 + z_3$$

is a linear functional on $\mathbf{F}^3$. We could write this linear functional in the form

$$\varphi(z) = \langle z, w \rangle$$

for every $z \in \mathbf{F}^3$, where $w = (2, -5, 1)$.

6.41 example: *linear functional on $\mathcal{P}_5(\mathbf{R})$*

The function $\varphi : \mathcal{P}_5(\mathbf{R}) \to \mathbf{R}$ defined by

$$\varphi(p) = \int_{-1}^{1} p(t)(\cos(\pi t))\, dt$$

is a linear functional on $\mathcal{P}_5(\mathbf{R})$.

ישי שור (1875–1941) פרסם הוכחה למסקנה הבאה בשנת 1909.

עבור מרחבים וקטוריים מרוכבים המסקנה הבאה היא יישום חשוב של המסקנה הקודמת. ראו תרגיל 20 לגרסה של משפט הפירוק של שור הניתנת ליישום על מספר אופרטורים בו־זמנית.

6.38 משפט הפירוק של שור

לכל אופרטור מעל מרחב מכפלה פנימית מרוכב מממד סופי יש מטריצה משולשית עליונה ביחס לבסיס אורתונורמלי כלשהו.

הוכחה מסקנה זו נובעת באופן ישיר מהמשפט היסודי של האלגברה בגרסתו השנייה (4.13) וממסקנה 6.37.

פונקציונלים ליניאריים מעל מרחבי מכפלה פנימית

מאחר שלהעתקות ליניאריות אל שדה הסקלרים $\mathbf{F}$ יש תפקיד מיוחד, הגדרנו עבורן שם מיוחד כמו גם עבור המרחב הווקטורי שלהן ביחידה 6ג. ההגדרות הללו מובאות כאן שוב למקרה שדילגתם על יחידה 6ג.

6.39 הגדרה: פונקציונל ליניארי, מרחב דואלי, V'

- פונקציונל ליניארי מעל V הוא העתקה ליניארית מ־ V אל $\mathbf{F}$.
- המרחב הדואלי של V, המסומן ב־ V', הוא המרחב הווקטורי של כל הפונקציונלים הליניאריים מעל V. במילים אחרות, $V' = \mathcal{L}(V, \mathbf{F})$.

6.40 דוגמה: פונקציונל ליניארי מעל $\mathbf{F}^3$

הפונקציה $\varphi : \mathbf{F}^3 \to \mathbf{F}$ המוגדרת על ידי

$$\varphi(z_1, z_2, z_3) = 2z_1 - 5z_2 + z_3$$

היא פונקציונל ליניארי מעל $\mathbf{F}^3$. נוכל לכתוב את הפונקציונל הליניארי הזה בצורה

$$\varphi(z) = \langle z, w \rangle$$

לכל $z \in \mathbf{F}^3$, כאשר $w = (2, -5, 1)$.

6.41 דוגמה: פונקציונל ליניארי מעל $\mathcal{P}_5(\mathbf{R})$

הפונקציה $\varphi : \mathcal{P}_5(\mathbf{R}) \to \mathbf{R}$ המוגדרת על ידי

$$\varphi(p) = \int_{-1}^{1} p(t)(\cos(\pi t))\, dt$$

היא פונקציונל ליניארי מעל $\mathcal{P}_5(\mathbf{R})$.

If $v \in V$, then the map that sends u to $\langle u, v\rangle$ is a linear functional on V. The next result states that every linear functional on V is of this form. For example, we can take $v = (2, -5, 1)$ in Example 6.40.

The next result is named in honor of Frigyes Riesz (1880–1956), who proved several theorems early in the twentieth century that look very much like the result below.

Suppose we make the vector space $\mathcal{P}_5(\mathbf{R})$ into an inner product space by defining $\langle p, q\rangle = \int_{-1}^{1} pq$. Let φ be as in Example 6.41. It is not obvious that there exists $q \in \mathcal{P}_5(\mathbf{R})$ such that

$$\int_{-1}^{1} p(t)(\cos(\pi t))\, dt = \langle p, q\rangle$$

for every $p \in \mathcal{P}_5(\mathbf{R})$ [we cannot take $q(t) = \cos(\pi t)$ because that choice of q is not an element of $\mathcal{P}_5(\mathbf{R})$]. The next result tells us the somewhat surprising result that there indeed exists a polynomial $q \in \mathcal{P}_5(\mathbf{R})$ such that the equation above holds for all $p \in \mathcal{P}_5(\mathbf{R})$.

6.42 *Riesz representation theorem*

Suppose V is finite-dimensional and φ is a linear functional on V. Then there is a unique vector $v \in V$ such that

$$\varphi(u) = \langle u, v\rangle$$

for every $u \in V$.

Proof First we show that there exists a vector $v \in V$ such that $\varphi(u) = \langle u, v\rangle$ for every $u \in V$. Let $e_1, \dots, e_n$ be an orthonormal basis of V. Then

$$\begin{aligned}\varphi(u) &= \varphi(\langle u, e_1\rangle e_1 + \cdots + \langle u, e_n\rangle e_n)\\ &= \langle u, e_1\rangle \varphi(e_1) + \cdots + \langle u, e_n\rangle \varphi(e_n)\\ &= \big\langle u, \overline{\varphi(e_1)} e_1 + \cdots + \overline{\varphi(e_n)} e_n\big\rangle\end{aligned}$$

for every $u \in V$, where the first equality comes from 6.30(a). Thus setting

6.43 $$v = \overline{\varphi(e_1)} e_1 + \cdots + \overline{\varphi(e_n)} e_n,$$

we have $\varphi(u) = \langle u, v\rangle$ for every $u \in V$, as desired.

Now we prove that only one vector $v \in V$ has the desired behavior. Suppose $v_1, v_2 \in V$ are such that

$$\varphi(u) = \langle u, v_1\rangle = \langle u, v_2\rangle$$

for every $u \in V$. Then

$$0 = \langle u, v_1\rangle - \langle u, v_2\rangle = \langle u, v_1 - v_2\rangle$$

for every $u \in V$. Taking $u = v_1 - v_2$ shows that $v_1 - v_2 = 0$. Thus $v_1 = v_2$, completing the proof of the uniqueness part of the result. ∎

המסקנה הבאה נקראת על שמו של פרידיש ריס (1880–1956), אשר הוכיח מספר משפטים בראשית המאה העשרים הדומים מאוד למסקנה זו.

אם $v \in V$, אזי ההעתקה השולחת את u אל $\langle u, v\rangle$ היא פונקציונל ליניארי מעל V. המסקנה הבאה קובעת כי כל פונקציונל ליניארי מעל V הוא מהצורה הזו. לדוגמה, ניקח $v = (2, -5, 1)$ בדוגמה 6.40.

נניח שאנו הופכים את המרחב הווקטורי $\mathcal{P}_5(\mathbf{R})$ למרחב מכפלה פנימית על ידי כך שנגדיר $\langle p, q\rangle = \int_{-1}^{1} pq$. נגדיר את φ כבדוגמה 6.41. אין הדבר מובן מאליו כי קיים $q \in \mathcal{P}_5(\mathbf{R})$ כך ש־

$$\int_{-1}^{1} p(t)(\cos(\pi t))\, dt = \langle p, q\rangle$$

לכל $p \in \mathcal{P}_5(\mathbf{R})$ [לא ניתן להגדיר $q(t) = \cos(\pi t)$ מאחר שבחירה זו של q אינה איבר של $\mathcal{P}_5(\mathbf{R})$]. המסקנה הבאה מלמדת אותנו תוצאה מפתיעה במידת מה: אכן קיים פולינום $q \in \mathcal{P}_5(\mathbf{R})$ כך שהמשוואה לעיל מתקיימת לכל $p \in \mathcal{P}_5(\mathbf{R})$.

6.42 משפט ההצגה של ריס

נניח ש־ V מממד סופי ו־ φ פונקציונל ליניארי מעל V. אזי קיים וקטור יחיד $v \in V$ כך ש־

$$\varphi(u) = \langle u, v\rangle$$

לכל $u \in V$.

הוכחה תחילה נראה שאכן קיים וקטור $v \in V$ כך ש־ $\varphi(u) = \langle u, v\rangle$ לכל $u \in V$. יהי $e_1, \dots, e_n$ בסיס אורתונורמלי של V. אזי

$$\begin{aligned}\varphi(u) &= \varphi(\langle u, e_1\rangle e_1 + \cdots + \langle u, e_n\rangle e_n)\\ &= \langle u, e_1\rangle \varphi(e_1) + \cdots + \langle u, e_n\rangle \varphi(e_n)\\ &= \left\langle u, \overline{\varphi(e_1)} e_1 + \cdots + \overline{\varphi(e_n)} e_n\right\rangle\end{aligned}$$

לכל $u \in V$, כאשר השוויון הראשון נובע מ־ 6.30(א). לפיכך בהגדרת

$$v = \overline{\varphi(e_1)} e_1 + \cdots + \overline{\varphi(e_n)} e_n, \tag{6.43}$$

נקבל $\varphi(u) = \langle u, v\rangle$ לכל $u \in V$, כפי שרצינו.

כעת נוכיח כי רק וקטור אחד $v \in V$ מקיים את התכונה הרצויה. נניח ש־ $v_1, v_2 \in V$ הם כך ש־

$$\varphi(u) = \langle u, v_1\rangle = \langle u, v_2\rangle$$

לכל $u \in V$. אזי

$$0 = \langle u, v_1\rangle - \langle u, v_2\rangle = \langle u, v_1 - v_2\rangle$$

לכל $u \in V$. בחירת $u = v_1 - v_2$ מראה לנו ש־ $v_1 - v_2 = 0$. לפיכך $v_1 = v_2$, ובזאת הושלמה הוכחת חלק היחידות של המסקנה. ■

6.44 example: *computation illustrating Riesz representation theorem*

Suppose we want to find a polynomial $q \in \mathcal{P}_2(\mathbf{R})$ such that

$$\int_{-1}^{1} p(t)(\cos(\pi t))\,dt = \int_{-1}^{1} pq \tag{6.45}$$

for every polynomial $p \in \mathcal{P}_2(\mathbf{R})$. To do this, we make $\mathcal{P}_2(\mathbf{R})$ into an inner product space by defining $\langle p, q\rangle$ to be the right side of the equation above for $p, q \in \mathcal{P}_2(\mathbf{R})$. Note that the left side of the equation above does not equal the inner product in $\mathcal{P}_2(\mathbf{R})$ of p and the function $t \mapsto \cos(\pi t)$ because this last function is not a polynomial.

Define a linear functional φ on $\mathcal{P}_2(\mathbf{R})$ by letting

$$\varphi(p) = \int_{-1}^{1} p(t)(\cos(\pi t))\,dt$$

for each $p \in \mathcal{P}_2(\mathbf{R})$. Now use the orthonormal basis from Example 6.34 and apply formula 6.43 from the proof of the Riesz representation theorem to see that if $p \in \mathcal{P}_2(\mathbf{R})$, then $\varphi(p) = \langle p, q\rangle$, where

$$q(x) = \left(\int_{-1}^{1} \sqrt{\tfrac{1}{2}}\cos(\pi t)\,dt\right)\sqrt{\tfrac{1}{2}} + \left(\int_{-1}^{1} \sqrt{\tfrac{3}{2}}\,t\cos(\pi t)\,dt\right)\sqrt{\tfrac{3}{2}}x$$
$$+ \left(\int_{-1}^{1} \sqrt{\tfrac{45}{8}}\left(t^2 - \tfrac{1}{3}\right)\cos(\pi t)\,dt\right)\sqrt{\tfrac{45}{8}}\left(x^2 - \tfrac{1}{3}\right).$$

A bit of calculus applied to the equation above shows that

$$q(x) = \tfrac{15}{2\pi^2}(1 - 3x^2).$$

The same procedure shows that if we want to find $q \in \mathcal{P}_5(\mathbf{R})$ such that 6.45 holds for all $p \in \mathcal{P}_5(\mathbf{R})$, then we should take

$$q(x) = \tfrac{105}{8\pi^4}\left((27 - 2\pi^2) + (24\pi^2 - 270)x^2 + (315 - 30\pi^2)x^4\right).$$

Suppose V is finite-dimensional and φ a linear functional on V. Then 6.43 gives a formula for the vector v that satisfies

$$\varphi(u) = \langle u, v\rangle$$

for all $u \in V$. Specifically, we have

$$v = \overline{\varphi(e_1)}e_1 + \cdots + \overline{\varphi(e_n)}e_n.$$

The right side of the equation above seems to depend on the orthonormal basis $e_1, \ldots, e_n$ as well as on φ. However, 6.42 tells us that v is uniquely determined by φ. Thus the right side of the equation above is the same regardless of which orthonormal basis $e_1, \ldots, e_n$ of V is chosen.

For two additional different proofs of the Riesz representation theorem, see 6.58 and also Exercise 13 in Section 6C.

6.44 דוגמה: חישוב המַמְחִישׁ את משפט ההצגה של ריס

נניח שאנו רוצים למצוא פולינום $q \in \mathcal{P}_2(\mathbf{R})$ כך ש־

$$\int_{-1}^{1} p(t)(\cos(\pi t))\,dt = \int_{-1}^{1} pq \tag{6.45}$$

לכל פולינום $p \in \mathcal{P}_2(\mathbf{R})$. לשם כך, נהפוך את $\mathcal{P}_2(\mathbf{R})$ למרחב מכפלה פנימית על ידי כך שנגדיר את $\langle p, q\rangle$ כאגף הימני של המשוואה לעיל עבור $p, q \in \mathcal{P}_2(\mathbf{R})$. נשים לב כי האגף השמאלי של המשוואה לעיל אינו שווה למכפלה הפנימית ב־ $\mathcal{P}_2(\mathbf{R})$ של p והפונקציה $t \mapsto \cos(\pi t)$ שכן הפונקציה האחרונה אינה פולינום.

נגדיר פונקציונל ליניארי φ מעל $\mathcal{P}_2(\mathbf{R})$ על ידי

$$\varphi(p) = \int_{-1}^{1} p(t)(\cos(\pi t))\,dt$$

לכל $p \in \mathcal{P}_2(\mathbf{R})$. כעת נשתמש בבסיס האורתונורמלי מדוגמה 6.34 וניישם עליו את נוסחה 6.43 מההוכחה של משפט ההצגה של ריס כדי לראות שאכן אם $p \in \mathcal{P}_2(\mathbf{R})$, אזי $\varphi(p) = \langle p, q\rangle$, כאשר

$$q(x) = \left(\int_{-1}^{1} \sqrt{\tfrac{1}{2}}\cos(\pi t)\,dt\right)\sqrt{\tfrac{1}{2}} + \left(\int_{-1}^{1} \sqrt{\tfrac{3}{2}}\,t\cos(\pi t)\,dt\right)\sqrt{\tfrac{3}{2}}x$$

$$+ \left(\int_{-1}^{1} \sqrt{\tfrac{45}{8}}\left(t^2 - \tfrac{1}{3}\right)\cos(\pi t)\,dt\right)\sqrt{\tfrac{45}{8}}\left(x^2 - \tfrac{1}{3}\right).$$

מעט חשבון אינפיניטסימלי המיושם על המשוואה מראה לנו כי

$$q(x) = \tfrac{15}{2\pi^2}(1 - 3x^2).$$

אותו התהליך מראה שאם ברצוננו למצוא $q \in \mathcal{P}_5(\mathbf{R})$ כך ש־ 6.45 מתקיים לכל $p \in \mathcal{P}_5(\mathbf{R})$, אזי עלינו לבחור

$$q(x) = \tfrac{105}{8\pi^4}\Big((27 - 2\pi^2) + (24\pi^2 - 270)x^2 + (315 - 30\pi^2)x^4\Big).$$

נניח ש־ V מממד סופי ו־ φ פונקציונל ליניארי מעל V. אזי 6.43 מעניקה לנו את הנוסחה לווקטור v המקיים

$$\varphi(u) = \langle u, v\rangle$$

לכל $u \in V$. ליתר דיוק, קיבלנו

$$v = \overline{\varphi(e_1)}e_1 + \cdots + \overline{\varphi(e_n)}e_n.$$

האגף הימני של המשוואה לעיל נראה תלוי בבסיס האורתונורמלי $e_1, \ldots, e_n$ כמו גם ב־ φ. אולם, 6.42 קובעת ש־ v נקבע באופן יחיד על ידי φ. לפיכך האגף הימני של המשוואה לעיל נותר בעינו ללא תלות בזהות הבסיס האורתונורמלי $e_1, \ldots, e_n$ של V שנבחר.

לשתי הוכחות נוספות למשפט ההצגה של ריס, ראו 6.58 ובנוסף תרגיל 13 ביחידה 10.

Exercises 6B

1 Suppose $e_1, \dots, e_m$ is a list of vectors in V such that

$$\|a_1 e_1 + \cdots + a_m e_m\|^2 = |a_1|^2 + \cdots + |a_m|^2$$

for all $a_1, \dots, a_m \in \mathbf{F}$. Show that $e_1, \dots, e_m$ is an orthonormal list.

This exercise provides a converse to 6.24.

2 (a) Suppose $\theta \in \mathbf{R}$. Show that both

$$(\cos\theta, \sin\theta), (-\sin\theta, \cos\theta) \quad \text{and} \quad (\cos\theta, \sin\theta), (\sin\theta, -\cos\theta)$$

are orthonormal bases of $\mathbf{R}^2$.

(b) Show that each orthonormal basis of $\mathbf{R}^2$ is of the form given by one of the two possibilities in (a).

3 Suppose $e_1, \dots, e_m$ is an orthonormal list in V and $v \in V$. Prove that

$$\|v\|^2 = |\langle v, e_1\rangle|^2 + \cdots + |\langle v, e_m\rangle|^2 \iff v \in \operatorname{span}(e_1, \dots, e_m).$$

4 Suppose n is a positive integer. Prove that

$$\frac{1}{\sqrt{2\pi}}, \frac{\cos x}{\sqrt{\pi}}, \frac{\cos 2x}{\sqrt{\pi}}, \dots, \frac{\cos nx}{\sqrt{\pi}}, \frac{\sin x}{\sqrt{\pi}}, \frac{\sin 2x}{\sqrt{\pi}}, \dots, \frac{\sin nx}{\sqrt{\pi}}$$

is an orthonormal list of vectors in $C[-\pi, \pi]$, the vector space of continuous real-valued functions on $[-\pi, \pi]$ with inner product

$$\langle f, g\rangle = \int_{-\pi}^{\pi} fg.$$

Hint: The following formulas should help.

$$(\sin x)(\cos y) = \frac{\sin(x-y) + \sin(x+y)}{2}$$

$$(\sin x)(\sin y) = \frac{\cos(x-y) - \cos(x+y)}{2}$$

$$(\cos x)(\cos y) = \frac{\cos(x-y) + \cos(x+y)}{2}$$

5 Suppose $f : [-\pi, \pi] \to \mathbf{R}$ is continuous. For each nonnegative integer k, define

$$a_k = \frac{1}{\sqrt{\pi}} \int_{-\pi}^{\pi} f(x)\cos(kx)\,dx \quad \text{and} \quad b_k = \frac{1}{\sqrt{\pi}} \int_{-\pi}^{\pi} f(x)\sin(kx)\,dx.$$

Prove that

$$\frac{a_0^2}{2} + \sum_{k=1}^{\infty} (a_k^2 + b_k^2) \le \int_{-\pi}^{\pi} f^2.$$

The inequality above is actually an equality for all continuous functions $f : [-\pi, \pi] \to \mathbf{R}$. However, proving that this inequality is an equality involves Fourier series techniques beyond the scope of this book.

תרגילים ו7

1 נניח ש־ $e_1, \dots, e_m$ רשימת וקטורים ב־ V כך ש־

$$\|a_1 e_1 + \cdots + a_m e_m\|^2 = |a_1|^2 + \cdots + |a_m|^2$$

לכל $a_1, \dots, a_m \in \mathbf{F}$. הראו ש־ $e_1, \dots, e_m$ רשימה אורתונורמלית.

תרגיל זה מספק את הכיוון ההפוך של התוצאה ב־ 6.24.

2 א. נניח ש־ $\theta \in \mathbf{R}$. הראו ששתי הרשימות

$$(\cos\theta, \sin\theta), (-\sin\theta, \cos\theta) \quad \text{ו־} \quad (\cos\theta, \sin\theta), (\sin\theta, -\cos\theta)$$

בסיסים אורתונורמליים של $\mathbf{R}^2$.

ב. הראו שכל בסיס אורתונורמלי של $\mathbf{R}^2$ הוא מהצורה הנתונה על ידי אחת משתי האפשרויות בסעיף (א).

3 נניח ש־ $e_1, \dots, e_m$ רשימה אורתונורמלית ב־ V ו־ $v \in V$. הוכיחו ש־

$$\|v\|^2 = |\langle v, e_1\rangle|^2 + \cdots + |\langle v, e_m\rangle|^2 \iff v \in \operatorname{span}(e_1, \dots, e_m).$$

4 נניח ש־ n מספר שלם חיובי. הוכיחו ש־

$$\frac{1}{\sqrt{2\pi}}, \frac{\cos x}{\sqrt{\pi}}, \frac{\cos 2x}{\sqrt{\pi}}, \dots, \frac{\cos nx}{\sqrt{\pi}}, \frac{\sin x}{\sqrt{\pi}}, \frac{\sin 2x}{\sqrt{\pi}}, \dots, \frac{\sin nx}{\sqrt{\pi}}$$

רשימה אורתונורמלית של וקטורים ב־ $C[-\pi, \pi]$, המרחב הווקטורי של פונקציות רציפות ממשיות מעל $[-\pi, \pi]$ עם מכפלה פנימית

$$\langle f, g\rangle = \int_{-\pi}^{\pi} fg.$$

רמז: העזרו בנוסחאות הבאות.

$$(\sin x)(\cos y) = \frac{\sin(x-y) + \sin(x+y)}{2}$$

$$(\sin x)(\sin y) = \frac{\cos(x-y) - \cos(x+y)}{2}$$

$$(\cos x)(\cos y) = \frac{\cos(x-y) + \cos(x+y)}{2}$$

5 נניח ש־ $f\colon [-\pi, \pi] \to \mathbf{R}$ רציפה. לכל מספר שלם אי־שלילי k, הגדירו

$$a_k = \frac{1}{\sqrt{\pi}} \int_{-\pi}^{\pi} f(x)\cos(kx)\,dx \quad \text{וגם} \quad b_k = \frac{1}{\sqrt{\pi}} \int_{-\pi}^{\pi} f(x)\sin(kx)\,dx.$$

הוכיחו ש־

$$\frac{a_0^2}{2} + \sum_{k=1}^{\infty} (a_k^2 + b_k^2) \le \int_{-\pi}^{\pi} f^2.$$

$f\colon [-\pi, \pi] \to \mathbf{R}$. האי־שוויון לעיל הוא למעשה שוויון לכל הפונקציות הרציפות אולם, להוכיח כי אי־שוויון זה הוא למעשה שוויון ידרוש שימוש בטכניקות טור פורייה החורגות מתחומו של ספר זה.

6 Suppose $e_1, \dots, e_n$ is an orthonormal basis of V.

(a) Prove that if $v_1, \dots, v_n$ are vectors in V such that

$$\|e_k - v_k\| < \frac{1}{\sqrt{n}}$$

for each k, then $v_1, \dots, v_n$ is a basis of V.

(b) Show that there exist $v_1, \dots, v_n \in V$ such that

$$\|e_k - v_k\| \le \frac{1}{\sqrt{n}}$$

for each k, but $v_1, \dots, v_n$ is not linearly independent.

This exercise states in (a) that an appropriately small perturbation of an orthonormal basis is a basis. Then (b) shows that the number $1/\sqrt{n}$ on the right side of the inequality in (a) cannot be higher.

7 Suppose $T \in \mathcal{L}(\mathbf{R}^3)$ has an upper-triangular matrix with respect to the basis $(1, 0, 0)$, $(1, 1, 1)$, $(1, 1, 2)$. Find an orthonormal basis of $\mathbf{R}^3$ with respect to which T has an upper-triangular matrix.

8 Make $\mathcal{P}_2(\mathbf{R})$ into an inner product space by defining $\langle p, q \rangle = \int_0^1 pq$ for all $p, q \in \mathcal{P}_2(\mathbf{R})$.

(a) Apply the Gram–Schmidt procedure to the basis $1, x, x^2$ to produce an orthonormal basis of $\mathcal{P}_2(\mathbf{R})$.

(b) The differentiation operator (the operator that takes p to p') on $\mathcal{P}_2(\mathbf{R})$ has an upper-triangular matrix with respect to the basis $1, x, x^2$, which is not an orthonormal basis. Find the matrix of the differentiation operator on $\mathcal{P}_2(\mathbf{R})$ with respect to the orthonormal basis produced in (a) and verify that this matrix is upper triangular, as expected from the proof of 6.37.

9 Suppose $e_1, \dots, e_m$ is the result of applying the Gram–Schmidt procedure to a linearly independent list $v_1, \dots, v_m$ in V. Prove that $\langle v_k, e_k \rangle > 0$ for each $k = 1, \dots, m$.

10 Suppose $v_1, \dots, v_m$ is a linearly independent list in V. Explain why the orthonormal list produced by the formulas of the Gram–Schmidt procedure (6.32) is the only orthonormal list $e_1, \dots, e_m$ in V such that $\langle v_k, e_k \rangle > 0$ and $\operatorname{span}(v_1, \dots, v_k) = \operatorname{span}(e_1, \dots, e_k)$ for each $k = 1, \dots, m$.

The result in this exercise is used in the proof of 7.58.

11 Find a polynomial $q \in \mathcal{P}_2(\mathbf{R})$ such that $p\left(\frac{1}{2}\right) = \int_0^1 pq$ for every $p \in \mathcal{P}_2(\mathbf{R})$.

12 Find a polynomial $q \in \mathcal{P}_2(\mathbf{R})$ such that

$$\int_0^1 p(x) \cos(\pi x)\, dx = \int_0^1 pq$$

for every $p \in \mathcal{P}_2(\mathbf{R})$.

6 נניח ש־ $e_1, \dots, e_n$ בסיס אורתונורמלי של V.

א. הוכיחו שאם $v_1, \dots, v_n$ וקטורים ב־ V כך ש־
$$\|e_k - v_k\| < \frac{1}{\sqrt{n}}$$
לכל k, אזי $v_1, \dots, v_n$ בסיס של V.

ב. הראו שקיימים $v_1, \dots, v_n \in V$ כך ש־
$$\|e_k - v_k\| \le \frac{1}{\sqrt{n}}$$
לכל k, אבל $v_1, \dots, v_n$ אינה בלתי תלויה ליניארית.

תרגיל זה קובע בסעיף (א) כי בסיס אורתונורמלי ששובש במידה זעירה הוא עדיין בסיס. לאחר מכן סעיף (ב) מראה שהמספר $1/\sqrt{n}$ באגף הימני של האי־שוויון בסעיף (א) אינו יכול להיות גבוה יותר.

7 נניח שיש ל־ $T \in \mathcal{L}(\mathbf{R}^3)$ מטריצה משולשית עליונה ביחס לבסיס $(1,1,2), (1,1,1), (1,0,0)$. מצאו בסיס אורתונורמלי של $\mathbf{R}^3$ אשר ביחס אליו יש ל־ T מטריצה משולשית עליונה.

8 הפכו את $\mathcal{P}_2(\mathbf{R})$ למרחב מכפלה פנימית באמצעות הגדרת $\langle p, q\rangle = \int_0^1 pq$ לכל $p, q \in \mathcal{P}_2(\mathbf{R})$.

א. הפעילו את תהליך גרם־שמידט על הבסיס $1, x, x^2$ ליצירת בסיס אורתונורמלי של $\mathcal{P}_2(\mathbf{R})$.

ב. לאופרטור הגזירה (האופרטור המעתיק p אל p') מעל $\mathcal{P}_2(\mathbf{R})$ יש מטריצה משולשית עליונה ביחס לבסיס $1, x, x^2$, שהוא אינו בסיס אורתונורמלי. מצאו את המטריצה של אופרטור הגזירה מעל $\mathcal{P}_2(\mathbf{R})$ ביחס לבסיס האורתונורמלי שיצרנו בסעיף (א) וודאו כי זו אכן מטריצה משולשית עליונה, כמצופה מההוכחה של 6.37.

9 נניח ש־ $e_1, \dots, e_m$ הוא תוצאת הפעלת תהליך גרם־שמידט על רשימה בלתי תלויה ליניארית $v_1, \dots, v_m$ ב־ V. הוכיחו ש־ $\langle v_k, e_k\rangle > 0$ לכל $k = 1, \dots, m$.

10 נניח ש־ $v_1, \dots, v_m$ רשימה בלתי תלויה ליניארית ב־ V. הסבירו מדוע הרשימה האורתונורמלית הנוצרת באמצעות הנוסחאות של תהליך גרם־שמידט (6.32) היא הרשימה האורתונורמלית היחידה $e_1, \dots, e_m$ ב־ V כך ש־ $\langle v_k, e_k\rangle > 0$ ו־ $\operatorname{span}(v_1, \dots, v_k) = \operatorname{span}(e_1, \dots, e_k)$ לכל $k = 1, \dots, m$.

בהוכחה של 7.58 אנו מבצעים שימוש במסקנה של תרגיל זה.

11 מצאו פולינום $q \in \mathcal{P}_2(\mathbf{R})$ כך ש־ $p\left(\frac{1}{2}\right) = \int_0^1 pq$ לכל $p \in \mathcal{P}_2(\mathbf{R})$.

12 מצאו פולינום $q \in \mathcal{P}_2(\mathbf{R})$ כך ש־
$$\int_0^1 p(x)\cos(\pi x)\,dx = \int_0^1 pq$$
לכל $p \in \mathcal{P}_2(\mathbf{R})$.

13 Show that a list $v_1, \dots, v_m$ of vectors in V is linearly dependent if and only if the Gram–Schmidt formula in 6.32 produces $f_k = 0$ for some $k \in \{1, \dots, m\}$.

This exercise gives an alternative to Gaussian elimination techniques for determining whether a list of vectors in an inner product space is linearly dependent.

14 Suppose V is a real inner product space and $v_1, \dots, v_m$ is a linearly independent list of vectors in V. Prove that there exist exactly 2^m orthonormal lists $e_1, \dots, e_m$ of vectors in V such that

$$\operatorname{span}(v_1, \dots, v_k) = \operatorname{span}(e_1, \dots, e_k)$$

for all $k \in \{1, \dots, m\}$.

15 Suppose $\langle \cdot, \cdot \rangle_1$ and $\langle \cdot, \cdot \rangle_2$ are inner products on V such that $\langle u, v \rangle_1 = 0$ if and only if $\langle u, v \rangle_2 = 0$. Prove that there is a positive number c such that $\langle u, v \rangle_1 = c \langle u, v \rangle_2$ for every $u, v \in V$.

This exercise shows that if two inner products have the same pairs of orthogonal vectors, then each of the inner products is a scalar multiple of the other inner product.

16 Suppose V is finite-dimensional. Suppose $\langle \cdot, \cdot \rangle_1, \langle \cdot, \cdot \rangle_2$ are inner products on V with corresponding norms $\| \cdot \|_1$ and $\| \cdot \|_2$. Prove that there exists a positive number c such that $\|v\|_1 \le c \|v\|_2$ for every $v \in V$.

17 Suppose $\mathbf{F} = \mathbf{C}$ and V is finite-dimensional. Prove that if T is an operator on V such that 1 is the only eigenvalue of T and $\|Tv\| \le \|v\|$ for all $v \in V$, then T is the identity operator.

18 Suppose $u_1, \dots, u_m$ is a linearly independent list in V. Show that there exists $v \in V$ such that $\langle u_k, v \rangle = 1$ for all $k \in \{1, \dots, m\}$.

19 Suppose $v_1, \dots, v_n$ is a basis of V. Prove that there exists a basis $u_1, \dots, u_n$ of V such that

$$\langle v_j, u_k \rangle = \begin{cases} 0 & \text{if } j \ne k, \\ 1 & \text{if } j = k. \end{cases}$$

20 Suppose $\mathbf{F} = \mathbf{C}$, V is finite-dimensional, and $\mathcal{E} \subseteq \mathcal{L}(V)$ is such that

$$ST = TS$$

for all $S, T \in \mathcal{E}$. Prove that there is an orthonormal basis of V with respect to which every element of $\mathcal{E}$ has an upper-triangular matrix.

This exercise strengthens Exercise 9(b) in Section 5E (in the context of inner product spaces) by asserting that the basis in that exercise can be chosen to be orthonormal.

21 Suppose $\mathbf{F} = \mathbf{C}$, V is finite-dimensional, $T \in \mathcal{L}(V)$, and all eigenvalues of T have absolute value less than 1. Let $\epsilon > 0$. Prove that there exists a positive integer m such that $\|T^m v\| \le \epsilon \|v\|$ for every $v \in V$.

13 הראו שרשימה $v_1, \dots, v_m$ של וקטורים ב־ V היא תלויה ליניארית אם־ורק־אם נוסחת גרם־שמידט ב־ 6.32 מפיקה $f_k = 0$ עבור $k \in \{1, \dots, m\}$ כלשהו.

תרגיל זה מספק חלופה לשיטת החילוץ של גאוס לזירוג מטריצות לקביעה האם רשימת וקטורים במרחב מכפלה פנימית תלויה ליניארית.

14 נניח ש־ V מרחב מכפלה פנימית ממשי ו־ $v_1, \dots, v_m$ רשימה בלתי תלויה ליניארית של וקטורים ב־ V. הוכיחו כי קיימות בדיוק 2^m רשימות אורתונורמליות $e_1, \dots, e_m$ של וקטורים ב־ V כך ש־

$$\operatorname{span}(v_1, \dots, v_k) = \operatorname{span}(e_1, \dots, e_k)$$

לכל $k \in \{1, \dots, m\}$.

15 נניח ש־ $\langle \cdot, \cdot \rangle_1$ ו־ $\langle \cdot, \cdot \rangle_2$ הן מכפלות פנימיות מעל V כך ש־ $\langle u, v \rangle_1 = 0$ אם־ורק־אם $\langle u, v \rangle_2 = 0$. הוכיחו שקיים מספר חיובי c כך ש־ $\langle u, v \rangle_1 = c \langle u, v \rangle_2$ לכל $u, v \in V$.

תרגיל זה מראה שאם לשתי מכפלות פנימיות אותם זוגות של וקטורים אורתוגונליים, אזי כל אחת מהמכפלות הפנימיות היא כפולת סקלר של המכפלה הפנימית האחרת.

16 נניח ש־ V מממד סופי. נניח ש־ $\langle \cdot, \cdot \rangle_1$, $\langle \cdot, \cdot \rangle_2$ מכפלות פנימיות מעל V עם הנורמות $\| \cdot \|_1$ ו־ $\| \cdot \|_2$ בהתאמה. הוכיחו שקיים מספר חיובי c כך ש־ $\|v\|_1 \le c\|v\|_2$ לכל $v \in V$.

17 נניח ש־ $\mathbf{F} = \mathbf{C}$ ו־ V מממד סופי. הוכיחו שאם T הוא אופרטור מעל V כך ש־ 1 הוא הערך העצמי היחיד של T ו־ $\|Tv\| \le \|v\|$ לכל $v \in V$, אזי T הוא אופרטור הזהות.

18 נניח ש־ $u_1, \dots, u_m$ רשימה בלתי תלויה ליניארית ב־ V. הראו שקיים $v \in V$ כך ש־ $\langle u_k, v \rangle = 1$ לכל $k \in \{1, \dots, m\}$.

19 נניח ש־ $v_1, \dots, v_n$ בסיס של V. הוכיחו שקיים בסיס $u_1, \dots, u_n$ של V כך ש־

$$\langle v_j, u_k \rangle = \begin{cases} 0 & \text{if } j \neq k, \\ 1 & \text{if } j = k. \end{cases}$$

20 נניח ש־ V, $\mathbf{F} = \mathbf{C}$ מממד סופי, ו־ $\mathcal{E} \subseteq \mathcal{L}(V)$ כך ש־

$$ST = TS$$

לכל $S, T \in \mathcal{E}$. הוכיחו שקיים בסיס אורתונורמלי של V אשר ביחס אליו לכל איבר ב־ $\mathcal{E}$ יש מטריצה משולשית עליונה.

תרגיל זה מחזק את תרגיל 9(ב) ביחידה ה5 (בהקשר של מרחבי מכפלה פנימית) בכך שקובע כי את הבסיס בתרגיל ההוא אפשר לבחור שיהיה אורתונורמלי.

21 נניח ש־ $\mathbf{F} = \mathbf{C}$, ו־ V מממד סופי, $T \in \mathcal{L}(V)$, והערך המוחלט של כל אחד מהערכים העצמיים של T קטן מ־ 1. יהי $\epsilon > 0$. הוכיחו שקיים מספר שלם חיובי m כך ש־ $\|T^m v\| \le \epsilon \|v\|$ לכל $v \in V$.

22 Suppose $C[-1,1]$ is the vector space of continuous real-valued functions on the interval $[-1,1]$ with inner product given by

$$\langle f,g\rangle = \int_{-1}^{1} fg$$

for all $f,g \in C[-1,1]$. Let φ be the linear functional on $C[-1,1]$ defined by $\varphi(f) = f(0)$. Show that there does not exist $g \in C[-1,1]$ such that

$$\varphi(f) = \langle f,g\rangle$$

for every $f \in C[-1,1]$.

This exercise shows that the Riesz representation theorem (6.42) does not hold on infinite-dimensional vector spaces without additional hypotheses on V and φ.

23 For all $u, v \in V$, define $d(u,v) = \|u - v\|$.

(a) Show that d is a metric on V.

(b) Show that if V is finite-dimensional, then d is a complete metric on V (meaning that every Cauchy sequence converges).

(c) Show that every finite-dimensional subspace of V is a closed subset of V (with respect to the metric d).

This exercise requires familiarity with metric spaces.

orthogonality at the Supreme Court

Law professor Richard Friedman presenting a case before the U.S. Supreme Court in 2010:

Mr. Friedman: I think that issue is entirely orthogonal to the issue here because the Commonwealth is acknowledging—
Chief Justice Roberts: I'm sorry. Entirely what?
Mr. Friedman: Orthogonal. Right angle. Unrelated. Irrelevant.
Chief Justice Roberts: Oh.
Justice Scalia: What was that adjective? I liked that.
Mr. Friedman: Orthogonal.
Chief Justice Roberts: Orthogonal.
Mr. Friedman: Right, right.
Justice Scalia: Orthogonal, ooh. (Laughter.)
Justice Kennedy: I knew this case presented us a problem. (Laughter.)

22 נניח ש־ $C[-1,1]$ הוא המרחב הווקטורי של פונקציות ממשיות רציפות מעל הקטע $[-1,1]$ עם מכפלה פנימית המוגדרת על ידי

$$\langle f,g\rangle = \int_{-1}^{1} fg$$

לכל $f,g \in C[-1,1]$. יהי φ הפונקציונל הליניארי מעל $C[-1,1]$ המוגדר על ידי $\varphi(f) = f(0)$. הראו שלא קיים $g \in C[-1,1]$ כך ש־

$$\varphi(f) = \langle f,g\rangle$$

לכל $f \in C[-1,1]$.

תרגיל זה מראה כי משפט ההצגה של ריס (6.42) אינו מתקיים מעל מרחב וקטורי מממד אינסופי ללא הנחות נוספות בנוגע ל־ V ו־ φ.

23 לכל $u,v \in V$, הגדירו $d(u,v) = \|u-v\|$.

א. הראו ש־ d היא מטריקה מעל V.

ב. הראו שאם V מממד סופי, אזי d מטריקה שלמה מעל V (כלומר, כל סדרת קושי היא סדרה מתכנסת).

ג. הראו שכל תת־מרחב מממד סופי של V הוא תת־קבוצה סגורה של V (ביחס למטריקה d).

תרגיל זה זורש היכרות עם מרחבים מטריים.

אורתוגונליות בבית המשפט העליון

פרופסור למשפטים ריצ'רד פרידמן מייצג תיק בפני בית המשפט העליון של ארה"ב בשנת 2010:

מר פריזמן: אני סבור שהעניין ההוא, הוא לחלוטין אורתוגונלי לסוגיה הנדונה מאחר שהמדינה כבר מכירה ב—

נשיא בית המשפט העליון השופט רוברטס: סליחה. לחלוטין מה?

מר פריזמן: אורתוגונלי. זווית ישרה. לא קשור. לא רלוונטי.

נשיא בית המשפט העליון השופט רוברטס: אההם.

השופט סקליה: מה היה שם התואר הזה? אהבתי אותו.

מר פריזמן: אורתוגונלי.

נשיא בית המשפט העליון השופט רוברטס: אורתוגונלי.

מר פריזמן: נכון, נכון.

השופט סקליה: אורתוגונלי, הו הו. (צחוק.)

השופט קנדי: ידעתי שהתיק הזה הולך להיות בעייתי. (צחוק.)

6C Orthogonal Complements and Minimization Problems

Orthogonal Complements

6.46 definition: *orthogonal complement*, $U^\perp$

If U is a subset of V, then the *orthogonal complement* of U, denoted by $U^\perp$, is the set of all vectors in V that are orthogonal to every vector in U:

$$U^\perp = \{v \in V : \langle u, v\rangle = 0 \text{ for every } u \in U\}.$$

The orthogonal complement $U^\perp$ depends on V as well as on U. However, the inner product space V should always be clear from the context and thus it can be omitted from the notation.

6.47 example: *orthogonal complements*

- If $V = \mathbf{R}^3$ and U is the subset of V consisting of the single point $(2, 3, 5)$, then $U^\perp$ is the plane $\{(x, y, z) \in \mathbf{R}^3 : 2x + 3y + 5z = 0\}$.
- If $V = \mathbf{R}^3$ and U is the plane $\{(x, y, z) \in \mathbf{R}^3 : 2x + 3y + 5z = 0\}$, then $U^\perp$ is the line $\{(2t, 3t, 5t) : t \in \mathbf{R}\}$.
- More generally, if U is a plane in $\mathbf{R}^3$ containing the origin, then $U^\perp$ is the line containing the origin that is perpendicular to U.
- If U is a line in $\mathbf{R}^3$ containing the origin, then $U^\perp$ is the plane containing the origin that is perpendicular to U.
- If $V = \mathbf{F}^5$ and $U = \{(a, b, 0, 0, 0) \in \mathbf{F}^5 : a, b \in \mathbf{F}\}$, then
$$U^\perp = \{(0, 0, x, y, z) \in \mathbf{F}^5 : x, y, z \in \mathbf{F}\}.$$
- If $e_1, \dots, e_m, f_1, \dots, f_n$ is an orthonormal basis of V, then
$$(\operatorname{span}(e_1, \dots, e_m))^\perp = \operatorname{span}(f_1, \dots, f_n).$$

We begin with some straightforward consequences of the definition.

6.48 *properties of orthogonal complement*

(a) If U is a subset of V, then $U^\perp$ is a subspace of V.

(b) $\{0\}^\perp = V$.

(c) $V^\perp = \{0\}$.

(d) If U is a subset of V, then $U \cap U^\perp \subseteq \{0\}$.

(e) If G and H are subsets of V and $G \subseteq H$, then $H^\perp \subseteq G^\perp$.

6ו משלימים אורתוגונליים ובעיות מזעור

משלימים אורתוגונליים

6.46 הגדרה: משלים אורתוגונלי, $U^\perp$

אם U תת־קבוצה של V, אזי המשלים האורתוגונלי של U, המסומן על ידי $U^\perp$, הוא קבוצת כל הווקטורים ב־ V האורתוגונליים לכל וקטור ב־ U:
$$U^\perp = \{v \in V : \langle u, v\rangle = 0 \quad \text{לכל } u \in U\}.$$

המשלים האורתוגונלי $U^\perp$ תלוי ב־ V כמו גם ב־ U. אולם, מרחב המכפלה הפנימית V אמור תמיד להיות ברור מההקשר, ולפיכך הוא מושמט מהסימון.

6.47 דוגמה: משלימים אורתוגונליים

- אם $V = \mathbf{R}^3$ ו־ U תת־קבוצה של V המורכבת מהנקודה היחידה $(2,3,5)$, אזי $U^\perp$ הוא המישור $\{(x,y,z) \in \mathbf{R}^3 : 2x + 3y + 5z = 0\}$.
- אם $V = \mathbf{R}^3$ ו־ U הוא המישור $\{(x,y,z) \in \mathbf{R}^3 : 2x + 3y + 5z = 0\}$, אזי $U^\perp$ הוא הישר $\{(2t, 3t, 5t) : t \in \mathbf{R}\}$.
- באופן כללי, אם U מישור ב־ $\mathbf{R}^3$ העובר דרך ראשית הצירים, אזי $U^\perp$ הוא הישר העובר דרך ראשית הצירים ומאונך ל־ U.
- אם U ישר ב־ $\mathbf{R}^3$ העובר דרך ראשית הצירים, אזי $U^\perp$ הוא המישור העובר דרך ראשית הצירים ומאונך ל־ U.
- אם $V = \mathbf{F}^5$ ו־ $U = \{(a,b,0,0,0) \in \mathbf{F}^5 : a, b \in \mathbf{F}\}$, אזי
$$U^\perp = \{(0,0,x,y,z) \in \mathbf{F}^5 : x, y, z \in \mathbf{F}\}.$$
- אם $e_1, \dots, e_m, f_1, \dots, f_n$ בסיס אורתונורמלי של V, אזי
$$(\operatorname{span}(e_1, \dots, e_m))^\perp = \operatorname{span}(f_1, \dots, f_n).$$

נפתח במספר השלכות ישירות הנובעות מההגדרה.

6.48 תכונות של משלים אורתוגונלי

א. אם U תת־קבוצה של V, אזי $U^\perp$ תת־מרחב של V.

ב. $\{0\}^\perp = V$.

ג. $V^\perp = \{0\}$.

ד. אם U תת־קבוצה של V, אזי $U \cap U^\perp \subseteq \{0\}$.

ה. אם G ו־ H תתי־קבוצה של V ו־ $G \subseteq H$, אזי $H^\perp \subseteq G^\perp$.

Proof

(a) Suppose U is a subset of V. Then $\langle u, 0\rangle = 0$ for every $u \in U$; thus $0 \in U^\perp$.

Suppose $v, w \in U^\perp$. If $u \in U$, then

$$\langle u, v + w\rangle = \langle u, v\rangle + \langle u, w\rangle = 0 + 0 = 0.$$

Thus $v + w \in U^\perp$, which shows that $U^\perp$ is closed under addition.

Similarly, suppose $\lambda \in \mathbf{F}$ and $v \in U^\perp$. If $u \in U$, then

$$\langle u, \lambda v\rangle = \overline{\lambda}\langle u, v\rangle = \overline{\lambda} \cdot 0 = 0.$$

Thus $\lambda v \in U^\perp$, which shows that $U^\perp$ is closed under scalar multiplication. Thus $U^\perp$ is a subspace of V.

(b) Suppose that $v \in V$. Then $\langle 0, v\rangle = 0$, which implies that $v \in \{0\}^\perp$. Thus $\{0\}^\perp = V$.

(c) Suppose that $v \in V^\perp$. Then $\langle v, v\rangle = 0$, which implies that $v = 0$. Thus $V^\perp = \{0\}$.

(d) Suppose U is a subset of V and $u \in U \cap U^\perp$. Then $\langle u, u\rangle = 0$, which implies that $u = 0$. Thus $U \cap U^\perp \subseteq \{0\}$.

(e) Suppose G and H are subsets of V and $G \subseteq H$. Suppose $v \in H^\perp$. Then $\langle u, v\rangle = 0$ for every $u \in H$, which implies that $\langle u, v\rangle = 0$ for every $u \in G$. Hence $v \in G^\perp$. Thus $H^\perp \subseteq G^\perp$. ∎

Recall that if U and W are subspaces of V, then V is the direct sum of U and W (written $V = U \oplus W$) if each element of V can be written in exactly one way as a vector in U plus a vector in W (see 1.41). Furthermore, this happens if and only if $V = U + W$ and $U \cap W = \{0\}$ (see 1.46).

The next result shows that every finite-dimensional subspace of V leads to a natural direct sum decomposition of V. See Exercise 16 for an example showing that the result below can fail without the hypothesis that the subspace U is finite-dimensional.

6.49 *direct sum of a subspace and its orthogonal complement*

Suppose U is a finite-dimensional subspace of V. Then

$$V = U \oplus U^\perp.$$

Proof First we will show that

$$V = U + U^\perp.$$

To do this, suppose that $v \in V$. Let $e_1, \ldots, e_m$ be an orthonormal basis of U. We want to write v as the sum of a vector in U and a vector orthogonal to U.

הוכחה

א. נניח ש־ U תת־קבוצה של V. אזי $\langle u,0\rangle = 0$ לכל $u \in U$; לפיכך $0 \in U^\perp$.

נניח ש־ $v,w \in U^\perp$. אם $u \in U$, אזי
$$\langle u, v+w\rangle = \langle u,v\rangle + \langle u,w\rangle = 0+0 = 0.$$
לפיכך $v+w \in U^\perp$, מה שמראה ש־ $U^\perp$ סגור תחת פעולת החיבור.

באופן דומה, נניח ש־ $\lambda \in \mathbf{F}$ ו־ $v \in U^\perp$. אם $u \in U$, אזי
$$\langle u, \lambda v\rangle = \overline{\lambda}\langle u,v\rangle = \overline{\lambda}\cdot 0 = 0.$$
לפיכך $\lambda v \in U^\perp$, מה שמראה ש־ $U^\perp$ סגור תחת פעולת כפל בסקלר. לפיכך $U^\perp$ תת־מרחב של V.

ב. נניח ש־ $v \in V$. אזי $\langle 0,v\rangle = 0$, מה שגורר ש־ $v \in \{0\}^\perp$. לפיכך $\{0\}^\perp = V$.

ג. נניח ש־ $v \in V^\perp$. אזי $\langle v,v\rangle = 0$, מה שגורר ש־ $v = 0$. לפיכך $V^\perp = \{0\}$.

ד. נניח ש־ U תת־קבוצה של V ו־ $u \in U \cap U^\perp$. אזי $\langle u,u\rangle = 0$, מה שגורר ש־ $u = 0$. לפיכך $U \cap U^\perp \subseteq \{0\}$.

ה. נניח ש־ G ו־ H תתי־קבוצה של V ו־ $G \subseteq H$. נניח ש־ $v \in H^\perp$. אזי $\langle u,v\rangle = 0$ לכל $u \in H$, מה שגורר ש־ $\langle u,v\rangle = 0$ לכל $u \in G$. לכן $v \in G^\perp$. לפיכך $H^\perp \subseteq G^\perp$. ∎

נזכיר שאם U ו־ W תתי־מרחב של V, אזי V הוא הסכום הישר של U ו־ W (נכתוב $V = U \oplus W$) אם כל איבר של V ניתן לכתיבה באופן יחיד כווקטור ב־ U ועוד וקטור ב־ W (ראו 1.41). יתר על כן, הדבר קורה אם־ורק־אם $V = U+W$ ו־ $U \cap W = \{0\}$ (ראו 1.46).

המסקנה הבאה מראה כי כל תת־מרחב מממד סופי של V מוביל לפירוק טבעי של V באמצעות סכום ישר. ראו תרגיל 16 לדוגמה הממחישה כי המסקנה להלן יכולה להיכשל ללא ההנחה כי התת־מרחב U מממד סופי.

6.49 סכום ישר של תת־מרחב והמשלים האורתוגונלי שלו

נניח ש־ U תת־מרחב של V מממד סופי. אזי
$$V = U \oplus U^\perp.$$

הוכחה תחילה נראה כי
$$V = U + U^\perp.$$
לצורך כך, נניח ש־ $v \in V$. יהי $e_1,\dots,e_m$ בסיס אורתונורמלי של U. נרצה לכתוב את v כסכום של וקטור ב־ U ווקטור אורתוגונלי ל־ U.

We have

6.50 $$v = \underbrace{\langle v, e_1\rangle e_1 + \cdots + \langle v, e_m\rangle e_m}_{u} + \underbrace{v - \langle v, e_1\rangle e_1 - \cdots - \langle v, e_m\rangle e_m}_{w}.$$

Let u and w be defined as in the equation above (as was done in the proof of 6.26). Because each $e_k \in U$, we see that $u \in U$. Because $e_1, \ldots, e_m$ is an orthonormal list, for each $k = 1, \ldots, m$ we have

$$\begin{aligned}\langle w, e_k\rangle &= \langle v, e_k\rangle - \langle v, e_k\rangle \\ &= 0.\end{aligned}$$

Thus w is orthogonal to every vector in $\operatorname{span}(e_1, \ldots, e_m)$, which shows that $w \in U^\perp$. Hence we have written $v = u + w$, where $u \in U$ and $w \in U^\perp$, completing the proof that $V = U + U^\perp$.

From 6.48(d), we know that $U \cap U^\perp = \{0\}$. Now equation $V = U + U^\perp$ implies that $V = U \oplus U^\perp$ (see 1.46). ∎

Now we can see how to compute $\dim U^\perp$ from $\dim U$.

6.51 *dimension of orthogonal complement*

Suppose V is finite-dimensional and U is a subspace of V. Then

$$\dim U^\perp = \dim V - \dim U.$$

Proof The formula for $\dim U^\perp$ follows immediately from 6.49 and 3.94. ∎

The next result is an important consequence of 6.49.

6.52 *orthogonal complement of the orthogonal complement*

Suppose U is a finite-dimensional subspace of V. Then

$$U = (U^\perp)^\perp.$$

Proof First we will show that

6.53 $$U \subseteq (U^\perp)^\perp.$$

To do this, suppose $u \in U$. Then $\langle u, w\rangle = 0$ for every $w \in U^\perp$ (by the definition of $U^\perp$). Because u is orthogonal to every vector in $U^\perp$, we have $u \in (U^\perp)^\perp$, completing the proof of 6.53.

To prove the inclusion in the other direction, suppose $v \in (U^\perp)^\perp$. By 6.49, we can write $v = u + w$, where $u \in U$ and $w \in U^\perp$. We have $v - u = w \in U^\perp$. Because $v \in (U^\perp)^\perp$ and $u \in (U^\perp)^\perp$ (from 6.53), we have $v - u \in (U^\perp)^\perp$. Thus $v - u \in U^\perp \cap (U^\perp)^\perp$, which implies that $v - u = 0$ [by 6.48(d)], which implies that $v = u$, which implies that $v \in U$. Thus $(U^\perp)^\perp \subseteq U$, which along with 6.53 completes the proof. ∎

מתקיים

$$v = \underbrace{\langle v, e_1\rangle e_1 + \cdots + \langle v, e_m\rangle e_m}_{u} + \underbrace{v - \langle v, e_1\rangle e_1 - \cdots - \langle v, e_m\rangle e_m}_{w}. \tag{6.50}$$

יהיו u ו־ w מוגדרים כפי שמופיע במשוואה לעיל (כפי שעשינו בהוכחת 6.26). מאחר שכל $e_k \in U$, אנו רואים כי $u \in U$. מאחר ש־ $e_1, \ldots, e_m$ רשימת וקטורים אורתונורמלית, לכל $k = 1, \ldots, m$ נקבל

$$\begin{aligned}\langle w, e_k\rangle &= \langle v, e_k\rangle - \langle v, e_k\rangle \\ &= 0.\end{aligned}$$

לפיכך w אורתוגונלי לכל וקטור ב־ $\operatorname{span}(e_1, \ldots, e_m)$, מה שמראה ש־ $w \in U^\perp$. לכן כתבנו $v = u + w$, כאשר $u \in U$ ו־ $w \in U^\perp$, ובכך השלמנו את ההוכחה לכך ש־ $V = U + U^\perp$.

מ־ 6.48(ד), אנו יודעים ש־ $U \cap U^\perp = \{0\}$. כעת המשוואה $V = U + U^\perp$ גוררת ש־ $V = U \oplus U^\perp$ (ראו 1.46). ■

כעת נוכל לראות איך ניתן לחשב את $\dim U^\perp$ מתוך $\dim U$.

6.51 ממד של משלים אורתוגונלי

נניח ש־ V מממד סופי ו־ U תת־מרחב של V. אזי
$$\dim U^\perp = \dim V - \dim U.$$

הוכחה הנוסחה עבור $\dim U^\perp$ נובעת ישירות מ־ 6.49 ו־ 3.94. ■

המסקנה הבאה היא תוצאה חשובה של 6.49.

6.52 משלים אורתוגונלי של המשלים האורתוגונלי

נניח ש־ U תת־מרחב של V מממד סופי. אזי
$$U = (U^\perp)^\perp.$$

הוכחה תחילה נראה כי

$$U \subseteq (U^\perp)^\perp. \tag{6.53}$$

לצורך כך, נניח ש־ $u \in U$. אזי $\langle u, w\rangle = 0$ לכל $w \in U^\perp$ (לפי הגדרת $U^\perp$). מאחר ש־ u אורתוגונלי לכל וקטור ב־ $U^\perp$, הרי ש־ $u \in (U^\perp)^\perp$, ובכך הושלמה הוכחת 6.53.

להוכחת ההכלה בכיוון השני, נניח ש־ $v \in (U^\perp)^\perp$. לפי 6.49, נוכל לכתוב $v = u + w$, כאשר $u \in U$ ו־ $w \in U^\perp$. מתקיים $v - u = w \in U^\perp$. מאחר ש־ $v \in (U^\perp)^\perp$ ו־ $u \in (U^\perp)^\perp$ (לפי 6.53), קיבלנו ש־ $v - u \in (U^\perp)^\perp$. לפיכך $v - u \in U^\perp \cap (U^\perp)^\perp$, מה שגורר ש־ $v - u = 0$ [לפי 6.48(ד)], מה שגורר ש־ $v = u$, מה שגורר ש־ $v \in U$. לפיכך $(U^\perp)^\perp \subseteq U$; נתון זה יחד עם 6.53 משלים את ההוכחה. ■

Suppose U is a subspace of V and we want to show that $U = V$. In some situations, the easiest way to do this is to show that the only vector orthogonal to U is 0, and then use the result below. For example, the result below is useful for Exercise 4.

Exercise 16(a) shows that the result below is not true without the hypothesis that U is finite-dimensional.

6.54 $U^\perp = \{0\} \iff U = V$ *(for U a finite-dimensional subspace of V)*

Suppose U is a finite-dimensional subspace of V. Then

$$U^\perp = \{0\} \iff U = V.$$

Proof First suppose $U^\perp = \{0\}$. Then by 6.52, $U = (U^\perp)^\perp = \{0\}^\perp = V$, as desired.

Conversely, if $U = V$, then $U^\perp = V^\perp = \{0\}$ by 6.48(c). ∎

We now define an operator P_U for each finite-dimensional subspace U of V.

6.55 definition: *orthogonal projection, P_U*

Suppose U is a finite-dimensional subspace of V. The *orthogonal projection* of V onto U is the operator $P_U \in \mathcal{L}(V)$ defined as follows: For each $v \in V$, write $v = u + w$, where $u \in U$ and $w \in U^\perp$. Then let $P_U v = u$.

The direct sum decomposition $V = U \oplus U^\perp$ given by 6.49 shows that each $v \in V$ can be uniquely written in the form $v = u + w$ with $u \in U$ and $w \in U^\perp$. Thus $P_U v$ is well defined. See the figure that accompanies the proof of 6.61 for the picture describing $P_U v$ that you should keep in mind.

6.56 example: *orthogonal projection onto one-dimensional subspace*

Suppose $u \in V$ with $u \neq 0$ and U is the one-dimensional subspace of V defined by $U = \operatorname{span}(u)$.

If $v \in V$, then

$$v = \frac{\langle v, u\rangle}{\|u\|^2} u + \left(v - \frac{\langle v, u\rangle}{\|u\|^2} u\right),$$

where the first term on the right is in $\operatorname{span}(u)$ (and thus is in U) and the second term on the right is orthogonal to u (and thus is in $U^\perp$). Thus $P_U v$ equals the first term on the right. In other words, we have the formula

$$P_U v = \frac{\langle v, u\rangle}{\|u\|^2} u$$

for every $v \in V$.

The formula above becomes $P_U u = u$ if $v = u$ and becomes $P_U v = 0$ if $v \in \{u\}^\perp$. These equations are special cases of (b) and (c) in the next result.

נניח ש־ U תת־מרחב של V ואנו מעוניינים להראות כי $U = V$. במצבים מסוימים, הדרך הקלה ביותר לעשות זאת היא להראות כי הווקטור היחידי האורתוגונלי ל־ U הוא וקטור ה־ 0, ואז להשתמש במסקנה להלן. לדוגמה, המסקנה להלן שימושית עבור תרגיל 4.

תרגיל 16(א) מראה כי המסקנה להלן אינה נכונה ללא ההנחה כי U מממד סופי.

6.54 $U = V \iff U^\perp = \{0\}$ (עבור U תת־מרחב של V מממד סופי)

נניח ש־ U תת־מרחב של V מממד סופי. אזי
$$U^\perp = \{0\} \iff U = V.$$

הוכחה תחילה נניח ש־ $U^\perp = \{0\}$. מכאן, לפי 6.52, $U = (U^\perp)^\perp = \{0\}^\perp = V$, כפי שרצינו.

בכיוון ההפוך, אם $U = V$, אזי $U^\perp = V^\perp = \{0\}$ לפי 6.48(ג). ■

כעת אנו מגדירים אופרטור P_U לכל תת־מרחב U של V מממד סופי.

6.55 הגדרה: היטל אורתוגונלי, P_U

נניח ש־ U תת־מרחב של V מממד סופי. ההיטל האורתוגונלי של V על U הוא האופרטור $P_U \in \mathcal{L}(V)$ המוגדר על ידי: לכל $v \in V$, נכתוב $v = u + w$, כאשר $u \in U$ ו־ $w \in U^\perp$ ואז נגדיר $P_U v = u$.

הפירוק באמצעות סכום ישר $V = U \oplus U^\perp$ שקיבלנו ב־ 6.49 מראה שכל $v \in V$ יכול להיכתב באופן יחיד בצורה $v = u + w$ עם $u \in U$ ו־ $w \in U^\perp$. לפיכך $P_U v$ מוגדר היטב. ראו את השרטוט המצורף להוכחה של 6.61 לתמונה המתארת את $P_U v$ שעליכם לנצור בתודעתכם.

6.56 דוגמה: היטל אורתוגונלי על תת־מרחב חד־ממדי

נניח ש־ $u \in V$ עם $u \neq 0$ ו־ U תת־מרחב חד־ממדי של V המוגדר על ידי $U = \operatorname{span}(u)$.

אם $v \in V$, אזי
$$v = \frac{\langle v, u\rangle}{\|u\|^2} u + \left(v - \frac{\langle v, u\rangle}{\|u\|^2} u\right),$$
כאשר הביטוי הראשון באגף הימני הוא ב־ $\operatorname{span}(u)$ (ולפיכך ב־ U) והביטוי השני באגף הימני הוא אורתוגונלי ל־ u (ולפיכך ב־ $U^\perp$). לפיכך $P_U v$ שווה לביטוי הראשון באגף הימני. במילים אחרות, בידינו הנוסחה
$$P_U v = \frac{\langle v, u\rangle}{\|u\|^2} u$$
לכל $v \in V$.

הנוסחה לעיל הופכת ל־ $P_U u = u$ אם $v = u$ והופכת ל־ $P_U v = 0$ אם $v \in \{u\}^\perp$. משוואות אלו הן מקרים פרטיים של (ב) ו־ (ג) במסקנה הבאה.

6.57 *properties of orthogonal projection P_U*

Suppose U is a finite-dimensional subspace of V. Then

(a) $P_U \in \mathcal{L}(V)$;

(b) $P_U u = u$ for every $u \in U$;

(c) $P_U w = 0$ for every $w \in U^\perp$;

(d) range $P_U = U$;

(e) null $P_U = U^\perp$;

(f) $v - P_U v \in U^\perp$ for every $v \in V$;

(g) $P_U^2 = P_U$;

(h) $\|P_U v\| \le \|v\|$ for every $v \in V$;

(i) if $e_1, \dots, e_m$ is an orthonormal basis of U and $v \in V$, then

$$P_U v = \langle v, e_1 \rangle e_1 + \cdots + \langle v, e_m \rangle e_m.$$

Proof

(a) To show that P_U is a linear map on V, suppose $v_1, v_2 \in V$. Write

$$v_1 = u_1 + w_1 \quad \text{and} \quad v_2 = u_2 + w_2$$

with $u_1, u_2 \in U$ and $w_1, w_2 \in U^\perp$. Thus $P_U v_1 = u_1$ and $P_U v_2 = u_2$. Now

$$v_1 + v_2 = (u_1 + u_2) + (w_1 + w_2),$$

where $u_1 + u_2 \in U$ and $w_1 + w_2 \in U^\perp$. Thus

$$P_U(v_1 + v_2) = u_1 + u_2 = P_U v_1 + P_U v_2.$$

Similarly, suppose $\lambda \in \mathbf{F}$ and $v \in V$. Write $v = u + w$, where $u \in U$ and $w \in U^\perp$. Then $\lambda v = \lambda u + \lambda w$ with $\lambda u \in U$ and $\lambda w \in U^\perp$. Thus $P_U(\lambda v) = \lambda u = \lambda P_U v$.

Hence P_U is a linear map from V to V.

(b) Suppose $u \in U$. We can write $u = u + 0$, where $u \in U$ and $0 \in U^\perp$. Thus $P_U u = u$.

(c) Suppose $w \in U^\perp$. We can write $w = 0 + w$, where $0 \in U$ and $w \in U^\perp$. Thus $P_U w = 0$.

(d) The definition of P_U implies that range $P_U \subseteq U$. Furthermore, (b) implies that $U \subseteq$ range P_U. Thus range $P_U = U$.

(e) The inclusion $U^\perp \subseteq$ null P_U follows from (c). To prove the inclusion in the other direction, note that if $v \in$ null P_U then the decomposition given by 6.49 must be $v = 0 + v$, where $0 \in U$ and $v \in U^\perp$. Thus null $P_U \subseteq U^\perp$.

6.57 מאפיינים של ההיטל האורתוגונלי P_U

נניח ש־ U תת־מרחב של V מממד סופי. אזי

א. $P_U \in \mathcal{L}(V)$.

ב. $P_U u = u$ לכל $u \in U$.

ג. $P_U w = 0$ לכל $w \in U^\perp$.

ד. $\operatorname{range} P_U = U$.

ה. $\operatorname{null} P_U = U^\perp$.

ו. $v - P_U v \in U^\perp$ לכל $v \in V$.

ז. $P_U^2 = P_U$.

ח. $\|P_U v\| \le \|v\|$ לכל $v \in V$.

ט. אם $e_1, \dots, e_m$ בסיס אורתונורמלי של U ו־ $v \in V$, אזי
$P_U v = \langle v, e_1 \rangle e_1 + \cdots + \langle v, e_m \rangle e_m$.

הוכחה

א. כדי להראות ש־ P_U העתקה ליניארית מעל V, נניח ש־ $v_1, v_2 \in V$. נכתוב
$$v_1 = u_1 + w_1 \quad \text{וגם} \quad v_2 = u_2 + w_2$$
עם $u_1, u_2 \in U$ ו־ $w_1, w_2 \in U^\perp$. לפיכך $P_U v_1 = u_1$ ו־ $P_U v_2 = u_2$. כעת
$$v_1 + v_2 = (u_1 + u_2) + (w_1 + w_2),$$
כאשר $u_1 + u_2 \in U$ ו־ $w_1 + w_2 \in U^\perp$. לפיכך
$$P_U(v_1 + v_2) = u_1 + u_2 = P_U v_1 + P_U v_2.$$

באופן דומה, נניח ש־ $\lambda \in \mathbf{F}$ ו־ $v \in V$. נכתוב $v = u + w$, כאשר $u \in U$ ו־ $w \in U^\perp$. אזי $\lambda v = \lambda u + \lambda w$ עם $\lambda u \in U$ ו־ $\lambda w \in U^\perp$. לפיכך $P_U(\lambda v) = \lambda u = \lambda P_U v$.

לכן P_U העתקה ליניארית מ־ V אל V.

ב. נניח ש־ $u \in U$. נוכל לכתוב $u = u + 0$, כאשר $u \in U$ ו־ $0 \in U^\perp$. לפיכך $P_U u = u$.

ג. נניח ש־ $w \in U^\perp$. נוכל לכתוב $w = 0 + w$, כאשר $0 \in U$ ו־ $w \in U^\perp$. לפיכך $P_U w = 0$.

ד. ההגדרה של P_U גוררת ש־ $\operatorname{range} P_U \subseteq U$. יתר על כן, (ב) גוררת ש־ $U \subseteq \operatorname{range} P_U$. לפיכך $\operatorname{range} P_U = U$.

ה. ההכלה $U^\perp \subseteq \operatorname{null} P_U$ נובעת מ־ (ג). להוכחת ההכלה בכיוון ההפוך, נשים לב כי אם $v \in \operatorname{null} P_U$ אזי הפירוק שקיבלנו על ידי 6.49 חייב להיות $v = 0 + v$, כאשר $0 \in U$ ו־ $v \in U^\perp$. לפיכך $\operatorname{null} P_U \subseteq U^\perp$.

(f) If $v \in V$ and $v = u + w$ with $u \in U$ and $w \in U^\perp$, then

$$v - P_U v = v - u = w \in U^\perp.$$

(g) If $v \in V$ and $v = u + w$ with $u \in U$ and $w \in U^\perp$, then

$$(P_U^2)v = P_U(P_U v) = P_U u = u = P_U v.$$

(h) If $v \in V$ and $v = u + w$ with $u \in U$ and $w \in U^\perp$, then

$$\|P_U v\|^2 = \|u\|^2 \le \|u\|^2 + \|w\|^2 = \|v\|^2,$$

where the last equality comes from the Pythagorean theorem.

(i) The formula for $P_U v$ follows from equation 6.50 in the proof of 6.49. ∎

In the previous section we proved the Riesz representation theorem (6.42), whose key part states that every linear functional on a finite-dimensional inner product space is given by taking the inner product with some fixed vector. Seeing a different proof often provides new insight. Thus we now give a new proof of the key part of the Riesz representation theorem using orthogonal complements instead of orthonormal bases as in our previous proof.

The restatement below of the Riesz representation theorem provides an identification of V with V'. We will prove only the "onto" part of the result below because the more routine "one-to-one" part of the result can be proved as in 6.42.

Intuition behind this new proof: If $\varphi \in V'$, $v \in V$, and $\varphi(u) = \langle u, v\rangle$ for all $u \in V$, then $v \in (\operatorname{null}\varphi)^\perp$. However, $(\operatorname{null}\varphi)^\perp$ is a one-dimensional subspace of V (except for the trivial case in which $\varphi = 0$), as follows from 6.51 and 3.21. Thus we can obtain v by choosing any nonzero element of $(\operatorname{null}\varphi)^\perp$ and then multiplying by an appropriate scalar, as is done in the proof below.

6.58 *Riesz representation theorem, revisited*

Suppose V is finite-dimensional. For each $v \in V$, define $\varphi_v \in V'$ by

$$\varphi_v(u) = \langle u, v\rangle$$

for each $u \in V$. Then $v \mapsto \varphi_v$ is a one-to-one function from V onto V'.

Proof To show that $v \mapsto \varphi_v$ is surjective, suppose $\varphi \in V'$. If $\varphi = 0$, then $\varphi = \varphi_0$. Thus assume $\varphi \ne 0$. Hence $\operatorname{null}\varphi \ne V$, which implies that $(\operatorname{null}\varphi)^\perp \ne \{0\}$ (by 6.49 with $U = \operatorname{null}\varphi$).

Caution: *The function $v \mapsto \varphi_v$ is a linear mapping from V to V' if $\mathbf{F} = \mathbf{R}$. However, this function is not linear if $\mathbf{F} = \mathbf{C}$ because $\varphi_{\lambda v} = \overline{\lambda}\varphi_v$ if $\lambda \in \mathbf{C}$.*

Let $w \in (\operatorname{null}\varphi)^\perp$ be such that $w \ne 0$. Let

$$v = \frac{\overline{\varphi(w)}}{\|w\|^2} w. \tag{6.59}$$

Then $v \in (\operatorname{null}\varphi)^\perp$. Also, $v \ne 0$ (because $w \notin \operatorname{null}\varphi$).

ו. אם $v \in V$ ו־ $v = u + w$ עם $u \in U$ ו־ $w \in U^\perp$, אזי
$v - P_U v = v - u = w \in U^\perp$.

ז. אם $v \in V$ ו־ $v = u + w$ עם $u \in U$ ו־ $w \in U^\perp$, אזי
$(P_U^2)v = P_U(P_U v) = P_U u = u = P_U v$.

ח. אם $v \in V$ ו־ $v = u + w$ עם $u \in U$ ו־ $w \in U^\perp$, אזי
$\|P_U v\|^2 = \|u\|^2 \le \|u\|^2 + \|w\|^2 = \|v\|^2$,
כאשר השוויון האחרון מגיע ממשפט פיתגורס.

ט. הנוסחה עבור $P_U v$ נובעת ממשוואה 6.50 בהוכחה של 6.49. ■

ביחידה הקודמת הוכחנו את משפט ההצגה של ריס (6.42), אשר חלקו העיקרי קובע כי כל פונקציונל ליניארי מעל מרחב מכפלה פנימית מממד סופי נתון על ידי המכפלה הפנימית עם וקטור קבוע כלשהו. היכרות עם הוכחה שונה לעיתים קרובות מספקת תובנות חדשות. לפיכך נציג כעת הוכחה נוספת לחלקו העיקרי של משפט ההצגה של ריס תוך שימוש במשלימים אורתוגונליים במקום בבסיסים אורתונורמליים כפי שעשינו בהוכחה הקודמת.

הניסוח מחדש של משפט ההצגה של ריס המובא להלן מספק זיהוי של V עם V'. אנו נוכיח רק את חלק ה"על" של המסקנה להלן מאחר שחלק ה"חד־חד־ערכית" השגרתי יותר ניתן להוכחה כמו ב־ 6.42.

אינטואיציה מאחורי הוכחה זו: אם $v \in V$, $\varphi \in V'$, ו־ $\varphi(u) = \langle u, v\rangle$ לכל $u \in V$, אזי $v \in (\operatorname{null}\varphi)^\perp$. אולם, $(\operatorname{null}\varphi)^\perp$ הוא תת־מרחב חד־ממדי של V (מלבד המקרה הטריוויאלי של $\varphi = 0$), כפי שנובע מ־ 6.51 ו־ 3.21. לפיכך נוכל לקבל את v בבחירת כל איבר השונה מאפס ב־ $(\operatorname{null}\varphi)^\perp$ ואז הכפלתו בסקלר המתאים, כפי שמבוצע בהוכחה להלן.

6.58 משפט ההצגה של ריס, היבט חדש

נניח ש־ V מממד סופי. לכל $v \in V$, נגדיר $\varphi_v \in V'$ על ידי
$$\varphi_v(u) = \langle u, v\rangle$$
לכל $u \in V$. אזי $v \mapsto \varphi_v$ היא פונקציה חד־חד־ערכית מ־ V על V'.

הוכחה כדי להראות ש־ $v \mapsto \varphi_v$ על, נניח ש־ $\varphi \in V'$. אם $\varphi = 0$, אזי $\varphi = \varphi_0$. לפיכך נניח ש־ $\varphi \ne 0$. לכן $\operatorname{null}\varphi \ne V$, מה שגורר ש־ $(\operatorname{null}\varphi)^\perp \ne \{0\}$ (לפי 6.49 עם $U = \operatorname{null}\varphi$).

יהי $w \in (\operatorname{null}\varphi)^\perp$ כך ש־ $w \ne 0$. נגדיר

$$v = \frac{\overline{\varphi(w)}}{\|w\|^2} w. \tag{6.59}$$

אזי $v \in (\operatorname{null}\varphi)^\perp$. בנוסף, $v \ne 0$ (מאחר ש־ $w \notin \operatorname{null}\varphi$).

זְהִירוּת: הפונקציה $v \mapsto \varphi_v$ היא העתקה ליניארית מ־ V אל V' אם $\mathbf{F} = \mathbf{R}$. אולם, פונקציה זו אינה ליניארית אם $\mathbf{F} = \mathbf{C}$ שכן $\varphi_{\lambda v} = \overline{\lambda}\varphi_v$ אם $\lambda \in \mathbf{C}$.

Taking the norm of both sides of 6.59 gives

6.60 $$\|v\| = \frac{|\varphi(w)|}{\|w\|}.$$

Applying φ to both sides of 6.59 and then using 6.60, we have

$$\varphi(v) = \frac{|\varphi(w)|^2}{\|w\|^2} = \|v\|^2.$$

Now suppose $u \in V$. Using the equation above, we have

$$u = \left(u - \frac{\varphi(u)}{\varphi(v)}v\right) + \frac{\varphi(u)}{\|v\|^2}v.$$

The first term in parentheses above is in null φ and hence is orthogonal to v. Thus taking the inner product of both sides of the equation above with v shows that

$$\langle u, v\rangle = \frac{\varphi(u)}{\|v\|^2}\langle v, v\rangle = \varphi(u).$$

Thus $\varphi = \varphi_v$, showing that $v \mapsto \varphi_v$ is surjective, as desired. ∎

See Exercise 13 for yet another proof of the Riesz representation theorem.

Minimization Problems

The following problem often arises: Given a subspace U of V and a point $v \in V$, find a point $u \in U$ such that $\|v - u\|$ is as small as possible. The next result shows that $u = P_U v$ is the unique solution of this minimization problem.

The remarkable simplicity of the solution to this minimization problem has led to many important applications of inner product spaces outside of pure mathematics.

6.61 *minimizing distance to a subspace*

Suppose U is a finite-dimensional subspace of V, $v \in V$, and $u \in U$. Then

$$\|v - P_U v\| \le \|v - u\|.$$

Furthermore, the inequality above is an equality if and only if $u = P_U v$.

Proof We have

6.62
$$\begin{aligned}\|v - P_U v\|^2 &\le \|v - P_U v\|^2 + \|P_U v - u\|^2\\ &= \|(v - P_U v) + (P_U v - u)\|^2\\ &= \|v - u\|^2,\end{aligned}$$

חישוב הנורמה של שני האגפים של 6.59 נותן לנו

$$\|v\| = \frac{|\varphi(w)|}{\|w\|}. \tag{6.60}$$

הפעלת φ על שני האגפים של 6.59 ואז שימוש ב־ 6.60 נותנים לנו

$$\varphi(v) = \frac{|\varphi(w)|^2}{\|w\|^2} = \|v\|^2.$$

כעת נניח ש־ $u \in V$. בשימוש במשוואה לעיל, נקבל

$$u = \left(u - \frac{\varphi(u)}{\varphi(v)}v\right) + \frac{\varphi(u)}{\|v\|^2}v.$$

הביטוי הראשון בסוגריים לעיל הוא ב־ φ null ולכן הוא אורתוגונלי ל־ v. לפיכך חישוב מכפלה פנימית של שני אגפי המשוואה לעיל עם v מראה ש־

$$\langle u, v\rangle = \frac{\varphi(u)}{\|v\|^2}\langle v, v\rangle = \varphi(u).$$

לפיכך $\varphi = \varphi_v$, מה שמראה כי $v \mapsto \varphi_v$ על, כפי שרצינו. ■

ראו תרגיל 13 להוכחה נוספת למשפט ההצגה של ריס.

בעיות מזעור

הבעיה להלן עולה לעיתים קרובות: בהינתן תת־מרחב U של V ונקודה $v \in V$, מצאו נקודה $u \in U$ כך ש־ $\|v - u\|$ קטן ככל האפשר. המסקנה הבאה מראה כי $u = P_U v$ הוא הפתרון היחיד לבעיית המזעור הזו.

הפשטות הראויה לציון של הפתרון לבעיית מזעור זו הובילה ליישומים חשובים רבים של מרחבי מכפלה פנימית מחוץ לגבולות המתמטיקה העיונית.

6.61 מזעור מרחק לתת־מרחב

נניח ש־ U תת־מרחב של V מממד סופי, $v \in V$, ו־ $u \in U$. אזי

$$\|v - P_U v\| \le \|v - u\|.$$

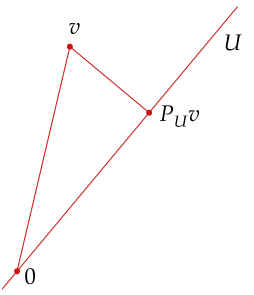
יתר על כן, האי־שוויון לעיל הוא שוויון אם־ורק־אם $u = P_U v$.

הוכחה מתקיים

$$\begin{aligned}\|v - P_U v\|^2 &\le \|v - P_U v\|^2 + \|P_U v - u\|^2 \\ &= \|(v - P_U v) + (P_U v - u)\|^2 \\ &= \|v - u\|^2,\end{aligned} \tag{6.62}$$

where the first line above holds because $0 \le \|P_U v - u\|^2$, the second line above comes from the Pythagorean theorem [which applies because $v - P_U v \in U^\perp$ by 6.57(f), and $P_U v - u \in U$], and the third line above holds by simple algebra. Taking square roots gives the desired inequality.

The inequality proved above is an equality if and only if 6.62 is an equality, which happens if and only if $\|P_U v - u\| = 0$, which happens if and only if $u = P_U v$. ∎



$P_U v$ is the closest point in U to v.

The last result is often combined with the formula 6.57(i) to compute explicit solutions to minimization problems, as in the following example.

6.63 example: *using linear algebra to approximate the sine function*

Suppose we want to find a polynomial u with real coefficients and of degree at most 5 that approximates the sine function as well as possible on the interval $[-\pi, \pi]$, in the sense that

$$\int_{-\pi}^{\pi} |\sin x - u(x)|^2 \, dx$$

is as small as possible.

Let $C[-\pi, \pi]$ denote the real inner product space of continuous real-valued functions on $[-\pi, \pi]$ with inner product

6.64 $$\langle f, g \rangle = \int_{-\pi}^{\pi} fg.$$

Let $v \in C[-\pi, \pi]$ be the function defined by $v(x) = \sin x$. Let U denote the subspace of $C[-\pi, \pi]$ consisting of the polynomials with real coefficients and of degree at most 5. Our problem can now be reformulated as follows:

Find $u \in U$ such that $\|v - u\|$ is as small as possible.

A computer that can integrate is useful here.

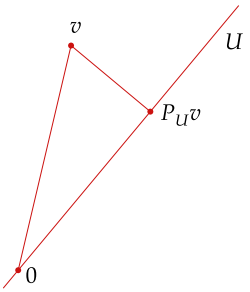
To compute the solution to our approximation problem, first apply the Gram–Schmidt procedure (using the inner product given by 6.64) to the basis $1, x, x^2, x^3, x^4, x^5$ of U, producing an orthonormal basis $e_1, e_2, e_3, e_4, e_5, e_6$ of U. Then, again using the inner product given by 6.64, compute $P_U v$ using 6.57(i) (with $m = 6$). Doing this computation shows that $P_U v$ is the function u defined by

6.65 $$u(x) = 0.987862x - 0.155271x^3 + 0.00564312x^5,$$

where the π's that appear in the exact answer have been replaced with a good decimal approximation. By 6.61, the polynomial u above is the best approximation to the sine function on $[-\pi, \pi]$ using polynomials of degree at most 5 (here "best approximation" means in the sense of minimizing $\int_{-\pi}^{\pi} |\sin x - u(x)|^2 \, dx$).

כאשר השורה הראשונה לעיל מתקיימת מאחר ש־ $0 \le \|P_U v - u\|^2$, השורה השנייה לעיל נובעת ממשפט פיתגורס [הניתן ליישום מאחר ש־ $v - P_U v \in U^\perp$ לפי 6.57(ו), ו־ $P_U v - u \in U$], והשורה השלישית מתקיימת על ידי אלגברה פשוטה. הוצאת שורשים מעניקה לנו את התוצאה המבוקשת.

האי־שוויון שהוכחנו לעיל הוא שוויון אם־ורק־אם 6.62 שוויון, שקורה אם־ורק־אם $\|P_U v - u\| = 0$, שקורה אם־ורק־אם $u = P_U v$. ∎



$P_U v$ הנקודה הקרובה ביותר ב־ U ל־ v.

המסקנה האחרונה משולבת לעיתים קרובות עם נוסחה 6.57(ט) לחישוב פתרונות מפורשים לבעיות מזעור, כמובא בדוגמה להלן.

6.63 דוגמה: שימוש באלגברה ליניארית לחישוב קירובים עבור הפונקציה סינוס

נניח שאנו מעוניינים למצוא פולינום u עם מקדמים ממשיים שמעלתו לכל היותר 5 שמספק את הקירוב הטוב ביותר האפשרי לפונקציית הסינוס מעל הקטע $[-\pi, \pi]$, במובן ש־

$$\int_{-\pi}^{\pi} |\sin x - u(x)|^2 \, dx$$

קטן ככל האפשר.

נסמן ב־ $C[-\pi, \pi]$ את מרחב המכפלה הפנימית הממשי של פונקציות ממשיות רציפות מעל $[-\pi, \pi]$ עם מכפלה פנימית

$$\langle f, g \rangle = \int_{-\pi}^{\pi} fg. \tag{6.64}$$

נגדיר את $v \in C[-\pi, \pi]$ להיות הפונקציה המוגדרת על ידי $v(x) = \sin x$. נסמן ב־ U את התת־מרחב של $C[-\pi, \pi]$ המורכב מהפולינומים עם מקדמים ממשיים שמעלתם 5 לכל היותר. כעת, נוכל לנסח מחדש את הבעיה שלנו באופן הבא:

מצאו $u \in U$ כך ש־ $\|v - u\|$ קטן ככל האפשר.

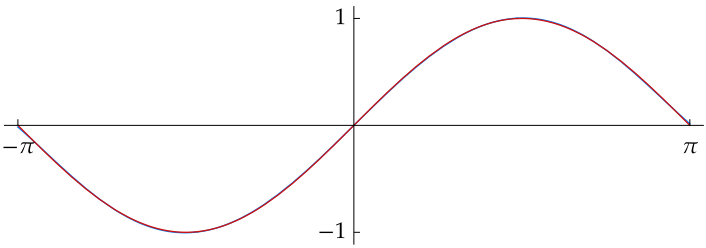
תוכנת מחשב המסוגלת לחשב אינטגרציה תהיה שימושית כאן.

לחישוב הפתרון לבעיית הקירוב שלנו, תחילה ניישם את תהליך גרם־שמידט (בשימוש במכפלה הפנימית שנתונה ב־ 6.64) על הבסיס $1, x, x^2, x^3, x^4, x^5$ של U, להפקת בסיס אורתונורמלי $e_1, e_2, e_3, e_4, e_5, e_6$ של U. בשלב זה, תוך שימוש חוזר במכפלה הפנימית הנתונה ב־ 6.64, נחשב את $P_U v$ תוך שימוש ב־ 6.57(ט) (עם $m = 6$). ביצוע חישוב זה מראה ש־ $P_U v$ היא הפונקציה u המוגדרת על ידי

$$u(x) = 0.987862x - 0.155271x^3 + 0.00564312x^5, \tag{6.65}$$

כאשר ה־ π־ים המופיעים בנוסחה המדויקת הוחלפו בקירוב עשרוני טוב. לפי 6.61, הפולינום u הוא הקירוב הטוב ביותר לפונקציית הסינוס מעל הקטע $[-\pi, \pi]$ תוך שימוש בפולינומים שמעלתם 5 לכל היותר ("הקירוב הטוב ביותר" המוזכר כאן הוא במובן של מזעור $\int_{-\pi}^{\pi} |\sin x - u(x)|^2 \, dx$).

To see how good this approximation is, the next figure shows the graphs of both the sine function and our approximation u given by 6.65 over the interval $[-\pi, \pi]$.



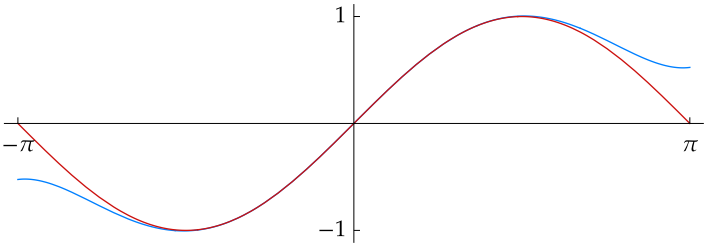
Graphs on $[-\pi, \pi]$ of the sine function (red) and its best fifth degree polynomial approximation u (blue) from 6.65.

Our approximation 6.65 is so accurate that the two graphs are almost identical—our eyes may see only one graph! Here the red graph is placed almost exactly over the blue graph. If you are viewing this on an electronic device, enlarge the picture above by 400% near π or $-\pi$ to see a small gap between the two graphs.

Another well-known approximation to the sine function by a polynomial of degree 5 is given by the Taylor polynomial p defined by

6.66
$$p(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!}.$$

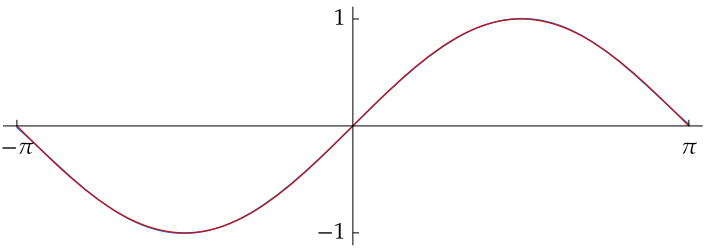
To see how good this approximation is, the next picture shows the graphs of both the sine function and the Taylor polynomial p over the interval $[-\pi, \pi]$.



Graphs on $[-\pi, \pi]$ of the sine function (red) and the Taylor polynomial (blue) from 6.66.

The Taylor polynomial of degree 5 is an excellent approximation to $\sin x$ for x near 0. But the picture above shows that for $|x| > 2$, the Taylor polynomial is not so accurate, especially compared to 6.65. For example, taking $x = 3$, our approximation 6.65 estimates $\sin 3$ with an error of approximately 0.001, but the Taylor polynomial 6.66 estimates $\sin 3$ with an error of approximately 0.4. Thus at $x = 3$, the error in the Taylor polynomial is hundreds of times larger than the error given by 6.65. Linear algebra has helped us discover an approximation to the sine function that improves upon what we learned in calculus!

כדי להיווכח עד כמה טוב הקירוב הזה, האיור הבא מציג את הגרפים של פונקציית הסינוס ושל הקירוב שלנו u הנתונה על ידי 6.65 מעל הקטע $[-\pi, \pi]$.



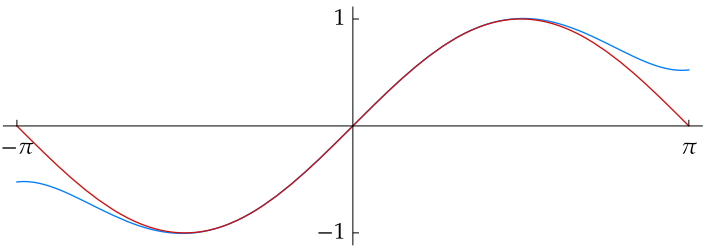
גרפים מעל $[-\pi, \pi]$ של פונקציית הסינוס (אדום) והקירוב ממעלה חמישית הטוב ביותר שלה u (כחול) מ־ 6.65.

הקירוב 6.65 שלנו הוא מדויק כל כך עד שהגרפים כמעט זהים—עינינו עשויות לראות גרף אחד בלבד! הגרף האדום מונח כמעט במדויק מעל הגרף הכחול. אם אתם צופים בכך באמצעות מכשיר אלקטרוני, הגדילו את תצוגת התמונה לעיל ב־ 400% ליד π או $-\pi$ כדי לחזות בפער קטן בין שני הגרפים.

קירוב ידוע ומוכר אחר לפונקציית הסינוס באמצעות פולינום ממעלה 5 מתקבל מפולינום טיילור p המוגדר על ידי

$$p(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!}. \tag{6.66}$$

כדי להיווכח עד כמה טוב קירוב זה, האיור הבא מציג את הגרפים של פונקציית הסינוס ושל פולינום טיילור p מעל הקטע $[-\pi, \pi]$.



גרף מעל $[-\pi, \pi]$ של פונקציית הסינוס (אדום)
ופולינום טיילור (כחול) מ־ 6.66.

פולינום טיילור ממעלה 5 הוא קירוב מצוין ל־ $\sin x$ עבור x הקרוב ל־ 0. אולם מהתמונה לעיל אנו רואים כי עבור $|x| > 2$, פולינום טיילור אינו מדויק עד כדי כך, בייחוד בהשוואה ל־ 6.65. לדוגמה, עבור $x = 3$, הקירוב שלנו 6.65 מעריך את $\sin 3$ עם סטייה של כ־ 0.001, אך פולינום טיילור 6.66 מעריך את $\sin 3$ עם סטייה של כ־ 0.4. לפיכך עבור $x = 3$, סטייה של פולינום טיילור גדולה יותר במאות פעמים מהסטייה בתוצאה שקיבלנו על ידי 6.65. אלגברה ליניארית סייעה לנו לגלות קירוב לפונקציית הסינוס משופר מזה שלמדנו בחשבון אינפיניטסימלי!

אלגברה ליניארית בדרך הישר, מהדורה רביעית, מחבר: שלדון אקסלר
Linear Algebra Done Right, fourth edition, by Sheldon Axler

Pseudoinverse

Suppose $T \in \mathcal{L}(V, W)$ and $w \in W$. Consider the problem of finding $v \in V$ such that

$$Tv = w.$$

For example, if $V = \mathbf{F}^n$ and $W = \mathbf{F}^m$, then the equation above could represent a system of m linear equations in n unknowns $v_1, \dots, v_n$, where $v = (v_1, \dots, v_n)$.

If T is invertible, then the unique solution to the equation above is $v = T^{-1}w$. However, if T is not invertible, then for some $w \in W$ there may not exist any solutions of the equation above, and for some $w \in W$ there may exist infinitely many solutions of the equation above.

If T is not invertible, then we can still try to do as well as possible with the equation above. For example, if the equation above has no solutions, then instead of solving the equation $Tv - w = 0$, we can try to find $v \in V$ such that $\|Tv - w\|$ is as small as possible. As another example, if the equation above has infinitely many solutions $v \in V$, then among all those solutions we can try to find one such that $\|v\|$ is as small as possible.

The pseudoinverse will provide the tool to solve the equation above as well as possible, even when T is not invertible. We need the next result to define the pseudoinverse.

In the next two proofs, we will use without further comment the result that if V is finite-dimensional and $T \in \mathcal{L}(V, W)$, then $\operatorname{null} T$, $(\operatorname{null} T)^\perp$, and $\operatorname{range} T$ are all finite-dimensional.

6.67 *restriction of a linear map to obtain a one-to-one and onto map*

Suppose V is finite-dimensional and $T \in \mathcal{L}(V, W)$. Then $T|_{(\operatorname{null} T)^\perp}$ is an injective map of $(\operatorname{null} T)^\perp$ onto $\operatorname{range} T$.

Proof Suppose that $v \in (\operatorname{null} T)^\perp$ and $T|_{(\operatorname{null} T)^\perp} v = 0$. Hence $Tv = 0$ and thus $v \in (\operatorname{null} T) \cap (\operatorname{null} T)^\perp$, which implies that $v = 0$ [by 6.48(d)]. Hence $\operatorname{null}(T|_{(\operatorname{null} T)^\perp}) = \{0\}$, which implies that $T|_{(\operatorname{null} T)^\perp}$ is injective, as desired.

Clearly $\operatorname{range}(T|_{(\operatorname{null} T)^\perp}) \subseteq \operatorname{range} T$. To prove the inclusion in the other direction, suppose $w \in \operatorname{range} T$. Hence there exists $v \in V$ such that $w = Tv$. There exist $u \in \operatorname{null} T$ and $x \in (\operatorname{null} T)^\perp$ such that $v = u + x$ (by 6.49). Now

$$T|_{(\operatorname{null} T)^\perp} x = Tx = Tv - Tu = w - 0 = w,$$

which shows that $w \in \operatorname{range} T|_{(\operatorname{null} T)^\perp}$. Hence $\operatorname{range} T \subseteq \operatorname{range} T|_{(\operatorname{null} T)^\perp}$, completing the proof that $\operatorname{range} T|_{(\operatorname{null} T)^\perp} = \operatorname{range} T$. ∎

Now we can define the pseudoinverse $T^\dagger$ (pronounced "T dagger") of a linear map T. In the next definition (and from now on), think of $T|_{(\operatorname{null} T)^\perp}$ as an invertible linear map from $(\operatorname{null} T)^\perp$ onto $\operatorname{range} T$, as is justified by the result above.

To produce the pseudoinverse notation $T^\dagger$ in TEX, type `T^\dagger`.

פְּסֵידוֹ־הופכי

נניח ש־ $T \in \mathcal{L}(V,W)$ ו־ $w \in W$. נדון בבעיית מציאת $v \in V$ כך ש־

$$Tv = w.$$

לדוגמה, אם $V = \mathbf{F}^n$ ו־ $W = \mathbf{F}^m$, אזי המשוואה לעיל יכולה לייצג מערכת של m משוואות ליניאריות ב־ n נעלמים $v_1, \dots, v_n$, כאשר $v = (v_1, \dots, v_n)$.

אם T הפיכה, אזי הפתרון היחיד של המשוואה לעיל הוא $v = T^{-1}w$. אולם, אם T אינה הפיכה, אזי עבור $w \in W$ כלשהו ייתכן שלמשוואה לעיל לא קיימים פתרונות, בעוד שעבור $w \in W$ אחר ייתכן שלמשוואה לעיל קיימים אינסוף פתרונות.

אם T אינה הפיכה, אזי עדיין נוכל לנסות להפיק את המיטב מן המשוואה לעיל. לדוגמה, אם למשוואה לעיל אין פתרונות, אזי במקום לפתור את המשוואה $Tv - w = 0$, נוכל לנסות למצוא $v \in V$ כך ש־ $\|Tv - w\|$ הוא קטן ככל האפשר. כדוגמה נוספת, אם למשוואה לעיל יש אינסוף פתרונות $v \in V$, אזי מִבֵּין כל הפתרונות הללו נוכל לנסות למצוא פתרון שעבורו $\|v\|$ קטן ככל האפשר.

הפסידו־הופכי יספק לנו את הכלי לפתרון המשוואה לעיל במידת האפשר, גם כאשר T אינה הפיכה. נזדקק למסקנה הבאה כדי להגדיר את הפסידו־הופכי.

בשתי ההוכחות הבאות, נשתמש, בלא לציין זאת בכל פעם, בתוצאה שאם V מממד סופי ו־ $T \in \mathcal{L}(V,W)$, אזי $\operatorname{null} T$, $(\operatorname{null} T)^\perp$, ו־ $\operatorname{range} T$ כולם מממד סופי אף הם.

6.67 צמצום תחום העתקה ליניארית לקבלת העתקה חד־חד־ערכית ועל

נניח ש־ V מממד סופי ו־ $T \in \mathcal{L}(V,W)$. אזי $T|_{(\operatorname{null} T)^\perp}$ העתקה חד־חד־ערכית של $(\operatorname{null} T)^\perp$ על $\operatorname{range} T$.

הוכחה נניח ש־ $v \in (\operatorname{null} T)^\perp$ ו־ $T|_{(\operatorname{null} T)^\perp} v = 0$. לכן $Tv = 0$ ולפיכך $v \in (\operatorname{null} T)^\perp \cap (\operatorname{null} T)$, מה שגורר ש־ $v = 0$ [לפי 6.48(ד)]. לכן $\operatorname{null}(T|_{(\operatorname{null} T)^\perp}) = \{0\}$, מה שגורר ש־ $T|_{(\operatorname{null} T)^\perp}$ חד־חד־ערכית, כפי שרצינו.

ברור כי $\operatorname{range}(T|_{(\operatorname{null} T)^\perp}) \subseteq \operatorname{range} T$. להוכחת ההכלה בכיוון השני, נניח ש־ $w \in \operatorname{range} T$. לכן קיים $v \in V$ כך ש־ $w = Tv$. קיימים $u \in \operatorname{null} T$ ו־ $x \in (\operatorname{null} T)^\perp$ כך ש־ $v = u + x$ (לפי 6.49). כעת

$$T|_{(\operatorname{null} T)^\perp} x = Tx = Tv - Tu = w - 0 = w,$$

מה שמראה ש־ $w \in \operatorname{range} T|_{(\operatorname{null} T)^\perp}$. לכן $\operatorname{range} T \subseteq \operatorname{range} T|_{(\operatorname{null} T)^\perp}$, ובכך הושלמה ההוכחה ש־ $\operatorname{range} T = \operatorname{range} T|_{(\operatorname{null} T)^\perp}$. ■

כעת נוכל להגדיר את הפְּסֵידוֹ־הופכי $T^\dagger$ (מבוטא "T dagger") של העתקה ליניארית T. בהגדרה הבאה (ומנקודה זו ואילך), חשבו על $T|_{(\operatorname{null} T)^\perp}$ כהעתקה ליניארית הפיכה מ־ $(\operatorname{null} T)^\perp$ על $\operatorname{range} T$, שהרי הצדקנו זאת במסקנה הקודמת.

להפקת הסימון לפְּסֵידוֹ־הופכי $T^\dagger$ ב־ TEX, כתבו `T^\dagger`.

6.68 definition: *pseudoinverse*, $T^\dagger$

Suppose that V is finite-dimensional and $T \in \mathcal{L}(V, W)$. The *pseudoinverse* $T^\dagger \in \mathcal{L}(W, V)$ of T is the linear map from W to V defined by

$$T^\dagger w = (T|_{(\operatorname{null} T)^\perp})^{-1} P_{\operatorname{range} T}\, w$$

for each $w \in W$.

Recall that $P_{\operatorname{range} T}\, w = 0$ if $w \in (\operatorname{range} T)^\perp$ and $P_{\operatorname{range} T}\, w = w$ if $w \in \operatorname{range} T$. Thus if $w \in (\operatorname{range} T)^\perp$, then $T^\dagger w = 0$, and if $w \in \operatorname{range} T$, then $T^\dagger w$ is the unique element of $(\operatorname{null} T)^\perp$ such that $T(T^\dagger w) = w$.

The pseudoinverse behaves much like an inverse, as we will see.

6.69 *algebraic properties of the pseudoinverse*

Suppose V is finite-dimensional and $T \in \mathcal{L}(V, W)$.

(a) If T is invertible, then $T^\dagger = T^{-1}$.

(b) $TT^\dagger = P_{\operatorname{range} T}$ = the orthogonal projection of W onto $\operatorname{range} T$.

(c) $T^\dagger T = P_{(\operatorname{null} T)^\perp}$ = the orthogonal projection of V onto $(\operatorname{null} T)^\perp$.

Proof

(a) Suppose T is invertible. Then $(\operatorname{null} T)^\perp = V$ and $\operatorname{range} T = W$. Thus $T|_{(\operatorname{null} T)^\perp} = T$ and $P_{\operatorname{range} T}$ is the identity operator on W. Hence $T^\dagger = T^{-1}$.

(b) Suppose $w \in \operatorname{range} T$. Thus

$$TT^\dagger w = T(T|_{(\operatorname{null} T)^\perp})^{-1} w = w = P_{\operatorname{range} T}\, w.$$

If $w \in (\operatorname{range} T)^\perp$, then $T^\dagger w = 0$. Hence $TT^\dagger w = 0 = P_{\operatorname{range} T}\, w$. Thus $TT^\dagger$ and $P_{\operatorname{range} T}$ agree on $\operatorname{range} T$ and on $(\operatorname{range} T)^\perp$. Hence these two linear maps are equal (by 6.49).

(c) Suppose $v \in (\operatorname{null} T)^\perp$. Because $Tv \in \operatorname{range} T$, the definition of $T^\dagger$ shows that

$$T^\dagger(Tv) = (T|_{(\operatorname{null} T)^\perp})^{-1}(Tv) = v = P_{(\operatorname{null} T)^\perp} v.$$

If $v \in \operatorname{null} T$, then $T^\dagger Tv = 0 = P_{(\operatorname{null} T)^\perp} v$. Thus $T^\dagger T$ and $P_{(\operatorname{null} T)^\perp}$ agree on $(\operatorname{null} T)^\perp$ and on $\operatorname{null} T$. Hence these two linear maps are equal (by 6.49). ∎

The pseudoinverse is also called the Moore–Penrose inverse.

Suppose that $T \in \mathcal{L}(V, W)$. If T is surjective, then $TT^\dagger$ is the identity operator on W, as follows from (b) in the result above. If T is injective, then $T^\dagger T$ is the identity operator on V, as follows from (c) in the result above. For additional algebraic properties of the pseudoinverse, see Exercises 19–23.

6.68 הגדרה: פְּסֵיזוֹ־הופכי, $T^\dagger$

נניח ש־ V מממד סופי ו־ $T \in \mathcal{L}(V,W)$. הפְּסֵיזוֹ־הופכי $T^\dagger \in \mathcal{L}(W,V)$ של T הוא ההעתקה הליניארית מ־ W אל V המוגדרת על ידי

$$T^\dagger w = (T|_{(\operatorname{null} T)^\perp})^{-1} P_{\operatorname{range} T}\, w$$

לכל $w \in W$.

נזכיר כי $P_{\operatorname{range} T}\, w = 0$ אם $w \in (\operatorname{range} T)^\perp$ ו־ $P_{\operatorname{range} T}\, w = w$ אם $w \in \operatorname{range} T$.
לפיכך אם $w \in (\operatorname{range} T)^\perp$, אזי $T^\dagger w = 0$, ואם $w \in \operatorname{range} T$, אזי $T^\dagger w$ הוא האיבר היחידי של $(\operatorname{null} T)^\perp$ כך ש־ $T(T^\dagger w) = w$.
התנהגות הפְּסֵיזוֹ־הופכי דומה לזו של ההופכית הסטנדרטית, כפי שנראה בהמשך.

6.69 תכונות אלגבריות של הפְּסֵיזוֹ־הופכי

נניח ש־ V מממד סופי ו־ $T \in \mathcal{L}(V,W)$.

א. אם T הפיכה, אזי $T^\dagger = T^{-1}$.

ב. $TT^\dagger = P_{\operatorname{range} T}$, שהוא ההיטל האורתוגונלי של W על $\operatorname{range} T$.

ג. $T^\dagger T = P_{(\operatorname{null} T)^\perp}$ שהוא ההיטל האורתוגונלי של V על $(\operatorname{null} T)^\perp$.

הוכחה

א. נניח ש־ T הפיכה. אזי $(\operatorname{null} T)^\perp = V$ ו־ $\operatorname{range} T = W$. לפיכך $T|_{(\operatorname{null} T)^\perp} = T$ ו־ $P_{\operatorname{range} T}$ הוא אופרטור הזהות מעל W. לכן $T^\dagger = T^{-1}$.

ב. נניח ש־ $w \in \operatorname{range} T$. לפיכך

$$TT^\dagger w = T(T|_{(\operatorname{null} T)^\perp})^{-1} w = w = P_{\operatorname{range} T}\, w.$$

אם $w \in (\operatorname{range} T)^\perp$, אזי $T^\dagger w = 0$. לכן $TT^\dagger w = 0 = P_{\operatorname{range} T}\, w$. לפיכך $TT^\dagger$ ו־ $P_{\operatorname{range} T}$ מתואמים בנוגע ל־ $\operatorname{range} T$ ו־ $(\operatorname{range} T)^\perp$. לכן שתי העתקות ליניאריות אלו שוות (לפי 6.49).

ג. נניח ש־ $v \in (\operatorname{null} T)^\perp$. מאחר ש־ $Tv \in \operatorname{range} T$, ההגדרה של $T^\dagger$ מראה ש־

$$T^\dagger(Tv) = (T|_{(\operatorname{null} T)^\perp})^{-1}(Tv) = v = P_{(\operatorname{null} T)^\perp} v.$$

אם $v \in \operatorname{null} T$, אזי $T^\dagger Tv = 0 = P_{(\operatorname{null} T)^\perp} v$. לפיכך $T^\dagger T$ ו־ $P_{(\operatorname{null} T)^\perp}$ מתואמים לגבי $(\operatorname{null} T)^\perp$ ו־ $\operatorname{null} T$. לכן שתי העתקות ליניאריות אלו שוות (לפי 6.49). ■

הפְּסֵיזוֹ־הופכי נקרא גם הופכי מור–פנרוז.

נניח ש־ $T \in \mathcal{L}(V,W)$. אם T על, אזי $TT^\dagger$ הוא אופרטור הזהות מעל W, כפי שנובע מסעיף (ב) במסקנה לעיל. אם T חד־חד־ערכית, אזי $T^\dagger T$ הוא אופרטור הזהות מעל V, כפי שנובע מסעיף (ג) במסקנה לעיל. לתכונות אלגבריות נוספות של הפסידו־הופכי, ראו תרגילים 19–23.

For $T \in \mathcal{L}(V, W)$ and $w \in W$, we now return to the problem of finding $v \in V$ that solves the equation

$$Tv = w.$$

As we noted earlier, if T is invertible, then $v = T^{-1}w$ is the unique solution, but if T is not invertible, then T^{-1} is not defined. However, the pseudoinverse $T^\dagger$ is defined. Taking $v = T^\dagger w$ makes Tv as close to w as possible, as shown by (a) of the next result. Thus the pseudoinverse provides what is called a *best fit* to the equation above.

Among all vectors $v \in V$ that make Tv as close as possible to w, the vector $T^\dagger w$ has the smallest norm, as shown by combining (b) in the next result with the condition for equality in (a).

6.70 *pseudoinverse provides best approximate solution or best solution*

Suppose V is finite-dimensional, $T \in \mathcal{L}(V, W)$, and $w \in W$.

(a) If $v \in V$, then

$$\|T(T^\dagger w) - w\| \le \|Tv - w\|,$$

with equality if and only if $v \in T^\dagger w + \operatorname{null} T$.

(b) If $v \in T^\dagger w + \operatorname{null} T$, then

$$\|T^\dagger w\| \le \|v\|,$$

with equality if and only if $v = T^\dagger w$.

Proof

(a) Suppose $v \in V$. Then

$$Tv - w = (Tv - TT^\dagger w) + (TT^\dagger w - w).$$

The first term in parentheses above is in range T. Because the operator $TT^\dagger$ is the orthogonal projection of W onto range T [by 6.69(b)], the second term in parentheses above is in $(\operatorname{range} T)^\perp$ [see 6.57(f)].

Thus the Pythagorean theorem implies the desired inequality that the norm of the second term in parentheses above is less than or equal to $\|Tv - w\|$, with equality if and only if the first term in parentheses above equals 0. Hence we have equality if and only if $v - T^\dagger w \in \operatorname{null} T$, which is equivalent to the statement that $v \in T^\dagger w + \operatorname{null} T$, completing the proof of (a).

(b) Suppose $v \in T^\dagger w + \operatorname{null} T$. Hence $v - T^\dagger w \in \operatorname{null} T$. Now

$$v = (v - T^\dagger w) + T^\dagger w.$$

The definition of $T^\dagger$ implies that $T^\dagger w \in (\operatorname{null} T)^\perp$. Thus the Pythagorean theorem implies that $\|T^\dagger w\| \le \|v\|$, with equality if and only if $v = T^\dagger w$. ∎

A formula for $T^\dagger$ will be given in the next chapter (see 7.78).

עבור $T \in \mathcal{L}(V,W)$ ו־ $w \in W$, נשוב כעת אל בעיית מציאת $v \in V$ הפותר את המשוואה

$$Tv = w.$$

כפי שציינו קודם לכן, אם T הפיכה, אזי $v = T^{-1}w$ הוא הפתרון היחיד, אך אם T אינה הפיכה, אזי T^{-1} אינה מוגדרת. אולם, הפסידו־הופכי $T^\dagger$ מוגדר. בחירת $v = T^\dagger w$ מקרבת את Tv ככל הניתן ל־ w, כפי שנראה בסעיף (א) של המסקנה הבאה. לפיכך הפְּסֵידוֹ־הופכי מספק את מה שקרוי ההתאמה הטובה ביותר עבור המשוואה לעיל.

מבין כל הווקטורים $v \in V$ הגורמים ל־ Tv להתקרב ככל האפשר ל־ w, לווקטור $T^\dagger w$ יש את הנורמה הקטנה ביותר, כפי שניתן לראות כשמצרפים את סעיף (ב) במסקנה הבאה עם תנאי השוויון בסעיף (א).

6.70 פְּסֵידוֹ־הופכי מספק את פתרון הקירוב הטוב ביותר או את הפתרון הטוב ביותר

נניח ש־ V מממד סופי, $T \in \mathcal{L}(V,W)$, ו־ $w \in W$.

א. אם $v \in V$, אזי

$$\|T(T^\dagger w) - w\| \le \|Tv - w\|,$$

עם שוויון אם־ורק־אם $v \in T^\dagger w + \operatorname{null} T$.

ב. אם $v \in T^\dagger w + \operatorname{null} T$, אזי

$$\|T^\dagger w\| \le \|v\|,$$

עם שוויון אם־ורק־אם $v = T^\dagger w$.

הוכחה

א. נניח ש־ $v \in V$. אזי

$$Tv - w = (Tv - TT^\dagger w) + (TT^\dagger w - w).$$

הביטוי הראשון בסוגריים לעיל הוא ב־ $\operatorname{range} T$. מאחר שהאופרטור $TT^\dagger$ הוא ההיטל האורתוגונלי של W על $\operatorname{range} T$ [לפי 6.69(ב)], הביטוי השני בסוגריים לעיל הוא ב־ $(\operatorname{range} T)^\perp$ [ראו 6.57(ו)].

לפיכך משפט פיתגורס גורר את האי־שוויון המבוקש שהנורמה של הביטוי השני בסוגריים לעיל היא קטנה או שווה ל־ $\|Tv - w\|$, עם שוויון אם־ורק־אם הביטוי הראשון בסוגריים לעיל שווה 0. לכן יש לנו שוויון אם־ורק־אם $v - T^\dagger w \in \operatorname{null} T$, שזו טענה השקולה לכך ש־ $v \in T^\dagger w + \operatorname{null} T$, ובכך הושלמה ההוכחה של סעיף (א).

ב. נניח ש־ $v \in T^\dagger w + \operatorname{null} T$. לכן $v - T^\dagger w \in \operatorname{null} T$. כעת

$$v = (v - T^\dagger w) + T^\dagger w.$$

ההגדרה של $T^\dagger$ גוררת ש־ $T^\dagger w \in (\operatorname{null} T)^\perp$. לפיכך משפט פיתגורס גורר ש־ $\|T^\dagger w\| \le \|v\|$, עם שוויון אם־ורק־אם $v = T^\dagger w$. ■

נוסחה עבור $T^\dagger$ תינתן בפרק הבא (ראו 7.78).

6.71 example: *pseudoinverse of a linear map from* $\mathbf{F}^4$ *to* $\mathbf{F}^3$

Suppose $T \in \mathcal{L}(\mathbf{F}^4, \mathbf{F}^3)$ is defined by

$$T(a, b, c, d) = (a + b + c, 2c + d, 0).$$

This linear map is neither injective nor surjective, but we can compute its pseudoinverse. To do this, first note that $\operatorname{range} T = \{(x, y, 0) : x, y \in \mathbf{F}\}$. Thus

$$P_{\operatorname{range} T}(x, y, z) = (x, y, 0)$$

for each $(x, y, z) \in \mathbf{F}^3$. Also,

$$\operatorname{null} T = \{(a, b, c, d) \in \mathbf{F}^4 : a + b + c = 0 \text{ and } 2c + d = 0\}.$$

The list $(-1, 1, 0, 0), (-1, 0, 1, -2)$ of two vectors in $\operatorname{null} T$ spans $\operatorname{null} T$ because if $(a, b, c, d) \in \operatorname{null} T$ then

$$(a, b, c, d) = b(-1, 1, 0, 0) + c(-1, 0, 1, -2).$$

Because the list $(-1, 1, 0, 0), (-1, 0, 1, -2)$ is linearly independent, this list is a basis of $\operatorname{null} T$.

Now suppose $(x, y, z) \in \mathbf{F}^3$. Then

$$T^\dagger(x, y, z) = (T|_{(\operatorname{null} T)^\perp})^{-1} P_{\operatorname{range} T}(x, y, z) = (T|_{(\operatorname{null} T)^\perp})^{-1}(x, y, 0). \tag{6.72}$$

The right side of the equation above is the vector $(a, b, c, d) \in \mathbf{F}^4$ such that $T(a, b, c, d) = (x, y, 0)$ and $(a, b, c, d) \in (\operatorname{null} T)^\perp$. In other words, a, b, c, d must satisfy the following equations:

$$\begin{aligned} a + b + c &= x \\ 2c + d &= y \\ -a + b &= 0 \\ -a + c - 2d &= 0, \end{aligned}$$

where the first two equations are equivalent to the equation $T(a, b, c, d) = (x, y, 0)$ and the last two equations come from the condition for (a, b, c, d) to be orthogonal to each of the basis vectors $(-1, 1, 0, 0), (-1, 0, 1, -2)$ in this basis of $\operatorname{null} T$. Thinking of x and y as constants and a, b, c, d as unknowns, we can solve the system above of four equations in four unknowns, getting

$$a = \tfrac{1}{11}(5x - 2y),\ b = \tfrac{1}{11}(5x - 2y),\ c = \tfrac{1}{11}(x + 4y),\ d = \tfrac{1}{11}(-2x + 3y).$$

Hence 6.72 tells us that

$$T^\dagger(x, y, z) = \tfrac{1}{11}(5x - 2y, 5x - 2y, x + 4y, -2x + 3y).$$

The formula above for $T^\dagger$ shows that $TT^\dagger(x, y, z) = (x, y, 0)$ for all $(x, y, z) \in \mathbf{F}^3$, which illustrates the equation $TT^\dagger = P_{\operatorname{range} T}$ from 6.69(b).

6.71 דוגמה: פְּסֵידוֹ־הופכי של העתקה ליניארית מ־ $\mathbf{F}^4$ אל $\mathbf{F}^3$

נניח ש־ $T \in \mathcal{L}(\mathbf{F}^4, \mathbf{F}^3)$ מוגדרת על ידי

$$T(a,b,c,d) = (a+b+c, 2c+d, 0).$$

העתקה ליניארית זו אינה חד־חד־ערכית ואינה על, אולם נוכל לחשב את הפְּסֵידוֹ־הופכי שלה. לצורך כך, תחילה נשים לב כי $\operatorname{range} T = \{(x,y,0) : x,y \in \mathbf{F}\}$. לפיכך

$$P_{\operatorname{range} T}(x,y,z) = (x,y,0)$$

לכל $(x,y,z) \in \mathbf{F}^3$. בנוסף,

$$\operatorname{null} T = \{(a,b,c,d) \in \mathbf{F}^4 : a+b+c = 0 \quad \text{וגם} \quad 2c+d=0\}.$$

הרשימה $(-1,1,0,0), (-1,0,1,-2)$ של שני וקטורים ב־ $\operatorname{null} T$ פורסת את $\operatorname{null} T$ מאחר שאם $(a,b,c,d) \in \operatorname{null} T$ אזי

$$(a,b,c,d) = b(-1,1,0,0) + c(-1,0,1,-2).$$

מאחר שהרשימה $(-1,1,0,0), (-1,0,1,-2)$ היא בלתי תלויה ליניארית, רשימה זו בסיס של $\operatorname{null} T$.

כעת נניח ש־ $(x,y,z) \in \mathbf{F}^3$. אזי

$$T^\dagger(x,y,z) = (T|_{(\operatorname{null} T)^\perp})^{-1} P_{\operatorname{range} T}(x,y,z) = (T|_{(\operatorname{null} T)^\perp})^{-1}(x,y,0). \tag{6.72}$$

האגף הימני של המשוואה לעיל הוא הווקטור $(a,b,c,d) \in \mathbf{F}^4$ כך ש־

$$T(a,b,c,d) = (x,y,0) \text{ ו־ } (a,b,c,d) \in (\operatorname{null} T)^\perp.$$

במילים אחרות, a, b, c, d חייבים לקיים את המשוואות הבאות:

$$\begin{aligned} a+b+c &= x \\ 2c+d &= y \\ -a+b &= 0 \\ -a+c-2d &= 0, \end{aligned}$$

כאשר שתי המשוואות הראשונות שקולות למשוואה $T(a,b,c,d) = (x,y,0)$ ושתי המשוואות האחרונות נובעות מהתנאי ש־ (a,b,c,d) יהיה אורתוגונלי לכל וקטור מווקטורי הבסיס $(-1,1,0,0), (-1,0,1,-2)$ בבסיס זה של $\operatorname{null} T$. במחשבה על x ו־ y כקבועים ועל a, b, c, d כנעלמים, נוכל לפתור את המערכת לעיל של ארבע משוואות בארבעה נעלמים ולקבל

$$a = \tfrac{1}{11}(5x-2y),\ b = \tfrac{1}{11}(5x-2y),\ c = \tfrac{1}{11}(x+4y),\ d = \tfrac{1}{11}(-2x+3y).$$

לכן 6.72 מורה לנו כי

$$T^\dagger(x,y,z) = \tfrac{1}{11}(5x-2y, 5x-2y, x+4y, -2x+3y).$$

הנוסחה לעיל עבור $T^\dagger$ מראה לנו ש־ $TT^\dagger(x,y,z) = (x,y,0)$ לכל $(x,y,z) \in \mathbf{F}^3$, מה שממחיש את המשוואה $TT^\dagger = P_{\operatorname{range} T}$ מ־ 6.69(ב).

Exercises 6C

1 Suppose $v_1, \dots, v_m \in V$. Prove that

$$\{v_1, \dots, v_m\}^\perp = (\operatorname{span}(v_1, \dots, v_m))^\perp.$$

2 Suppose U is a subspace of V with basis $u_1, \dots, u_m$ and

$$u_1, \dots, u_m, v_1, \dots, v_n$$

is a basis of V. Prove that if the Gram–Schmidt procedure is applied to the basis of V above, producing a list $e_1, \dots, e_m, f_1, \dots, f_n$, then $e_1, \dots, e_m$ is an orthonormal basis of U and $f_1, \dots, f_n$ is an orthonormal basis of $U^\perp$.

3 Suppose U is the subspace of $\mathbf{R}^4$ defined by

$$U = \operatorname{span}((1, 2, 3, -4), (-5, 4, 3, 2)).$$

Find an orthonormal basis of U and an orthonormal basis of $U^\perp$.

4 Suppose $e_1, \dots, e_n$ is a list of vectors in V with $\|e_k\| = 1$ for each $k = 1, \dots, n$ and

$$\|v\|^2 = |\langle v, e_1\rangle|^2 + \cdots + |\langle v, e_n\rangle|^2$$

for all $v \in V$. Prove that $e_1, \dots, e_n$ is an orthonormal basis of V.

This exercise provides a converse to 6.30(b).

5 Suppose that V is finite-dimensional and U is a subspace of V. Show that $P_{U^\perp} = I - P_U$, where I is the identity operator on V.

6 Suppose V is finite-dimensional and $T \in \mathcal{L}(V, W)$. Show that

$$T = TP_{(\operatorname{null} T)^\perp} = P_{\operatorname{range} T}T.$$

7 Suppose that X and Y are finite-dimensional subspaces of V. Prove that $P_X P_Y = 0$ if and only if $\langle x, y\rangle = 0$ for all $x \in X$ and all $y \in Y$.

8 Suppose U is a finite-dimensional subspace of V and $v \in V$. Define a linear functional $\varphi\colon U \to \mathbf{F}$ by

$$\varphi(u) = \langle u, v\rangle$$

for all $u \in U$. By the Riesz representation theorem (6.42) as applied to the inner product space U, there exists a unique vector $w \in U$ such that

$$\varphi(u) = \langle u, w\rangle$$

for all $u \in U$. Show that $w = P_U v$.

9 Suppose V is finite-dimensional. Suppose $P \in \mathcal{L}(V)$ is such that $P^2 = P$ and every vector in $\operatorname{null} P$ is orthogonal to every vector in $\operatorname{range} P$. Prove that there exists a subspace U of V such that $P = P_U$.

תרגילים ו.ג

1 נניח ש־ $v_1, \dots, v_m \in V$. הוכיחו ש־
$$\{v_1, \dots, v_m\}^\perp = (\operatorname{span}(v_1, \dots, v_m))^\perp.$$

2 נניח ש־ U תת־מרחב של V עם בסיס $u_1, \dots, u_m$ ו־
$$u_1, \dots, u_m, v_1, \dots, v_n$$
בסיס של V. הוכיחו שאם ניישם את תהליך גרם־שמידט על הבסיס של V לעיל וכתוצאה נקבל את הרשימה $e_1, \dots, e_m, f_1, \dots, f_n$, אזי $e_1, \dots, e_m$ בסיס אורתונורמלי של U ו־ $f_1, \dots, f_n$ בסיס אורתונורמלי של $U^\perp$.

3 נניח ש־ U התת־מרחב של $\mathbf{R}^4$ המוגדר על ידי
$$U = \operatorname{span}((1, 2, 3, -4), (-5, 4, 3, 2)).$$
מצאו בסיס אורתונורמלי של U ובסיס אורתונורמלי של $U^\perp$.

4 נניח ש־ $e_1, \dots, e_n$ רשימת וקטורים ב־ V עם $\|e_k\| = 1$ לכל $k = 1, \dots, n$ ו־
$$\|v\|^2 = |\langle v, e_1\rangle|^2 + \cdots + |\langle v, e_n\rangle|^2$$
לכל $v \in V$. הוכיחו ש־ $e_1, \dots, e_n$ בסיס אורתונורמלי של V.

תרגיל זה מספק את הכיוון ההפוך של 6.30(ב).

5 נניח ש־ V מממד סופי ו־ U תת־מרחב של V. הראו ש־ $P_{U^\perp} = I - P_U$, כאשר I הוא אופרטור הזהות מעל V.

6 נניח ש־ V מממד סופי ו־ $T \in \mathcal{L}(V, W)$. הראו ש־
$$T = TP_{(\operatorname{null} T)^\perp} = P_{\operatorname{range} T}T.$$

7 נניח ש־ X ו־ Y תתי־מרחב של V מממד סופי. הוכיחו ש־ $P_X P_Y = 0$ אם־ורק־אם $\langle x, y\rangle = 0$ לכל $x \in X$ ולכל $y \in Y$.

8 נניח ש־ U תת־מרחב של V מממד סופי ו־ $v \in V$. הגדירו פונקציונל ליניארי $\varphi : U \to \mathbf{F}$ על ידי
$$\varphi(u) = \langle u, v\rangle$$
לכל $u \in U$. לפי משפט ההצגה של ריס (6.42) כאשר הוא מיושם על מרחב המכפלה הפנימית U, קיים וקטור יחיד $w \in U$ כך ש־
$$\varphi(u) = \langle u, w\rangle$$
לכל $u \in U$. הראו ש־ $w = P_U v$.

9 נניח ש־ V מממד סופי. נניח ש־ $P \in \mathcal{L}(V)$ כך ש־ $P^2 = P$ וכל וקטור ב־ $\operatorname{null} P$ אורתוגונלי לכל וקטור ב־ $\operatorname{range} P$. הוכיחו שקיים תת־מרחב U של V כך ש־ $P = P_U$.

10 Suppose V is finite-dimensional and $P \in \mathcal{L}(V)$ is such that $P^2 = P$ and

$$\|Pv\| \le \|v\|$$

for every $v \in V$. Prove that there exists a subspace U of V such that $P = P_U$.

11 Suppose $T \in \mathcal{L}(V)$ and U is a finite-dimensional subspace of V. Prove that

$$U \text{ is invariant under } T \iff P_U T P_U = T P_U.$$

12 Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and U is a subspace of V. Prove that

$$U \text{ and } U^\perp \text{ are both invariant under } T \iff P_U T = T P_U.$$

13 Suppose $\mathbf{F} = \mathbf{R}$ and V is finite-dimensional. For each $v \in V$, let φ_v denote the linear functional on V defined by

$$\varphi_v(u) = \langle u, v \rangle$$

for all $u \in V$.

(a) Show that $v \mapsto \varphi_v$ is an injective linear map from V to V'.

(b) Use (a) and a dimension-counting argument to show that $v \mapsto \varphi_v$ is an isomorphism from V onto V'.

The purpose of this exercise is to give an alternative proof of the Riesz representation theorem (6.42 and 6.58) when $\mathbf{F} = \mathbf{R}$. Thus you should not use the Riesz representation theorem as a tool in your solution.

14 Suppose that $e_1, \dots, e_n$ is an orthonormal basis of V. Explain why the dual basis (see 3.112) of $e_1, \dots, e_n$ is $e_1, \dots, e_n$ under the identification of V' with V provided by the Riesz representation theorem (6.58).

15 In $\mathbf{R}^4$, let

$$U = \operatorname{span}((1,1,0,0), (1,1,1,2)).$$

Find $u \in U$ such that $\|u - (1,2,3,4)\|$ is as small as possible.

16 Suppose $C[-1,1]$ is the vector space of continuous real-valued functions on the interval $[-1,1]$ with inner product given by

$$\langle f, g \rangle = \int_{-1}^{1} fg$$

for all $f, g \in C[-1,1]$. Let U be the subspace of $C[-1,1]$ defined by

$$U = \{f \in C[-1,1] : f(0) = 0\}.$$

(a) Show that $U^\perp = \{0\}$.

(b) Show that 6.49 and 6.52 do not hold without the finite-dimensional hypothesis.

10 נניח ש־ V מממד סופי ו־ $P \in \mathcal{L}(V)$ כך ש־ $P^2 = P$ וגם
$$\|Pv\| \le \|v\|$$
לכל $v \in V$. הוכיחו שקיים תת־מרחב U של V כך ש־ $P = P_U$.

11 נניח ש־ $T \in \mathcal{L}(V)$ ו־ U תת־מרחב של V מממד סופי. הוכיחו ש־ U שמור תחת T אם־ורק־אם $P_U T P_U = T P_U$.

12 נניח ש־ V מממד סופי, $T \in \mathcal{L}(V)$, ו־ U תת־מרחב של V. הוכיחו ש־ U ו־ $U^\perp$ שניהם שמורים תחת T אם־ורק־אם $P_U T = T P_U$.

13 נניח ש־ $\mathbf{F} = \mathbf{R}$ ו־ V מממד סופי. לכל $v \in V$, יהי φ_v הפונקציונל הליניארי מעל V המוגדר על ידי
$$\varphi_v(u) = \langle u, v \rangle$$
לכל $u \in V$.

א. הראו ש־ $v \mapsto \varphi_v$ העתקה ליניארית חד־חד־ערכית מ־ V אל V'.

ב. השתמשו בסעיף (א) ובטיעון של ספירת ממדים כדי להראות ש־ $v \mapsto \varphi_v$ הוא איזומורפיזם מ־ V על V'.

מטרת תרגיל זה היא לספק הוכחה חלופית למשפט ההצגה של ריס (6.42 ו־ 6.58) כאשר $\mathbf{F} = \mathbf{R}$. לפיכך אין להשתמש במשפט ההצגה של ריס ככלי בפתרון שלכם.

14 נניח ש־ $e_1, \dots, e_n$ בסיס אורתונורמלי של V. הסבירו מדוע הבסיס הדואלי (ראו 3.112) של $e_1, \dots, e_n$ הוא $e_1, \dots, e_n$ תחת הזיהוי של V' עם V הנתון על ידי משפט ההצגה של ריס (6.58).

15 ב־ $\mathbf{R}^4$, יהי
$$U = \operatorname{span}((1,1,0,0),(1,1,1,2)).$$
מצאו $u \in U$ כך ש־ $\|u - (1,2,3,4)\|$ קטן ככל האפשר.

16 נניח ש־ $C[-1,1]$ הוא המרחב הווקטורי של פונקציות ממשיות רציפות מעל הקטע $[-1,1]$ עם מכפלה פנימית הנתונה על ידי
$$\langle f, g \rangle = \int_{-1}^{1} fg$$
לכל $f, g \in C[-1,1]$. יהי U התת־מרחב של $C[-1,1]$ המוגדר על ידי
$$U = \{f \in C[-1,1] : f(0) = 0\}.$$

א. הראו ש־ $U^\perp = \{0\}$.

ב. הראו ש־ 6.49 ו־ 6.52 אינם מתקיימים ללא הנחת הממד הסופי.

17 Find $p \in \mathcal{P}_3(\mathbf{R})$ such that $p(0) = 0$, $p'(0) = 0$, and $\int_0^1 |2 + 3x - p(x)|^2 \, dx$ is as small as possible.

18 Find $p \in \mathcal{P}_5(\mathbf{R})$ that makes $\int_{-\pi}^{\pi} |\sin x - p(x)|^2 \, dx$ as small as possible.

The polynomial 6.65 is an excellent approximation to the answer to this exercise, but here you are asked to find the exact solution, which involves powers of π. A computer that can perform symbolic integration should help.

19 Suppose V is finite-dimensional and $P \in \mathcal{L}(V)$ is an orthogonal projection of V onto some subspace of V. Prove that $P^\dagger = P$.

20 Suppose V is finite-dimensional and $T \in \mathcal{L}(V, W)$. Show that

$$\text{null } T^\dagger = (\text{range } T)^\perp \quad \text{and} \quad \text{range } T^\dagger = (\text{null } T)^\perp.$$

21 Suppose $T \in \mathcal{L}(\mathbf{F}^3, \mathbf{F}^2)$ is defined by

$$T(a, b, c) = (a + b + c, 2b + 3c).$$

(a) For $(x, y) \in \mathbf{F}^2$, find a formula for $T^\dagger(x, y)$.

(b) Verify that the equation $TT^\dagger = P_{\text{range } T}$ from 6.69(b) holds with the formula for $T^\dagger$ obtained in (a).

(c) Verify that the equation $T^\dagger T = P_{(\text{null } T)^\perp}$ from 6.69(c) holds with the formula for $T^\dagger$ obtained in (a).

22 Suppose V is finite-dimensional and $T \in \mathcal{L}(V, W)$. Prove that

$$TT^\dagger T = T \quad \text{and} \quad T^\dagger T T^\dagger = T^\dagger.$$

Both formulas above clearly hold if T is invertible because in that case we can replace $T^\dagger$ with T^{-1}.

23 Suppose V and W are finite-dimensional and $T \in \mathcal{L}(V, W)$. Prove that

$$(T^\dagger)^\dagger = T.$$

The equation above is analogous to the equation $(T^{-1})^{-1} = T$ that holds if T is invertible.

17 מצאו $p \in \mathcal{P}_3(\mathbf{R})$ כך ש־ $p(0) = 0,\ p'(0) = 0$, ו־ $\int_0^1 |2 + 3x - p(x)|^2\, dx$ קטן ככל האפשר.

18 מצאו $p \in \mathcal{P}_5(\mathbf{R})$ הגורם ל־ $\int_{-\pi}^{\pi} |\sin x - p(x)|^2\, dx$ להיות קטן ככל האפשר.

הפולינום 6.65 הוא קירוב מצוין לתשובה לתרגיל זה, אבל כאן אתם נדרשים למצוא פתרון מדויק, מה שדורש טיפול בחזקות של π. תוכנת מחשב המסוגלת לבצע חישוב סימבולי של אינטגרציה אמורה לסייע.

19 נניח ש־ V מממד סופי ו־ $P \in \mathcal{L}(V)$ הוא ההיטל האורתוגונלי של V על תת־מרחב כלשהו של V. הוכיחו ש־ $P = P^\dagger$.

20 נניח ש־ V מממד סופי ו־ $T \in \mathcal{L}(V, W)$. הראו ש־

$$\operatorname{range} T^\dagger = (\operatorname{null} T)^\perp \quad \text{וגם} \quad \operatorname{null} T^\dagger = (\operatorname{range} T)^\perp.$$

21 נניח ש־ $T \in \mathcal{L}(\mathbf{F}^3, \mathbf{F}^2)$ המוגדרת על ידי

$$T(a, b, c) = (a + b + c, 2b + 3c).$$

א. עבור $(x, y) \in \mathbf{F}^2$, מצאו נוסחה עבור $T^\dagger(x, y)$.

ב. ודאו שהמשוואה $TT^\dagger = P_{\operatorname{range} T}$ מ־ 6.69(ב) מתקיימת עם הנוסחה עבור $T^\dagger$ שקיבלנו בסעיף (א).

ג. ודאו שהמשוואה $T^\dagger T = P_{(\operatorname{null} T)^\perp}$ מ־ 6.69(ג) מתקיימת עם הנוסחה עבור $T^\dagger$ שקיבלנו בסעיף (א).

22 נניח ש־ V מממד סופי ו־ $T \in \mathcal{L}(V, W)$. הוכיחו ש־

$$T^\dagger T T^\dagger = T^\dagger \quad \text{וגם} \quad T T^\dagger T = T.$$

שתי הנוסחאות לעיל מתקיימות בבירור אם T הפיכה שכן במקרה כזה אנו יכולים להחליף את $T^\dagger$ ב־ T^{-1}.

23 נניח ש־ V ו־ W מממד סופי ו־ $T \in \mathcal{L}(V, W)$. הוכיחו ש־

$$(T^\dagger)^\dagger = T.$$

המשוואה לעיל היא אנלוגית למשוואה $(T^{-1})^{-1} = T$ המתקיימת אם T הפיכה.

Chapter 7

Operators on Inner Product Spaces

The deepest results related to inner product spaces deal with the subject to which we now turn—linear maps and operators on inner product spaces. As we will see, good theorems can be proved by exploiting properties of the adjoint.

The hugely important spectral theorem will provide a complete description of self-adjoint operators on real inner product spaces and of normal operators on complex inner product spaces. We will then use the spectral theorem to help understand positive operators and unitary operators, which will lead to unitary matrices and matrix factorizations. The spectral theorem will also lead to the popular singular value decomposition, which will lead to the polar decomposition.

The most important results in the rest of this book are valid only in finite dimensions. Thus from now on we assume that V and W are finite-dimensional.

standing assumptions for this chapter

- **F** denotes **R** or **C**.
- V and W are nonzero finite-dimensional inner product spaces over **F**.

Petar Milošević CC BY-SA

Market square in Lviv, a city that has had several names and has been in several countries because of changing international borders. From 1772 until 1918, the city was in Austria and was called Lemberg. Between World War I and World War II, the city was in Poland and was called Lwów. During this time, mathematicians in Lwów, particularly Stefan Banach (1892–1945) and his colleagues, developed the basic results of modern functional analysis, using tools of analysis to study infinite-dimensional vector spaces. Since the end of World War II, Lviv has been in Ukraine, which was part of the Soviet Union until Ukraine became an independent country in 1991.

פרק 7

אופרטורים מעל מרחבי מכפלה פנימית

המסקנות העמוקות ביותר הקשורות למרחבי מכפלה פנימית עוסקות בנושא שאליו אנו ניגשים כעת—העתקות ליניאריות ואופרטורים מעל מרחבי מכפלה פנימית. כפי שנראה, משפטים טובים ניתנים להוכחה באמצעות ניצול תכונותיו של הצמוד.

משפט הפירוק הספקטרלי, בעל החשיבות העצומה, יספק לנו תיאור מלא של אופרטורים הצמודים לעצמם מעל מרחבי מכפלה פנימית ממשיים ושל אופרטורים נורמליים מעל מרחבי מכפלה פנימית מרוכבים. לאחר מכן נשתמש במשפט הפירוק הספקטרלי לסיוע בהבנת אופרטורים חיוביים ואופרטורים אוניטריים, שיובילו אותנו אל מטריצות אוניטריות ופירוק לגורמים של מטריצות. משפט הפירוק הספקטרלי יוביל אותנו בנוסף אל הפירוק לערכים סינגולריים הפופולרי, שיוביל אותנו אל הפירוק הפולרי.

המסקנות החשובות ביותר מנקודה זו ואילך בספר תקפות רק בממדים סופיים. לפיכך מכאן ואילך אנו מניחים כי V ו־ W הם מממד סופי.

הנחות המוצא לפרק זה

- $\mathbf{F}$ מסמן $\mathbf{R}$ או $\mathbf{C}$.
- V ו־ W מרחבי מכפלה פנימית מממד סופי מעל $\mathbf{F}$ שאינם מרחב האפס.

Petar Milošević CC BY-SA

כיכר השוק בלביב, עיר שהיו לה מספר שמות ושכנה בתחומן של מספר ארצות כתוצאה משינויי גבולות בינלאומיים. משנת 1772 ועד לשנת 1918, היה שמה למברג והייתה לעיר השנייה בחשיבותה במחוז גליציה של האימפריה האוסטרית. בין מלחמת העולם הראשונה למלחמת העולם השנייה, הייתה העיר נתונה לשליטה פולנית ונקראה לבוב. במהלך תקופה זו, מתמטיקאים בלבוב, בפרט סטפן בַּנַך (1892–1945) וחבריו המתמטיקאים, פיתחו את המסקנות הבסיסיות של אנליזה פונקציונלית מודרנית, תוך שימוש בכלי אנליזה לחקר מרחבים וקטוריים מממד אינסופי. בתום מלחמת העולם השנייה סופחה לביב לאוקראינה, שהייתה חלק מברית המועצות עד שהפכה אוקראינה למדינה עצמאית בשנת 1991.

אלגברה ליניארית בדרך הישר, מהדורה רביעית, מחבר: שלדון אקסלר
Linear Algebra Done Right, fourth edition, by Sheldon Axler

7A Self-Adjoint and Normal Operators

Adjoints

7.1 definition: *adjoint*, T^*

Suppose $T \in \mathcal{L}(V, W)$. The *adjoint* of T is the function $T^* : W \to V$ such that

$$\langle Tv, w\rangle = \langle v, T^*w\rangle$$

for every $v \in V$ and every $w \in W$.

*The word **adjoint** has another meaning in linear algebra. In case you encounter the second meaning elsewhere, be warned that the two meanings for adjoint are unrelated to each other.*

To see why the definition above makes sense, suppose $T \in \mathcal{L}(V, W)$. Fix $w \in W$. Consider the linear functional

$$v \mapsto \langle Tv, w\rangle$$

on V that maps $v \in V$ to $\langle Tv, w\rangle$; this linear functional depends on T and w. By the Riesz representation theorem (6.42), there exists a unique vector in V such that this linear functional is given by taking the inner product with it. We call this unique vector T^*w. In other words, T^*w is the unique vector in V such that

$$\langle Tv, w\rangle = \langle v, T^*w\rangle$$

for every $v \in V$.

In the equation above, the inner product on the left takes place in W and the inner product on the right takes place in V. However, we use the same notation $\langle \cdot, \cdot \rangle$ for both inner products.

7.2 example: *adjoint of a linear map from* $\mathbf{R}^3$ *to* $\mathbf{R}^2$

Define $T : \mathbf{R}^3 \to \mathbf{R}^2$ by

$$T(x_1, x_2, x_3) = (x_2 + 3x_3, 2x_1).$$

To compute T^*, suppose $(x_1, x_2, x_3) \in \mathbf{R}^3$ and $(y_1, y_2) \in \mathbf{R}^2$. Then

$$\begin{aligned}\langle T(x_1, x_2, x_3), (y_1, y_2)\rangle &= \langle (x_2 + 3x_3, 2x_1), (y_1, y_2)\rangle \\ &= x_2y_1 + 3x_3y_1 + 2x_1y_2 \\ &= \langle (x_1, x_2, x_3), (2y_2, y_1, 3y_1)\rangle.\end{aligned}$$

The equation above and the definition of the adjoint imply that

$$T^*(y_1, y_2) = (2y_2, y_1, 3y_1).$$

ז אופרטורים צמודים לעצמם ואופרטורים נורמליים

צמודים

7.1 הגדרה: צמוד, T^*

נניח ש־ $T \in \mathcal{L}(V, W)$. הצמודה של T היא הפונקציה $T^* : W \to V$ כך ש־

$$\langle Tv, w\rangle = \langle v, T^*w\rangle$$

לכל $v \in V$ ולכל $w \in W$.

כדי לראות מדוע ההגדרה לעיל הגיונית, נניח ש־ $T \in \mathcal{L}(V, W)$. נְקַבֵּעַ $w \in W$. נתבונן בפונקציונל הליניארי

$$v \mapsto \langle Tv, w\rangle$$

מעל V המעתיק את $v \in V$ אל $\langle Tv, w\rangle$; פונקציונל ליניארי זה תלוי ב־ T וב־ w. לפי משפט ההצגה של ריס (6.42), קיים וקטור יחיד ב־ V כך שפונקציונל ליניארי זה נתון על ידי המכפלה הפנימית איתו. לווקטור יחיד זה אנו קוראים T^*w. במילים אחרות, T^*w הוא הווקטור היחיד ב־ V כך ש־

$$\langle Tv, w\rangle = \langle v, T^*w\rangle$$

לכל $v \in V$.

בספרות בשפה האנגלית, הצמוד (של העתקות, אופרטורים ומטריצות) נקרא **adjoint**. למונח זה יש משמעות נוספת באלגברה ליניארית. אם נתקלתם כבר במונח זה במשמעותו השנייה, שימו לב כי אין כל קשר בין שתי משמעויות אלו.

במשוואה לעיל, המכפלה הפנימית באגף השמאלי מתרחשת ב־ W והמכפלה הפנימית באגף הימני מתרחשת ב־ V. למרות זאת, אנו משתמשים באותו הסימון $\langle \cdot, \cdot \rangle$ עבור שתי המכפלות הפנימיות.

7.2 דוגמה: צמודה של העתקה ליניארית מ־ $\mathbf{R}^3$ אל $\mathbf{R}^2$

נגדיר $T : \mathbf{R}^3 \to \mathbf{R}^2$ על ידי

$$T(x_1, x_2, x_3) = (x_2 + 3x_3, 2x_1).$$

לחישוב T^*, נניח ש־ $(x_1, x_2, x_3) \in \mathbf{R}^3$ ו־ $(y_1, y_2) \in \mathbf{R}^2$. אזי

$$\begin{aligned}\langle T(x_1, x_2, x_3), (y_1, y_2)\rangle &= \langle (x_2 + 3x_3, 2x_1), (y_1, y_2)\rangle \\ &= x_2y_1 + 3x_3y_1 + 2x_1y_2 \\ &= \langle (x_1, x_2, x_3), (2y_2, y_1, 3y_1)\rangle.\end{aligned}$$

המשוואה לעיל והגדרת הצמוד גוררות ש־

$$T^*(y_1, y_2) = (2y_2, y_1, 3y_1).$$

7.3 example: *adjoint of a linear map with range of dimension at most* 1

Fix $u \in V$ and $x \in W$. Define $T \in \mathcal{L}(V, W)$ by

$$Tv = \langle v, u \rangle x$$

for each $v \in V$. To compute T^*, suppose $v \in V$ and $w \in W$. Then

$$\begin{aligned} \langle Tv, w \rangle &= \big\langle \langle v, u \rangle x, w \big\rangle \\ &= \langle v, u \rangle \langle x, w \rangle \\ &= \big\langle v, \langle w, x \rangle u \big\rangle. \end{aligned}$$

Thus

$$T^*w = \langle w, x \rangle u.$$

In the two examples above, T^* turned out to be not just a function from W to V but a linear map from W to V. This behavior is true in general, as shown by the next result.

The two examples above and the proof below use a common technique for computing T^*: start with a formula for $\langle Tv, w \rangle$ then manipulate it to get just v in the first slot; the entry in the second slot will then be T^*w.

7.4 *adjoint of a linear map is a linear map*

If $T \in \mathcal{L}(V, W)$, then $T^* \in \mathcal{L}(W, V)$.

Proof Suppose $T \in \mathcal{L}(V, W)$. If $v \in V$ and $w_1, w_2 \in W$, then

$$\begin{aligned} \langle Tv, w_1 + w_2 \rangle &= \langle Tv, w_1 \rangle + \langle Tv, w_2 \rangle \\ &= \langle v, T^*w_1 \rangle + \langle v, T^*w_2 \rangle \\ &= \langle v, T^*w_1 + T^*w_2 \rangle. \end{aligned}$$

The equation above shows that

$$T^*(w_1 + w_2) = T^*w_1 + T^*w_2.$$

If $v \in V$, $\lambda \in \mathbf{F}$, and $w \in W$, then

$$\begin{aligned} \langle Tv, \lambda w \rangle &= \overline{\lambda} \langle Tv, w \rangle \\ &= \overline{\lambda} \langle v, T^*w \rangle \\ &= \langle v, \lambda T^*w \rangle. \end{aligned}$$

The equation above shows that

$$T^*(\lambda w) = \lambda T^*w.$$

Thus T^* is a linear map, as desired. ∎

7.3 דוגמה: צמודה של העתקה ליניארית עם טווח חד־ממדי לכל היותר

נְקַבֵּעַ $u \in V$ ו־ $x \in W$. נגדיר $T \in \mathcal{L}(V, W)$ על ידי

$$Tv = \langle v, u \rangle x$$

לכל $v \in V$. לחישוב T^*, נניח ש־ $v \in V$ ו־ $w \in W$. אזי

$$\begin{aligned}\langle Tv, w \rangle &= \langle \langle v, u \rangle x, w \rangle \\ &= \langle v, u \rangle \langle x, w \rangle \\ &= \langle v, \langle w, x \rangle u \rangle.\end{aligned}$$

לפיכך

$$T^* w = \langle w, x \rangle u.$$

בשתי הדוגמאות לעיל, התברר ש־ T^* אינה רק פונקציה מ־ W אל V אלא העתקה ליניארית מ־ W אל V. הדבר נכון באופן כללי, כפי שמראה לנו המסקנה הבאה.

לשתי הדוגמאות לעיל ולהוכחה להלן טכניקה משותפת לחישוב T^*: מתחילים עם נוסחה עבור $\langle Tv, w \rangle$ ומבצעים פעולות אלגבריות כדי לבודד את v ברכיב הראשון; האיבר שמתקבל ברכיב השני יהיה $T^* w$.

7.4 צמודה של העתקה ליניארית היא העתקה ליניארית

אם $T \in \mathcal{L}(V, W)$, אזי $T^* \in \mathcal{L}(W, V)$.

הוכחה נניח ש־ $T \in \mathcal{L}(V, W)$. אם $v \in V$ ו־ $w_1, w_2 \in W$, אזי

$$\begin{aligned}\langle Tv, w_1 + w_2 \rangle &= \langle Tv, w_1 \rangle + \langle Tv, w_2 \rangle \\ &= \langle v, T^* w_1 \rangle + \langle v, T^* w_2 \rangle \\ &= \langle v, T^* w_1 + T^* w_2 \rangle.\end{aligned}$$

המשוואה לעיל מראה ש־

$$T^*(w_1 + w_2) = T^* w_1 + T^* w_2.$$

אם $\lambda \in \mathbf{F}$, $v \in V$, ו־ $w \in W$, אזי

$$\begin{aligned}\langle Tv, \lambda w \rangle &= \overline{\lambda} \langle Tv, w \rangle \\ &= \overline{\lambda} \langle v, T^* w \rangle \\ &= \langle v, \lambda T^* w \rangle.\end{aligned}$$

המשוואה לעיל מראה ש־

$$T^*(\lambda w) = \lambda T^* w.$$

לפיכך T^* העתקה ליניארית, כפי שרצינו. ■

7.5 *properties of the adjoint*

Suppose $T \in \mathcal{L}(V, W)$. Then

(a) $(S + T)^* = S^* + T^*$ for all $S \in \mathcal{L}(V, W)$;

(b) $(\lambda T)^* = \overline{\lambda} T^*$ for all $\lambda \in \mathbf{F}$;

(c) $(T^*)^* = T$;

(d) $(ST)^* = T^* S^*$ for all $S \in \mathcal{L}(W, U)$ (here U is a finite-dimensional inner product space over $\mathbf{F}$);

(e) $I^* = I$, where I is the identity operator on V;

(f) if T is invertible, then T^* is invertible and $(T^*)^{-1} = (T^{-1})^*$.

Proof Suppose $v \in V$ and $w \in W$.

(a) If $S \in \mathcal{L}(V, W)$, then

$$\begin{aligned}\langle (S + T)v, w\rangle &= \langle Sv, w\rangle + \langle Tv, w\rangle \\ &= \langle v, S^*w\rangle + \langle v, T^*w\rangle \\ &= \langle v, S^*w + T^*w\rangle.\end{aligned}$$

Thus $(S + T)^* w = S^*w + T^*w$, as desired.

(b) If $\lambda \in \mathbf{F}$, then

$$\langle (\lambda T)v, w\rangle = \lambda\langle Tv, w\rangle = \lambda\langle v, T^*w\rangle = \langle v, \overline{\lambda} T^*w\rangle.$$

Thus $(\lambda T)^* w = \overline{\lambda} T^*w$, as desired.

(c) We have

$$\langle T^*w, v\rangle = \overline{\langle v, T^*w\rangle} = \overline{\langle Tv, w\rangle} = \langle w, Tv\rangle.$$

Thus $(T^*)^* v = Tv$, as desired.

(d) Suppose $S \in \mathcal{L}(W, U)$ and $u \in U$. Then

$$\langle (ST)v, u\rangle = \langle S(Tv), u\rangle = \langle Tv, S^*u\rangle = \langle v, T^*(S^*u)\rangle.$$

Thus $(ST)^* u = T^*(S^*u)$, as desired.

(e) Suppose $u \in V$. Then

$$\langle Iu, v\rangle = \langle u, v\rangle.$$

Thus $I^* v = v$, as desired.

(f) Suppose T is invertible. Take adjoints of both sides of the equation $T^{-1}T = I$, then use (d) and (e) to show that $T^*(T^{-1})^* = I$. Similarly, the equation $TT^{-1} = I$ implies $(T^{-1})^* T^* = I$. Thus $(T^{-1})^*$ is the inverse of T^*, as desired. ∎

If $\mathbf{F} = \mathbf{R}$, then the map $T \mapsto T^*$ is a linear map from $\mathcal{L}(V, W)$ to $\mathcal{L}(W, V)$, as follows from (a) and (b) of the result above. However, if $\mathbf{F} = \mathbf{C}$, then this map is not linear because of the complex conjugate that appears in (b).

7.5 תכונות של הצמוד

נניח ש־ $T \in \mathcal{L}(V,W)$. אזי

א. $(S+T)^* = S^* + T^*$ לכל $S \in \mathcal{L}(V,W)$

ב. $(\lambda T)^* = \overline{\lambda} T^*$ לכל $\lambda \in \mathbf{F}$

ג. $(T^*)^* = T$

ד. $(ST)^* = T^* S^*$ לכל $S \in \mathcal{L}(W,U)$ (כאן U הוא מרחב מכפלה פנימית מממד סופי מעל $\mathbf{F}$)

ה. $I^* = I$, כאשר I הוא אופרטור הזהות מעל V

ו. אם T הפיכה, אזי T^* הפיכה ו־ $(T^*)^{-1} = (T^{-1})^*$.

הוכחה נניח ש־ $v \in V$ ו־ $w \in W$.

א. אם $S \in \mathcal{L}(V,W)$, אזי

$$\begin{aligned}\langle (S+T)v, w\rangle &= \langle Sv, w\rangle + \langle Tv, w\rangle \\ &= \langle v, S^*w\rangle + \langle v, T^*w\rangle \\ &= \langle v, S^*w + T^*w\rangle.\end{aligned}$$

לפיכך $(S+T)^*w = S^*w + T^*w$, כפי שרצינו.

ב. אם $\lambda \in \mathbf{F}$, אזי

$$\langle (\lambda T)v, w\rangle = \lambda \langle Tv, w\rangle = \lambda \langle v, T^*w\rangle = \langle v, \overline{\lambda} T^*w\rangle.$$

לפיכך $(\lambda T)^*w = \overline{\lambda} T^*w$, כפי שרצינו.

ג. מתקיים

$$\langle T^*w, v\rangle = \overline{\langle v, T^*w\rangle} = \overline{\langle Tv, w\rangle} = \langle w, Tv\rangle.$$

לפיכך $(T^*)^*v = Tv$, כפי שרצינו.

ד. נניח ש־ $S \in \mathcal{L}(W,U)$ ו־ $u \in U$. אזי

$$\langle (ST)v, u\rangle = \langle S(Tv), u\rangle = \langle Tv, S^*u\rangle = \langle v, T^*(S^*u)\rangle.$$

לפיכך $(ST)^*u = T^*(S^*u)$, כפי שרצינו.

ה. נניח ש־ $u \in V$. אזי

$$\langle Iu, v\rangle = \langle u, v\rangle.$$

לפיכך $I^*v = v$, כפי שרצינו.

ו. נניח ש־ T הפיכה. ניקח את הצמודים של שני אגפי המשוואה $T^{-1}T = I$ ואז נשתמש ב־ (ד) ו־ (ה) כדי להראות כי $T^*(T^{-1})^* = I$. באופן דומה, המשוואה $TT^{-1} = I$ גוררת ש־ $(T^{-1})^*T^* = I$. לפיכך $(T^{-1})^*$ ההופכית של T^*, כפי שרצינו. ■

אם $\mathbf{F} = \mathbf{R}$, אזי ההעתקה $T \mapsto T^*$ היא העתקה ליניארית מ־ $\mathcal{L}(V,W)$ אל $\mathcal{L}(W,V)$, כפי שנובע מסעיפים (א) ו־ (ב) של המסקנה לעיל. אולם, אם $\mathbf{F} = \mathbf{C}$, אזי העתקה זו אינה ליניארית לאור הצמוד המרוכב המופיע בסעיף (ב).

The next result shows the relationship between the null space and the range of a linear map and its adjoint.

7.6 *null space and range of T^**

Suppose $T \in \mathcal{L}(V, W)$. Then

(a) $\operatorname{null} T^* = (\operatorname{range} T)^\perp$;

(b) $\operatorname{range} T^* = (\operatorname{null} T)^\perp$;

(c) $\operatorname{null} T = (\operatorname{range} T^*)^\perp$;

(d) $\operatorname{range} T = (\operatorname{null} T^*)^\perp$.

Proof We begin by proving (a). Let $w \in W$. Then

$$\begin{aligned} w \in \operatorname{null} T^* &\iff T^* w = 0 \\ &\iff \langle v, T^* w \rangle = 0 \text{ for all } v \in V \\ &\iff \langle Tv, w \rangle = 0 \text{ for all } v \in V \\ &\iff w \in (\operatorname{range} T)^\perp. \end{aligned}$$

Thus $\operatorname{null} T^* = (\operatorname{range} T)^\perp$, proving (a).

If we take the orthogonal complement of both sides of (a), we get (d), where we have used 6.52. Replacing T with T^* in (a) gives (c), where we have used 7.5(c). Finally, replacing T with T^* in (d) gives (b). ∎

As we will soon see, the next definition is intimately connected to the matrix of the adjoint of a linear map.

7.7 definition: *conjugate transpose, A^**

The *conjugate transpose* of an m-by-n matrix A is the n-by-m matrix A^* obtained by interchanging the rows and columns and then taking the complex conjugate of each entry. In other words, if $j \in \{1, \dots, n\}$ and $k \in \{1, \dots, m\}$, then

$$(A^*)_{j,k} = \overline{A_{k,j}}.$$

7.8 example: *conjugate transpose of a 2-by-3 matrix*

The conjugate transpose of the 2-by-3 matrix $\begin{pmatrix} 2 & 3+4i & 7 \\ 6 & 5 & 8i \end{pmatrix}$ is the 3-by-2 matrix

$$\begin{pmatrix} 2 & 6 \\ 3-4i & 5 \\ 7 & -8i \end{pmatrix}.$$

If a matrix A has only real entries, then $A^* = A^t$, where A^t denotes the transpose of A (the matrix obtained by interchanging the rows and the columns).

המסקנה הבאה מראה את היחס בין מרחב האפס והטווח של העתקה ליניארית לבין אלו של צמודתה.

7.6 מרחב אפס וטווח של T^*

נניח ש־ $T \in \mathcal{L}(V, W)$. אזי

א. $\operatorname{null} T^* = (\operatorname{range} T)^\perp$

ב. $\operatorname{range} T^* = (\operatorname{null} T)^\perp$

ג. $\operatorname{null} T = (\operatorname{range} T^*)^\perp$

ד. $\operatorname{range} T = (\operatorname{null} T^*)^\perp$.

הוכחה נפתח בהוכחת סעיף (א). יהי $w \in W$. אזי

$$\begin{aligned} T^*w = 0 &\iff w \in \operatorname{null} T^* \\ &\iff \langle v, T^*w\rangle = 0 \text{ לכל } v \in V \\ &\iff \langle Tv, w\rangle = 0 \text{ לכל } v \in V \\ &\iff w \in (\operatorname{range} T)^\perp. \end{aligned}$$

לפיכך $\operatorname{null} T^* = (\operatorname{range} T)^\perp$, ובכך הוכחנו את סעיף (א).

אם ניקח את המשלים האורתוגונלי של שני האגפים של סעיף (א), אנו מקבלים את (ד), כאשר השתמשנו ב־ 6.52. בהחלפת T ב־ T^* בסעיף (א) נקבל את (ג), כאשר השתמשנו ב־ 7.5(ג). לסיום, החלפת T ב־ T^* בסעיף (ד) תתן לנו את סעיף (ב). ■

כפי שנראה בקרוב, ההגדרה הבאה קשורה באופן הדוק למטריצה של הצמודה של העתקה ליניארית.

7.7 הגדרה: צמודה־משוחלפת, A^*

הצמודה־משוחלפת של A מטריצה $m \times n$ היא A^* מטריצה $n \times m$ המתקבלת על ידי החלפה בין השורות והעמודות והצבת הצמוד המרוכב של כל איבר במטריצה. במילים אחרות, אם $j \in \{1, \dots, n\}$ ו־ $k \in \{1, \dots, m\}$, אזי

$$(A^*)_{j,k} = \overline{A_{k,j}}.$$

7.8 דוגמה: צמודה־משוחלפת של מטריצה 2×3

הצמודה־משוחלפת של המטריצה 2×3

$$\begin{pmatrix} 2 & 3+4i & 7 \\ 6 & 5 & 8i \end{pmatrix}$$

היא המטריצה 3×2

$$\begin{pmatrix} 2 & 6 \\ 3-4i & 5 \\ 7 & -8i \end{pmatrix}.$$

אם למטריצה A יש רק איברים ממשיים, אזי $A^* = A^t$, כאשר A^t מסמנת את המשוחלפת של A (המטריצה המתקבלת על ידי החלפה בין השורות והעמודות).

The next result shows how to compute the matrix of T^* from the matrix of T. **Caution:** With respect to nonorthonormal bases, the matrix of T^* does not necessarily equal the conjugate transpose of the matrix of T.

The adjoint of a linear map does not depend on a choice of basis. Thus we frequently emphasize adjoints of linear maps instead of transposes or conjugate transposes of matrices.

7.9 *matrix of T^* equals conjugate transpose of matrix of T*

Let $T \in \mathcal{L}(V, W)$. Suppose $e_1, \dots, e_n$ is an orthonormal basis of V and $f_1, \dots, f_m$ is an orthonormal basis of W. Then $\mathcal{M}(T^*, (f_1, \dots, f_m), (e_1, \dots, e_n))$ is the conjugate transpose of $\mathcal{M}(T, (e_1, \dots, e_n), (f_1, \dots, f_m))$. In other words,

$$\mathcal{M}(T^*) = (\mathcal{M}(T))^*.$$

Proof In this proof, we will write $\mathcal{M}(T)$ and $\mathcal{M}(T^*)$ instead of the longer expressions $\mathcal{M}(T, (e_1, \dots, e_n), (f_1, \dots, f_m))$ and $\mathcal{M}(T^*, (f_1, \dots, f_m), (e_1, \dots, e_n))$.

Recall that we obtain the kth column of $\mathcal{M}(T)$ by writing Te_k as a linear combination of the f_j's; the scalars used in this linear combination then become the kth column of $\mathcal{M}(T)$. Because $f_1, \dots, f_m$ is an orthonormal basis of W, we know how to write Te_k as a linear combination of the f_j's [see 6.30(a)]:

$$Te_k = \langle Te_k, f_1 \rangle f_1 + \cdots + \langle Te_k, f_m \rangle f_m.$$

Thus

$$\text{the entry in row } j, \text{ column } k, \text{ of } \mathcal{M}(T) \text{ is } \langle Te_k, f_j \rangle.$$

In the statement above, replace T with T^* and interchange $e_1, \dots, e_n$ and $f_1, \dots, f_m$. This shows that the entry in row j, column k, of $\mathcal{M}(T^*)$ is $\langle T^* f_k, e_j \rangle$, which equals $\langle f_k, Te_j \rangle$, which equals $\overline{\langle Te_j, f_k \rangle}$, which equals the complex conjugate of the entry in row k, column j, of $\mathcal{M}(T)$. Thus $\mathcal{M}(T^*) = (\mathcal{M}(T))^*$. ∎

The Riesz representation theorem as stated in 6.58 provides an identification of V with its dual space V' defined in 3.110. Under this identification, the orthogonal complement $U^\perp$ of a subset $U \subseteq V$ corresponds to the annihilator U^0 of U. If U is a subspace of V, then the formulas for the dimensions of $U^\perp$ and U^0 become identical under this identification—see 3.125 and 6.51.

Because orthogonal complements and adjoints are easier to deal with than annihilators and dual maps, there is no need to work with annihilators and dual maps in the context of inner product spaces.

Suppose $T : V \to W$ is a linear map. Under the identification of V with V' and the identification of W with W', the adjoint map $T^* : W \to V$ corresponds to the dual map $T' : W' \to V'$ defined in 3.118, as Exercise 32 asks you to verify. Under this identification, the formulas for null T^* and range T^* [7.6(a) and (b)] then become identical to the formulas for null T' and range T' [3.128(a) and 3.130(b)]. Furthermore, the theorem about the matrix of T^* (7.9) is analogous to the theorem about the matrix of T' (3.132).

המסקנה הבאה מראה כיצד לחשב את המטריצה של T^* בהתבסס על המטריצה של T. **זְהִירוּת:** ביחס לבסיסים שאינם אורתונורמליים, המטריצה של T^* לא בהכרח תהיה שווה למטריצה הצמודה־משוחלפת של המטריצה של T.

ההעתקה הצמודה של העתקה ליניארית אינה תלויה בבחירה של בסיס. לפיכך לעיתים קרובות נשים דגש על ההעתקות הצמודות של העתקות ליניאריות במקום על מטריצות משוחלפות או מטריצות צמודות־משוחלפות.

7.9 מטריצה של T^* שווה לצמודה־משוחלפת של המטריצה של T

תהי $T \in \mathcal{L}(V, W)$. נניח ש־ $e_1, \dots, e_n$ בסיס אורתונורמלי של V ו־ $f_1, \dots, f_m$ בסיס אורתונורמלי של W. אזי $\mathcal{M}(T^*, (f_1, \dots, f_m), (e_1, \dots, e_n))$ היא הצמודה־משוחלפת של $\mathcal{M}(T, (e_1, \dots, e_n), (f_1, \dots, f_m))$. במילים אחרות,

$$\mathcal{M}(T^*) = (\mathcal{M}(T))^*.$$

הוכחה בהוכחה זו, נכתוב $\mathcal{M}(T)$ ו־ $\mathcal{M}(T^*)$ במקום לכתוב את הביטויים הארוכים $\mathcal{M}(T, (e_1, \dots, e_n), (f_1, \dots, f_m))$ ו־ $\mathcal{M}(T^*, (f_1, \dots, f_m), (e_1, \dots, e_n))$.

נזכיר כי העמודה ה־ k^{th} של $\mathcal{M}(T)$ מתקבלת באמצעות כתיבת Te_k כצירוף ליניארי של ה־ f_j־ים; הסקלרים שבהם השתמשנו בצירוף ליניארי זה מוצבים בעמודה ה־ k^{th} של $\mathcal{M}(T)$. מאחר ש־ $f_1, \dots, f_m$ הוא בסיס אורתונורמלי של W, אנו יודעים כיצד לכתוב את Te_k כצירוף ליניארי של ה־ f_j־ים [ראו 6.30(א)]:

$$Te_k = \langle Te_k, f_1 \rangle f_1 + \cdots + \langle Te_k, f_m \rangle f_m.$$

לפיכך

האיבר בשורה j, עמודה k, של $\mathcal{M}(T)$ הוא $\langle Te_k, f_j \rangle$.

במשפט לעיל, נחליף את T ב־ T^* ונחליף בין $e_1, \dots, e_n$ ל־ $f_1, \dots, f_m$. כך נראה כי האיבר בשורה j, עמודה k, של $\mathcal{M}(T^*)$ הוא $\langle T^* f_k, e_j \rangle$, שהוא שווה ל־ $\langle f_k, Te_j \rangle$, שהוא שווה ל־ $\overline{\langle Te_j, f_k \rangle}$, שהוא שווה לצמוד המרוכב של האיבר בשורה k, עמודה j, של $\mathcal{M}(T)$. לפיכך $\mathcal{M}(T^*) = (\mathcal{M}(T))^*$. ■

משפט ההצגה של ריס כפי שנוסח ב־ 6.58 מספק זיהוי של V עם המרחב הדואלי שלו V' כמוגדר ב־ 3.110. תחת זיהוי זה, המשלים האורתוגונלי $U^\perp$ של התת־קבוצה $U \subseteq V$ תואם למאפס U^0 של U. אם U תת־מרחב של V, אזי נוסחאות הממדים של $U^\perp$ ושל U^0 הופכות לְזֶהוֹת תחת זיהוי זה—ראו 3.125 ו־ 6.51.

מאחר שמשלימים אורתוגונליים וצמודים יותר קלים לעבודה מאשר מְמַאְפְּסִים והעתקות דואליות, בהקשר של מרחבי מכפלה פנימית אין צורך לעבוד עם מְמַאְפְּסִים והעתקות דואליות.

נניח ש־ $T : V \to W$ העתקה ליניארית. תחת הזיהוי של V עם V' והזיהוי של W עם W', ההעתקה הצמודה $T^* : W \to V$ תואמת להעתקה הדואלית $T' : W' \to V'$ המוגדרת ב־ 3.118, כפי שתרגיל 32 מבקש מכם לוודא. תחת זיהוי זה, הנוסחאות עבור $\operatorname{null} T^*$ ו־ $\operatorname{range} T^*$ [ראו 7.6(א) ו־ (ב)] הופכות לזהות לנוסחאות עבור $\operatorname{null} T'$ ו־ $\operatorname{range} T'$ [ראו 3.128(א) ו־ 3.130(ב)]. יתר על כן, המשפט לגבי המטריצה של T^* (ראו 7.9) אנלוגי למשפט לגבי המטריצה של T' (ראו 3.132).

Self-Adjoint Operators

Now we switch our attention to operators on inner product spaces. Instead of considering linear maps from V to W, we will focus on linear maps from V to V; recall that such linear maps are called operators.

7.10 definition: *self-adjoint*

An operator $T \in \mathcal{L}(V)$ is called *self-adjoint* if $T = T^*$.

If $T \in \mathcal{L}(V)$ and $e_1, \dots, e_n$ is an orthonormal basis of V, then T is self-adjoint if and only if $\mathcal{M}(T, (e_1, \dots, e_n)) = \mathcal{M}(T, (e_1, \dots, e_n))^*$, as follows from 7.9.

7.11 example: *determining whether T is self-adjoint from its matrix*

Suppose $c \in \mathbf{F}$ and T is the operator on $\mathbf{F}^2$ whose matrix (with respect to the standard basis) is

$$\mathcal{M}(T) = \begin{pmatrix} 2 & c \\ 3 & 7 \end{pmatrix}.$$

The matrix of T^* (with respect to the standard basis) is

$$\mathcal{M}(T^*) = \begin{pmatrix} 2 & 3 \\ \overline{c} & 7 \end{pmatrix}.$$

Thus $\mathcal{M}(T) = \mathcal{M}(T^*)$ if and only if $c = 3$. Hence the operator T is self-adjoint if and only if $c = 3$.

A good analogy to keep in mind is that the adjoint on $\mathcal{L}(V)$ plays a role similar to that of the complex conjugate on $\mathbf{C}$. A complex number z is real if and only if $z = \overline{z}$; thus a self-adjoint operator ($T = T^*$) is analogous to a real number.

We will see that the analogy discussed above is reflected in some important properties of self-adjoint operators, beginning with eigenvalues in the next result.

If $\mathbf{F} = \mathbf{R}$, then by definition every eigenvalue is real, so the next result is interesting only when $\mathbf{F} = \mathbf{C}$.

An operator $T \in \mathcal{L}(V)$ is self-adjoint if and only if

$$\langle Tv, w \rangle = \langle v, Tw \rangle$$

for all $v, w \in V$.

7.12 *eigenvalues of self-adjoint operators*

Every eigenvalue of a self-adjoint operator is real.

Proof Suppose T is a self-adjoint operator on V. Let λ be an eigenvalue of T, and let v be a nonzero vector in V such that $Tv = \lambda v$. Then

$$\lambda \|v\|^2 = \langle \lambda v, v \rangle = \langle Tv, v \rangle = \langle v, Tv \rangle = \langle v, \lambda v \rangle = \overline{\lambda} \|v\|^2.$$

Thus $\lambda = \overline{\lambda}$, which means that λ is real, as desired. ∎

אופרטורים צמודים לעצמם

כעת אנו מפנים את תשומת ליבנו אל אופרטורים מעל מרחבי מכפלה פנימית. במקום להתבונן בהעתקות ליניאריות מ־ V אל W, אנו נתמקד בהעתקות ליניאריות מ־ V אל V; נזכיר כי העתקות ליניאריות שכאלה נקראות אופרטורים.

7.10 הגדרה: צמוד לעצמו

אופרטור $T \in \mathcal{L}(V)$ נקרא צמוד לעצמו אם $T = T^*$.

אם $T \in \mathcal{L}(V)$ ו־ $e_1, \dots, e_n$ בסיס אורתונורמלי של V, אזי T צמוד לעצמו אם־ורק־אם $\mathcal{M}(T, (e_1, \dots, e_n)) = \mathcal{M}(T, (e_1, \dots, e_n))^*$, כפי שנובע מ־ 7.9.

7.11 דוגמה: קביעה האם T צמוד לעצמו על פי המטריצה שלו

נניח ש־ $c \in \mathbf{F}$ ו־ T האופרטור מעל $\mathbf{F}^2$ שהמטריצה שלו (ביחס לבסיס הסטנדרטי) היא

$$\mathcal{M}(T) = \begin{pmatrix} 2 & c \\ 3 & 7 \end{pmatrix}.$$

המטריצה של T^* (ביחס לבסיס הסטנדרטי) היא

$$\mathcal{M}(T^*) = \begin{pmatrix} 2 & 3 \\ \bar{c} & 7 \end{pmatrix}.$$

לפיכך $\mathcal{M}(T) = \mathcal{M}(T^*)$ אם־ורק־אם $c = 3$. לכן האופרטור T צמוד לעצמו אם־ורק־אם $c = 3$.

אנלוגיה שימושית שכדאי לזכור היא כי לפעולת הצמוד מעל $\mathcal{L}(V)$ תפקיד דומה לזה של הצמוד המרוכב מעל $\mathbf{C}$. מספר מרוכב z הוא ממשי אם־ורק־אם $z = \bar{z}$; לפיכך אופרטור צמוד לעצמו ($T = T^*$) אנלוגי למספר ממשי.

אופרטור $T \in \mathcal{L}(V)$ צמוד לעצמו אם־ורק־אם

$$\langle Tv, w \rangle = \langle v, Tw \rangle$$

לכל $v, w \in V$.

אנו נראה כי האנלוגיה שהוזכרה לעיל באה לידי ביטוי במספר תכונות חשובות של אופרטורים צמודים לעצמם, כשהראשונה שבהן קשורה לערכים עצמיים במסקנה הבאה.

אם $\mathbf{F} = \mathbf{R}$, אזי לפי הגדרה כל ערך עצמי הוא ממשי, ולכן יש עניין במסקנה הבאה רק כאשר $\mathbf{F} = \mathbf{C}$.

7.12 ערכים עצמיים של אופרטורים צמודים לעצמם

כל ערך עצמי של אופרטור צמוד לעצמו הוא ממשי.

הוכחה נניח ש־ T אופרטור צמוד לעצמו מעל V. יהי λ ערך עצמי של T, ויהי v וקטור השונה מאפס ב־ V כך ש־ $Tv = \lambda v$. אזי

$$\lambda \|v\|^2 = \langle \lambda v, v \rangle = \langle Tv, v \rangle = \langle v, Tv \rangle = \langle v, \lambda v \rangle = \bar{\lambda} \|v\|^2.$$

לפיכך $\lambda = \bar{\lambda}$, כלומר λ מספר ממשי, כפי שרצינו. ■

The next result is false for real inner product spaces. As an example, consider the operator $T \in \mathcal{L}(\mathbf{R}^2)$ that is a counterclockwise rotation of 90° around the origin; thus $T(x, y) = (-y, x)$. Notice that Tv is orthogonal to v for every $v \in \mathbf{R}^2$, even though $T \neq 0$.

7.13 *Tv is orthogonal to v for all v $\iff$ $T = 0$ (assuming $\mathbf{F} = \mathbf{C}$)*

Suppose V is a complex inner product space and $T \in \mathcal{L}(V)$. Then

$$\langle Tv, v\rangle = 0 \text{ for every } v \in V \iff T = 0.$$

Proof If $u, w \in V$, then

$$\langle Tu, w\rangle = \frac{\langle T(u+w), u+w\rangle - \langle T(u-w), u-w\rangle}{4} + \frac{\langle T(u+iw), u+iw\rangle - \langle T(u-iw), u-iw\rangle}{4} i,$$

as can be verified by computing the right side. Note that each term on the right side is of the form $\langle Tv, v\rangle$ for appropriate $v \in V$.

Now suppose $\langle Tv, v\rangle = 0$ for every $v \in V$. Then the equation above implies that $\langle Tu, w\rangle = 0$ for all $u, w \in V$, which then implies that $Tu = 0$ for every $u \in V$ (take $w = Tu$). Hence $T = 0$, as desired. ∎

The next result is false for real inner product spaces, as shown by considering any operator on a real inner product space that is not self-adjoint.

The next result provides another good example of how self-adjoint operators behave like real numbers.

7.14 *$\langle Tv, v\rangle$ is real for all v $\iff$ T is self-adjoint (assuming $\mathbf{F} = \mathbf{C}$)*

Suppose V is a complex inner product space and $T \in \mathcal{L}(V)$. Then

$$T \text{ is self-adjoint} \iff \langle Tv, v\rangle \in \mathbf{R} \text{ for every } v \in V.$$

Proof If $v \in V$, then

$$\langle T^*v, v\rangle = \overline{\langle v, T^*v\rangle} = \overline{\langle Tv, v\rangle}. \tag{7.15}$$

Now

$$\begin{aligned} T \text{ is self-adjoint} &\iff T - T^* = 0 \\ &\iff \langle (T - T^*)v, v\rangle = 0 \text{ for every } v \in V \\ &\iff \langle Tv, v\rangle - \overline{\langle Tv, v\rangle} = 0 \text{ for every } v \in V \\ &\iff \langle Tv, v\rangle \in \mathbf{R} \text{ for every } v \in V, \end{aligned}$$

where the second equivalence follows from 7.13 as applied to $T - T^*$ and the third equivalence follows from 7.15. ∎

המסקנה הבאה אינה נכונה עבור מרחבי מכפלה פנימית ממשיים. כדוגמה, נתבונן באופרטור $T \in \mathcal{L}(\mathbf{R}^2)$ שמבצע סיבוב של 90° סביב ראשית הצירים נגד כיוון השעון; לפיכך $T(x,y) = (-y,x)$. שימו לב ש־ Tv אורתוגונלי ל־ v לכל $v \in \mathbf{R}^2$, אף ש־ $T \neq 0$.

7.13 Tv אורתוגונלי ל־ v לכל v אם־ורק־אם $T = 0$ (בהנחה ש־ $\mathbf{F} = \mathbf{C}$)

נניח ש־ V מרחב מכפלה פנימית מרוכב ו־ $T \in \mathcal{L}(V)$. אזי

$$T = 0 \iff \langle Tv, v\rangle = 0 \text{ לכל } v \in V.$$

הוכחה אם $u, w \in V$, אזי

$$\langle Tu, w\rangle = \frac{\langle T(u+w), u+w\rangle - \langle T(u-w), u-w\rangle}{4} + \frac{\langle T(u+iw), u+iw\rangle - \langle T(u-iw), u-iw\rangle}{4}\, i,$$

כפי שניתן לוודא על ידי חישוב האגף הימני. שימו לב שכל ביטוי באגף הימני הוא מהצורה $\langle Tv, v\rangle$ עבור $v \in V$ מתאים.

כעת נניח ש־ $\langle Tv, v\rangle = 0$ לכל $v \in V$. אזי מהמשוואה לעיל נובע ש־ $\langle Tu, w\rangle = 0$ לכל $u, w \in V$, ומכאן ש־ $Tu = 0$ לכל $u \in V$ (ניקח $w = Tu$). לכן $T = 0$, כפי שרצינו. ■

המסקנה הבאה אינה נכונה עבור מרחבי מכפלה פנימית ממשיים, כפי שרואים כאשר מתבוננים בכל אופרטור מעל מרחב מכפלה פנימית ממשי שאינו צמוד לעצמו.

המסקנה הבאה מספקת דוגמה טובה נוספת לכך שאופרטורים צמודים לעצמם מתנהגים כמו מספרים ממשיים.

7.14 $\langle Tv, v\rangle$ ממשי לכל v אם־ורק־אם T צמוד לעצמו (בהנחה ש־ $\mathbf{F} = \mathbf{C}$)

נניח ש־ V מרחב מכפלה פנימית מרוכב ו־ $T \in \mathcal{L}(V)$. אזי

$$T \text{ צמוד לעצמו} \iff \langle Tv, v\rangle \in \mathbf{R} \text{ לכל } v \in V.$$

הוכחה אם $v \in V$, אזי

$$\langle T^*v, v\rangle = \overline{\langle v, T^*v\rangle} = \overline{\langle Tv, v\rangle}. \tag{7.15}$$

כעת

$$\begin{aligned} T \text{ צמוד לעצמו} &\iff T - T^* = 0 \\ &\iff \langle (T-T^*)v, v\rangle = 0 \text{ לכל } v \in V \\ &\iff \langle Tv, v\rangle - \overline{\langle Tv, v\rangle} = 0 \text{ לכל } v \in V \\ &\iff \langle Tv, v\rangle \in \mathbf{R} \text{ לכל } v \in V, \end{aligned}$$

כאשר השקילות השנייה נובעת מ־ 7.13 המופעל על $T - T^*$ והשקילות השלישית נובעת מ־ 7.15. ■

On a real inner product space V, a nonzero operator T might satisfy $\langle Tv, v\rangle = 0$ for all $v \in V$. However, the next result shows that this cannot happen for a self-adjoint operator.

7.16 *T self-adjoint and $\langle Tv, v\rangle = 0$ for all v $\iff$ $T = 0$*

Suppose T is a self-adjoint operator on V. Then

$$\langle Tv, v\rangle = 0 \text{ for every } v \in V \iff T = 0.$$

Proof We have already proved this (without the hypothesis that T is self-adjoint) when V is a complex inner product space (see 7.13). Thus we can assume that V is a real inner product space. If $u, w \in V$, then

$$\langle Tu, w\rangle = \frac{\langle T(u+w), u+w\rangle - \langle T(u-w), u-w\rangle}{4}, \tag{7.17}$$

as can be proved by computing the right side using the equation

$$\langle Tw, u\rangle = \langle w, Tu\rangle = \langle Tu, w\rangle,$$

where the first equality holds because T is self-adjoint and the second equality holds because we are working in a real inner product space.

Now suppose $\langle Tv, v\rangle = 0$ for every $v \in V$. Because each term on the right side of 7.17 is of the form $\langle Tv, v\rangle$ for appropriate v, this implies that $\langle Tu, w\rangle = 0$ for all $u, w \in V$. This implies that $Tu = 0$ for every $u \in V$ (take $w = Tu$). Hence $T = 0$, as desired. ∎

Normal Operators

7.18 definition: *normal*

- An operator on an inner product space is called *normal* if it commutes with its adjoint.
- In other words, $T \in \mathcal{L}(V)$ is normal if $TT^* = T^*T$.

Every self-adjoint operator is normal, because if T is self-adjoint then $T^* = T$ and hence T commutes with T^*.

7.19 example: *an operator that is normal but not self-adjoint*

Let T be the operator on $\mathbf{F}^2$ whose matrix (with respect to the standard basis) is

$$\begin{pmatrix} 2 & -3 \\ 3 & 2 \end{pmatrix}.$$

Thus $T(w, z) = (2w - 3z, 3w + 2z)$.

מעל מרחב מכפלה פנימית ממשי V, ייתכן שאופרטור T השונה מאפס יקיים $\langle Tv, v\rangle = 0$ לכל $v \in V$. בכל זאת, המסקנה הבאה מראה שהדבר אינו ייתכן עבור אופרטור הצמוד לעצמו.

7.16 T צמוד לעצמו ו־ $\langle Tv, v\rangle = 0$ לכל v אם־ורק־אם $T = 0$

נניח ש־ T אופרטור צמוד לעצמו מעל V. אזי

$$\langle Tv, v\rangle = 0 \text{ לכל } v \in V \iff T = 0.$$

הוכחה כאשר V מרחב מכפלה פנימית מרוכב (ללא ההנחה ש־ T צמוד לעצמו) כבר הוכחנו זאת (ראו 7.13). לפיכך אנו יכולים להניח ש־ V מרחב מכפלה פנימית ממשי. אם $u, w \in V$, אזי

$$\langle Tu, w\rangle = \frac{\langle T(u+w), u+w\rangle - \langle T(u-w), u-w\rangle}{4}, \tag{7.17}$$

כפי שניתן להוכיח על ידי חישוב האגף הימני תוך שימוש במשוואה

$$\langle Tw, u\rangle = \langle w, Tu\rangle = \langle Tu, w\rangle,$$

כאשר השוויון הראשון מתקיים מאחר ש־ T צמוד לעצמו והשוויון השני מתקיים מאחר שאנו עובדים במרחב מכפלה פנימית ממשי.

כעת נניח ש־ $\langle Tv, v\rangle = 0$ לכל $v \in V$. מאחר שכל ביטוי באגף הימני של 7.17 הוא מהצורה $\langle Tv, v\rangle$ עבור v מתאים, מכאן נובע ש־ $\langle Tu, w\rangle = 0$ לכל $u, w \in V$. מכאן נובע כי $Tu = 0$ לכל $u \in V$ (ניקח $w = Tu$). לכן $T = 0$, כפי שרצינו. ■

אופרטורים נורמליים

7.18 הגדרה: נורמלי

- אופרטור מעל מרחב מכפלה פנימית נקרא נורמלי אם הוא מתחלף עם האופרטור הצמוד שלו.
- במילים אחרות, $T \in \mathcal{L}(V)$ נורמלי אם $TT^* = T^*T$.

כל אופרטור צמוד לעצמו הוא נורמלי, מאחר שאם T צמוד לעצמו אזי $T^* = T$ ולכן T מתחלף עם T^*.

7.19 דוגמה: אופרטור נורמלי שאינו צמוד לעצמו

יהי T האופרטור מעל $\mathbf{F}^2$ שהמטריצה שלו (ביחס לבסיס הסטנדרטי) היא

$$\begin{pmatrix} 2 & -3 \\ 3 & 2 \end{pmatrix}.$$

לפיכך $T(w, z) = (2w - 3z, 3w + 2z)$.

This operator T is not self-adjoint because the entry in row 2, column 1 (which equals 3) does not equal the complex conjugate of the entry in row 1, column 2 (which equals -3).

The matrix of TT^* equals

$$\begin{pmatrix} 2 & -3 \\ 3 & 2 \end{pmatrix}\begin{pmatrix} 2 & 3 \\ -3 & 2 \end{pmatrix}, \text{ which equals } \begin{pmatrix} 13 & 0 \\ 0 & 13 \end{pmatrix}.$$

Similarly, the matrix of T^*T equals

$$\begin{pmatrix} 2 & 3 \\ -3 & 2 \end{pmatrix}\begin{pmatrix} 2 & -3 \\ 3 & 2 \end{pmatrix}, \text{ which equals } \begin{pmatrix} 13 & 0 \\ 0 & 13 \end{pmatrix}.$$

Because TT^* and T^*T have the same matrix, we see that $TT^* = T^*T$. Thus T is normal.

In the next section we will see why normal operators are worthy of special attention. The next result provides a useful characterization of normal operators.

7.20 *T is normal if and only if Tv and T^*v have the same norm*

Suppose $T \in \mathcal{L}(V)$. Then

$$T \text{ is normal} \iff \|Tv\| = \|T^*v\| \text{ for every } v \in V.$$

Proof We have

$$\begin{aligned} T \text{ is normal} &\iff T^*T - TT^* = 0 \\ &\iff \langle (T^*T - TT^*)v, v\rangle = 0 \text{ for every } v \in V \\ &\iff \langle T^*Tv, v\rangle = \langle TT^*v, v\rangle \text{ for every } v \in V \\ &\iff \langle Tv, Tv\rangle = \langle T^*v, T^*v\rangle \text{ for every } v \in V \\ &\iff \|Tv\|^2 = \|T^*v\|^2 \text{ for every } v \in V \\ &\iff \|Tv\| = \|T^*v\| \text{ for every } v \in V, \end{aligned}$$

where we used 7.16 to establish the second equivalence (note that the operator $T^*T - TT^*$ is self-adjoint). ∎

The next result presents several consequences of the result above. Compare (e) of the next result to Exercise 3. That exercise states that the eigenvalues of the adjoint of each operator are equal (as a set) to the complex conjugates of the eigenvalues of the operator. The exercise says nothing about eigenvectors, because an operator and its adjoint may have different eigenvectors. However, (e) of the next result implies that a normal operator and its adjoint have the same eigenvectors.

אופרטור T זה אינו צמוד לעצמו מאחר שהאיבר בשורה 2, עמודה 1 (השווה ל־ 3) אינו שווה לצמוד המרוכב של האיבר בשורה 1, עמודה 2 (השווה ל־ -3).

המטריצה של TT^* שווה ל־

$$\begin{pmatrix} 2 & -3 \\ 3 & 2 \end{pmatrix}\begin{pmatrix} 2 & 3 \\ -3 & 2 \end{pmatrix} = \begin{pmatrix} 13 & 0 \\ 0 & 13 \end{pmatrix}.$$

באופן דומה, המטריצה של T^*T שווה ל־

$$\begin{pmatrix} 2 & 3 \\ -3 & 2 \end{pmatrix}\begin{pmatrix} 2 & -3 \\ 3 & 2 \end{pmatrix} = \begin{pmatrix} 13 & 0 \\ 0 & 13 \end{pmatrix}.$$

מאחר של־ TT^* ול־ T^*T אותה מטריצה, אנו רואים ש־ $TT^* = T^*T$. לפיכך T אופרטור נורמלי.

ביחידה הבאה אנו נראה מדוע אופרטורים נורמליים ראויים לתשומת־לב מיוחדת. המסקנה הבאה מספקת אפיון שימושי לאופרטורים נורמליים.

7.20 T נורמלי אם־ורק־אם ל־ Tv ול־ T^*v יש אותה הנורמה

נניח ש־ $T \in \mathcal{L}(V)$. אזי

$$T \text{ נורמלי} \iff \|Tv\| = \|T^*v\| \text{ לכל } v \in V.$$

הוכחה מתקיים

$$\begin{aligned}
T \text{ נורמלי} &\iff T^*T - TT^* = 0 \\
&\iff \langle (T^*T - TT^*)v, v\rangle = 0 \text{ לכל } v \in V \\
&\iff \langle T^*Tv, v\rangle = \langle TT^*v, v\rangle \text{ לכל } v \in V \\
&\iff \langle Tv, Tv\rangle = \langle T^*v, T^*v\rangle \text{ לכל } v \in V \\
&\iff \|Tv\|^2 = \|T^*v\|^2 \text{ לכל } v \in V \\
&\iff \|Tv\| = \|T^*v\| \text{ לכל } v \in V,
\end{aligned}$$

כאשר השתמשנו ב־ 7.16 כדי לבסס את השקילות השנייה (שימו לב שהאופרטור $T^*T - TT^*$ צמוד לעצמו). ■

המסקנה הבאה מציגה מספר השלכות של המסקנה לעיל. השוו את סעיף (ה) של המסקנה הבאה לתרגיל 3. תרגיל זה קובע כי הערכים העצמיים של האופרטור הצמוד של כל אופרטור שווים (כקבוצה) לצמודים המרוכבים של הערכים העצמיים של האופרטור. התרגיל לא מציין דבר לגבי הווקטורים העצמיים, מאחר שלאופרטור ולאופרטור הצמוד שלו ייתכן שיהיו וקטורים עצמיים שונים. בכל זאת, מסעיף (ה) של המסקנה הבאה נובע שלאופרטור נורמלי ולאופרטור הצמוד שלו אותם וקטורים עצמיים.

7.21 *range, null space, and eigenvectors of a normal operator*

Suppose $T \in \mathcal{L}(V)$ is normal. Then

(a) $\operatorname{null} T = \operatorname{null} T^*$;

(b) $\operatorname{range} T = \operatorname{range} T^*$;

(c) $V = \operatorname{null} T \oplus \operatorname{range} T$;

(d) $T - \lambda I$ is normal for every $\lambda \in \mathbf{F}$;

(e) if $v \in V$ and $\lambda \in \mathbf{F}$, then $Tv = \lambda v$ if and only if $T^*v = \overline{\lambda} v$.

Proof

(a) Suppose $v \in V$. Then

$$v \in \operatorname{null} T \iff \|Tv\| = 0 \iff \|T^*v\| = 0 \iff v \in \operatorname{null} T^*,$$

where the middle equivalence above follows from 7.20. Thus $\operatorname{null} T = \operatorname{null} T^*$.

(b) We have

$$\operatorname{range} T = (\operatorname{null} T^*)^\perp = (\operatorname{null} T)^\perp = \operatorname{range} T^*,$$

where the first equality comes from 7.6(d), the second equality comes from (a) in this result, and the third equality comes from 7.6(b).

(c) We have

$$V = (\operatorname{null} T) \oplus (\operatorname{null} T)^\perp = \operatorname{null} T \oplus \operatorname{range} T^* = \operatorname{null} T \oplus \operatorname{range} T,$$

where the first equality comes from 6.49, the second equality comes from 7.6(b), and the third equality comes from (b) in this result.

(d) Suppose $\lambda \in \mathbf{F}$. Then

$$\begin{aligned}
(T - \lambda I)(T - \lambda I)^* &= (T - \lambda I)(T^* - \overline{\lambda} I) \\
&= TT^* - \overline{\lambda} T - \lambda T^* + |\lambda|^2 I \\
&= T^*T - \overline{\lambda} T - \lambda T^* + |\lambda|^2 I \\
&= (T^* - \overline{\lambda} I)(T - \lambda I) \\
&= (T - \lambda I)^*(T - \lambda I).
\end{aligned}$$

Thus $T - \lambda I$ commutes with its adjoint. Hence $T - \lambda I$ is normal.

(e) Suppose $v \in V$ and $\lambda \in \mathbf{F}$. Then (d) and 7.20 imply that

$$\|(T - \lambda I)v\| = \|(T - \lambda I)^* v\| = \|(T^* - \overline{\lambda} I)v\|.$$

Thus $\|(T - \lambda I)v\| = 0$ if and only if $\|(T^* - \overline{\lambda} I)v\| = 0$. Hence $Tv = \lambda v$ if and only if $T^*v = \overline{\lambda} v$.

7.21 טווח, מרחב האפס, וקטורים עצמיים של אופרטור נורמלי

נניח ש־ $T \in \mathcal{L}(V)$ אופרטור נורמלי. אזי

א. $\operatorname{null} T = \operatorname{null} T^*$

ב. $\operatorname{range} T = \operatorname{range} T^*$

ג. $V = \operatorname{null} T \oplus \operatorname{range} T$

ד. $T - \lambda I$ נורמלי לכל $\lambda \in \mathbf{F}$

ה. אם $v \in V$ ו־ $\lambda \in \mathbf{F}$, אזי $Tv = \lambda v$ אם־ורק־אם $T^*v = \overline{\lambda} v$.

הוכחה

א. נניח ש־ $v \in V$. אזי

$$v \in \operatorname{null} T \iff \|Tv\| = 0 \iff \|T^*v\| = 0 \iff v \in \operatorname{null} T^*,$$

כאשר השקילות האמצעית לעיל נובעת מ־ 7.20. לפיכך $\operatorname{null} T = \operatorname{null} T^*$.

ב. מתקיים

$$\operatorname{range} T = (\operatorname{null} T^*)^\perp = (\operatorname{null} T)^\perp = \operatorname{range} T^*,$$

כאשר השוויון הראשון נובע מ־ 7.6(ד), השוויון השני נובע מסעיף (א) של מסקנה זו, והשוויון השלישי נובע מ־ 7.6(ב).

ג. מתקיים

$$V = (\operatorname{null} T) \oplus (\operatorname{null} T)^\perp = \operatorname{null} T \oplus \operatorname{range} T^* = \operatorname{null} T \oplus \operatorname{range} T,$$

כאשר השוויון הראשון נובע מ־ 6.49, השוויון השני נובע מ־ 7.6(ב), והשוויון השלישי נובע מסעיף (ב) של מסקנה זו.

ד. נניח ש־ $\lambda \in \mathbf{F}$. אזי

$$\begin{aligned}(T - \lambda I)(T - \lambda I)^* &= (T - \lambda I)(T^* - \overline{\lambda} I)\\ &= TT^* - \overline{\lambda} T - \lambda T^* + |\lambda|^2 I\\ &= T^*T - \overline{\lambda} T - \lambda T^* + |\lambda|^2 I\\ &= (T^* - \overline{\lambda} I)(T - \lambda I)\\ &= (T - \lambda I)^*(T - \lambda I).\end{aligned}$$

לפיכך $T - \lambda I$ מתחלף עם האופרטור הצמוד שלו. לכן $T - \lambda I$ נורמלי.

ה. נניח ש־ $v \in V$ ו־ $\lambda \in \mathbf{F}$. אזי סעיף (ד) ו־ 7.20 גוררים ש־

$$\|(T - \lambda I)v\| = \|(T - \lambda I)^* v\| = \|(T^* - \overline{\lambda} I)v\|.$$

לפיכך $\|(T - \lambda I)v\| = 0$ אם־ורק־אם $\|(T^* - \overline{\lambda} I)v\| = 0$. לכן $Tv = \lambda v$ אם־ורק־אם $T^*v = \overline{\lambda} v$. ■

Because every self-adjoint operator is normal, the next result applies in particular to self-adjoint operators.

7.22 *orthogonal eigenvectors for normal operators*

Suppose $T \in \mathcal{L}(V)$ is normal. Then eigenvectors of T corresponding to distinct eigenvalues are orthogonal.

Proof Suppose α, β are distinct eigenvalues of T, with corresponding eigenvectors u, v. Thus $Tu = \alpha u$ and $Tv = \beta v$. From 7.21(e) we have $T^*v = \overline{\beta} v$. Thus

$$\begin{aligned}(\alpha - \beta)\langle u, v\rangle &= \langle \alpha u, v\rangle - \langle u, \overline{\beta} v\rangle \\ &= \langle Tu, v\rangle - \langle u, T^*v\rangle \\ &= 0.\end{aligned}$$

Because $\alpha \neq \beta$, the equation above implies that $\langle u, v\rangle = 0$. Thus u and v are orthogonal, as desired. ∎

As stated here, the next result makes sense only when $\mathbf{F} = \mathbf{C}$. However, see Exercise 12 for a version that makes sense when $\mathbf{F} = \mathbf{C}$ and when $\mathbf{F} = \mathbf{R}$.

Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$. Under the analogy between $\mathcal{L}(V)$ and $\mathbf{C}$, with the adjoint on $\mathcal{L}(V)$ playing a similar role to that of the complex conjugate on $\mathbf{C}$, the operators A and B as defined by 7.24 correspond to the real and imaginary parts of T. Thus the informal title of the result below should make sense.

7.23 *T is normal $\iff$ the real and imaginary parts of T commute*

Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$. Then T is normal if and only if there exist commuting self-adjoint operators A and B such that $T = A + iB$.

Proof First suppose T is normal. Let

$$A = \frac{T + T^*}{2} \quad \text{and} \quad B = \frac{T - T^*}{2i}. \tag{7.24}$$

Then A and B are self-adjoint and $T = A + iB$. A quick computation shows that

$$AB - BA = \frac{T^*T - TT^*}{2i}. \tag{7.25}$$

Because T is normal, the right side of the equation above equals 0. Thus the operators A and B commute, as desired.

To prove the implication in the other direction, now suppose there exist commuting self-adjoint operators A and B such that $T = A + iB$. Then $T^* = A - iB$. Adding the last two equations and then dividing by 2 produces the equation for A in 7.24. Subtracting the last two equations and then dividing by $2i$ produces the equation for B in 7.24. Now 7.24 implies 7.25. Because B and A commute, 7.25 implies that T is normal, as desired. ∎

מאחר שכל אופרטור צמוד לעצמו הוא נורמלי, המסקנה הבאה מתקיימת בפרט עבור אופרטורים צמודים לעצמם.

7.22 וקטורים עצמיים אורתוגונליים עבור אופרטורים נורמליים

נניח ש־ $T \in \mathcal{L}(V)$ נורמלי. אזי וקטורים עצמיים של T השייכים לערכים עצמיים שונים הם אורתוגונליים.

הוכחה נניח ש־ α, β ערכים עצמיים שונים של T, עם וקטורים עצמיים מתאימים u, v. לפיכך $Tu = \alpha u$ ו־ $Tv = \beta v$. מ־ 7.21(ה) נתון ש־ $T^*v = \overline{\beta} v$. לפיכך

$$\begin{aligned}(\alpha - \beta)\langle u, v\rangle &= \langle \alpha u, v\rangle - \langle u, \overline{\beta} v\rangle \\ &= \langle Tu, v\rangle - \langle u, T^*v\rangle \\ &= 0.\end{aligned}$$

מאחר ש־ $\alpha \neq \beta$, מהמשוואה לעיל נובע ש־ $\langle u, v\rangle = 0$. לפיכך u ו־ v אורתוגונליים, כפי שרצינו. ■

כפי שמצוין כאן, למסקנה הבאה יש משמעות רק כאשר $\mathbf{F} = \mathbf{C}$. בכל זאת, ראו תרגיל 12 לגרסה שבה יש למשפט משמעות כאשר $\mathbf{F} = \mathbf{C}$ וגם כאשר $\mathbf{F} = \mathbf{R}$.

נניח ש־ $\mathbf{F} = \mathbf{C}$ ו־ $T \in \mathcal{L}(V)$. בהשראת האנלוגיה בין $\mathcal{L}(V)$ ובין $\mathbf{C}$, כאשר תפקיד האופרטור הצמוד מעל $\mathcal{L}(V)$ דומה לזה של הצמוד המרוכב מעל $\mathbf{C}$, האופרטורים A ו־ B כפי שהם מוגדרים לפי 7.24 מתאימים לחלקים הממשי והמדומה של T. לפיכך יש הגיון בכותרת הבלתי רשמית של המסקנה להלן.

7.23 T נורמלי אם־ורק־אם החלקים הממשי והמדומה של T מתחלפים

נניח ש־ $\mathbf{F} = \mathbf{C}$ ו־ $T \in \mathcal{L}(V)$. אזי T נורמלי אם־ורק־אם קיימים אופרטורים מתחלפים הצמודים לעצמם A ו־ B כך ש־ $T = A + iB$.

הוכחה תחילה נניח כי T נורמלי. יהיו

$$A = \frac{T + T^*}{2} \quad \text{ו־} \quad B = \frac{T - T^*}{2i}. \tag{7.24}$$

אזי A ו־ B צמודים לעצמם ו־ $T = A + iB$. חישוב מהיר מראה ש־

$$AB - BA = \frac{T^*T - TT^*}{2i}. \tag{7.25}$$

מאחר ש־ T נורמלי, האגף הימני של המשוואה לעיל שווה 0. לפיכך האופרטורים A ו־ B מתחלפים, כפי שרצינו.

להוכחת הגרירה בכיוון השני, כעת נניח שקיימים אופרטורים מתחלפים הצמודים לעצמם A ו־ B כך ש־ $T = A + iB$. אזי $T^* = A - iB$. חיבור שתי המשוואות האחרונות ואז חלוקה ב־ 2 מפיקה את המשוואה עבור A ב־ 7.24. חיסור שתי המשוואות האחרונות ואז חלוקה ב־ $2i$ מפיקה את המשוואה עבור B ב־ 7.24. כעת 7.24 גורר את 7.25. מאחר ש־ B ו־ A מתחלפים, 7.25 גורר ש־ T נורמלי, כפי שרצינו. ■

Exercises 7A

1 Suppose n is a positive integer. Define $T \in \mathcal{L}(\mathbf{F}^n)$ by

$$T(z_1, \dots, z_n) = (0, z_1, \dots, z_{n-1}).$$

Find a formula for $T^*(z_1, \dots, z_n)$.

2 Suppose $T \in \mathcal{L}(V, W)$. Prove that

$$T = 0 \iff T^* = 0 \iff T^*T = 0 \iff TT^* = 0.$$

3 Suppose $T \in \mathcal{L}(V)$ and $\lambda \in \mathbf{F}$. Prove that

$$\lambda \text{ is an eigenvalue of } T \iff \overline{\lambda} \text{ is an eigenvalue of } T^*.$$

4 Suppose $T \in \mathcal{L}(V)$ and U is a subspace of V. Prove that

$$U \text{ is invariant under } T \iff U^\perp \text{ is invariant under } T^*.$$

5 Suppose $T \in \mathcal{L}(V, W)$. Suppose $e_1, \dots, e_n$ is an orthonormal basis of V and $f_1, \dots, f_m$ is an orthonormal basis of W. Prove that

$$\|Te_1\|^2 + \cdots + \|Te_n\|^2 = \|T^*f_1\|^2 + \cdots + \|T^*f_m\|^2.$$

The numbers $\|Te_1\|^2, \dots, \|Te_n\|^2$ in the equation above depend on the orthonormal basis $e_1, \dots, e_n$, but the right side of the equation does not depend on $e_1, \dots, e_n$. Thus the equation above shows that the sum on the left side does not depend on which orthonormal basis $e_1, \dots, e_n$ is used.

6 Suppose $T \in \mathcal{L}(V, W)$. Prove that
(a) T is injective $\iff$ T^* is surjective;
(b) T is surjective $\iff$ T^* is injective.

7 Prove that if $T \in \mathcal{L}(V, W)$, then
(a) $\dim \operatorname{null} T^* = \dim \operatorname{null} T + \dim W - \dim V$;
(b) $\dim \operatorname{range} T^* = \dim \operatorname{range} T$.

8 Suppose A is an m-by-n matrix with entries in $\mathbf{F}$. Use (b) in Exercise 7 to prove that the row rank of A equals the column rank of A.

This exercise asks for yet another alternative proof of a result that was previously proved in 3.57 and 3.133.

9 Prove that the product of two self-adjoint operators on V is self-adjoint if and only if the two operators commute.

10 Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$. Prove that T is self-adjoint if and only if

$$\langle Tv, v \rangle = \langle T^*v, v \rangle$$

for all $v \in V$.

תרגילים 7א

1 נניח ש־ n מספר שלם חיובי. נגדיר $T \in \mathcal{L}(\mathbf{F}^n)$ על ידי
$$T(z_1, \dots, z_n) = (0, z_1, \dots, z_{n-1}).$$
מצאו נוסחה עבור $T^*(z_1, \dots, z_n)$.

2 נניח ש־ $T \in \mathcal{L}(V, W)$. הוכיחו כי
$$T = 0 \iff T^* = 0 \iff T^*T = 0 \iff TT^* = 0.$$

3 נניח ש־ $T \in \mathcal{L}(V)$ ו־ $\lambda \in \mathbf{F}$. הוכיחו כי

λ הוא ערך עצמי של T אם־ורק־אם $\overline{\lambda}$ הוא ערך עצמי של T^*.

4 נניח ש־ $T \in \mathcal{L}(V)$ ו־ U תת־מרחב של V. הוכיחו כי

U שמור תחת T אם־ורק־אם $U^\perp$ שמור תחת T^*.

5 נניח ש־ $T \in \mathcal{L}(V, W)$. נניח ש־ $e_1, \dots, e_n$ בסיס אורתונורמלי של V ו־ $f_1, \dots, f_m$ בסיס אורתונורמלי של W. הוכיחו כי
$$\|Te_1\|^2 + \cdots + \|Te_n\|^2 = \|T^*f_1\|^2 + \cdots + \|T^*f_m\|^2.$$

המספרים $\|Te_1\|^2, \dots, \|Te_n\|^2$ במשוואה לעיל תלויים בבסיס אורתונורמלי $e_1, \dots, e_n$, אך האגף הימני של המשוואה אינו תלוי ב־ $e_1, \dots, e_n$. לפיכך המשוואה לעיל מראה שהסכום באגף השמאלי אינו תלוי באיזה בסיס אורתונורמלי $e_1, \dots, e_n$ השתמשנו.

6 נניח ש־ $T \in \mathcal{L}(V, W)$. הוכיחו כי

א. T חד־חד־ערכית $\iff$ T^* עַל

ב. T עַל $\iff$ T^* חד־חד־ערכית.

7 הוכיחו כי אם $T \in \mathcal{L}(V, W)$, אזי

א. $\dim \operatorname{null} T^* = \dim \operatorname{null} T + \dim W - \dim V$

ב. $\dim \operatorname{range} T^* = \dim \operatorname{range} T$.

8 נניח ש־ A מטריצה $m \times n$ עם איברים ב־ $\mathbf{F}$. השתמשו בסעיף (ב) בתרגיל 7 כדי להוכיח כי דרגת השורות של A שווה לדרגת העמודות של A.

תרגיל זה מבקש הוכחה חלופית נוספת למסקנה שכבר הוכחה ב־ 3.57 ו־ 3.133.

9 הוכיחו כי המכפלה של שני אופרטורים צמודים לעצמם מעל V היא אופרטור צמוד לעצמו אם־ורק־אם שני האופרטורים מתחלפים.

10 נניח ש־ $\mathbf{F} = \mathbf{C}$ ו־ $T \in \mathcal{L}(V)$. הוכיחו כי T צמוד לעצמו אם־ורק־אם
$$\langle Tv, v \rangle = \langle T^*v, v \rangle$$
לכל $v \in V$.

11 Define an operator $S: \mathbf{F}^2 \to \mathbf{F}^2$ by $S(w, z) = (-z, w)$.

(a) Find a formula for S^*.

(b) Show that S is normal but not self-adjoint.

(c) Find all eigenvalues of S.

If $\mathbf{F} = \mathbf{R}$, then S is the operator on $\mathbf{R}^2$ of counterclockwise rotation by $90°$.

12 An operator $B \in \mathcal{L}(V)$ is called *skew* if

$$B^* = -B.$$

Suppose that $T \in \mathcal{L}(V)$. Prove that T is normal if and only if there exist commuting operators A and B such that A is self-adjoint, B is a skew operator, and $T = A + B$.

13 Suppose $\mathbf{F} = \mathbf{R}$. Define $\mathcal{A} \in \mathcal{L}(\mathcal{L}(V))$ by $\mathcal{A}T = T^*$ for all $T \in \mathcal{L}(V)$.

(a) Find all eigenvalues of $\mathcal{A}$.

(b) Find the minimal polynomial of $\mathcal{A}$.

14 Define an inner product on $\mathcal{P}_2(\mathbf{R})$ by $\langle p, q\rangle = \int_0^1 pq$. Define an operator $T \in \mathcal{L}(\mathcal{P}_2(\mathbf{R}))$ by

$$T(ax^2 + bx + c) = bx.$$

(a) Show that with this inner product, the operator T is not self-adjoint.

(b) The matrix of T with respect to the basis $1, x, x^2$ is

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

This matrix equals its conjugate transpose, even though T is not self-adjoint. Explain why this is not a contradiction.

15 Suppose $T \in \mathcal{L}(V)$ is invertible. Prove that

(a) T is self-adjoint $\iff$ T^{-1} is self-adjoint;

(b) T is normal $\iff$ T^{-1} is normal.

16 Suppose $\mathbf{F} = \mathbf{R}$.

(a) Show that the set of self-adjoint operators on V is a subspace of $\mathcal{L}(V)$.

(b) What is the dimension of the subspace of $\mathcal{L}(V)$ in (a) [in terms of $\dim V$]?

17 Suppose $\mathbf{F} = \mathbf{C}$. Show that the set of self-adjoint operators on V is not a subspace of $\mathcal{L}(V)$.

18 Suppose $\dim V \geq 2$. Show that the set of normal operators on V is not a subspace of $\mathcal{L}(V)$.

11 הגדירו אופרטור $S: \mathbf{F}^2 \to \mathbf{F}^2$ על ידי $S(w,z) = (-z,w)$.

א. מצאו נוסחה עבור S^*.

ב. הראו כי S נורמלי אך אינו צמוד לעצמו.

ג. מצאו את כל הערכים העצמיים של S.

אם $\mathbf{F} = \mathbf{R}$, אזי S הוא האופרטור מעל $\mathbf{R}^2$ המבצע סיבוב של $90°$ נגד כיוון השעון.

12 אופרטור $B \in \mathcal{L}(V)$ נקרא *אנטי־צמוד* אם

$$B^* = -B.$$

נניח ש־ $T \in \mathcal{L}(V)$. הוכיחו כי T נורמלי אם־ורק־אם קיימים אופרטורים מתחלפים A ו־ B כך ש־ A צמוד לעצמו, B אנטי־צמוד, ו־ $T = A + B$.

13 נניח ש־ $\mathbf{F} = \mathbf{R}$. נגדיר $\mathcal{A} \in \mathcal{L}(\mathcal{L}(V))$ על ידי $\mathcal{A}T = T^*$ לכל $T \in \mathcal{L}(V)$.

א. מצאו את כל הערכים העצמיים של $\mathcal{A}$.

ב. מצאו את הפולינום המינימלי של $\mathcal{A}$.

14 הגדירו מכפלה פנימית מעל $\mathcal{P}_2(\mathbf{R})$ על ידי $\langle p,q\rangle = \int_0^1 pq$. הגדירו אופרטור $T \in \mathcal{L}(\mathcal{P}_2(\mathbf{R}))$ על ידי

$$T(ax^2 + bx + c) = bx.$$

א. הראו כי עם מכפלה פנימית זו, האופרטור T אינו צמוד לעצמו.

ב. המטריצה של T ביחס לבסיס $1, x, x^2$ היא

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

מטריצה זו שווה לצמודה־משוחלפת שלה, אף ש־ T אינו צמוד לעצמו. הסבירו מדוע אין הדבר מהווה סתירה.

15 נניח ש־ $T \in \mathcal{L}(V)$ הפיך. הוכיחו כי

א. T צמוד לעצמו $\iff$ T^{-1} צמוד לעצמו.

ב. T נורמלי $\iff$ T^{-1} נורמלי.

16 נניח ש־ $\mathbf{F} = \mathbf{R}$.

א. הראו כי הקבוצה של האופרטורים הצמודים לעצמם מעל V היא תת־מרחב של $\mathcal{L}(V)$.

ב. מהו הממד של התת־מרחב של $\mathcal{L}(V)$ בסעיף (א) [במונחים של $\dim V$]?

17 נניח ש־ $\mathbf{F} = \mathbf{C}$. הראו כי הקבוצה של האופרטורים הצמודים לעצמם מעל V אינה תת־מרחב של $\mathcal{L}(V)$.

18 נניח ש־ $\dim V \geq 2$. הראו כי הקבוצה של האופרטורים הנורמליים מעל V אינה תת־מרחב של $\mathcal{L}(V)$.

19 Suppose $T \in \mathcal{L}(V)$ and $\|T^*v\| \le \|Tv\|$ for every $v \in V$. Prove that T is normal.

This exercise fails on infinite-dimensional inner product spaces, leading to what are called hyponormal operators, which have a well-developed theory.

20 Suppose $P \in \mathcal{L}(V)$ is such that $P^2 = P$. Prove that the following are equivalent.

(a) P is self-adjoint.

(b) P is normal.

(c) There is a subspace U of V such that $P = P_U$.

21 Suppose $D: \mathcal{P}_8(\mathbf{R}) \to \mathcal{P}_8(\mathbf{R})$ is the differentiation operator defined by $Dp = p'$. Prove that there does not exist an inner product on $\mathcal{P}_8(\mathbf{R})$ that makes D a normal operator.

22 Give an example of an operator $T \in \mathcal{L}(\mathbf{R}^3)$ such that T is normal but not self-adjoint.

23 Suppose T is a normal operator on V. Suppose also that $v, w \in V$ satisfy the equations

$$\|v\| = \|w\| = 2, \quad Tv = 3v, \quad Tw = 4w.$$

Show that $\|T(v + w)\| = 10$.

24 Suppose $T \in \mathcal{L}(V)$ and

$$a_0 + a_1 z + a_2 z^2 + \cdots + a_{m-1} z^{m-1} + z^m$$

is the minimal polynomial of T. Prove that the minimal polynomial of T^* is

$$\overline{a_0} + \overline{a_1}\, z + \overline{a_2}\, z^2 + \cdots + \overline{a_{m-1}}\, z^{m-1} + z^m.$$

This exercise shows that the minimal polynomial of T^* equals the minimal polynomial of T if $\mathbf{F} = \mathbf{R}$.

25 Suppose $T \in \mathcal{L}(V)$. Prove that T is diagonalizable if and only if T^* is diagonalizable.

26 Fix $u, x \in V$. Define $T \in \mathcal{L}(V)$ by $Tv = \langle v, u\rangle x$ for every $v \in V$.

(a) Prove that if V is a real vector space, then T is self-adjoint if and only if the list u, x is linearly dependent.

(b) Prove that T is normal if and only if the list u, x is linearly dependent.

27 Suppose $T \in \mathcal{L}(V)$ is normal. Prove that

$$\operatorname{null} T^k = \operatorname{null} T \quad \text{and} \quad \operatorname{range} T^k = \operatorname{range} T$$

for every positive integer k.

28 Suppose $T \in \mathcal{L}(V)$ is normal. Prove that if $\lambda \in \mathbf{F}$, then the minimal polynomial of T is not a polynomial multiple of $(x - \lambda)^2$.

19 נניח ש־ $T \in \mathcal{L}(V)$ ו־ $\|Tv\| \le \|T^*v\|$ לכל $v \in V$. הוכיחו כי T נורמלי.

תרגיל זה אינו מתקיים במרחבי מכפלה פנימית מממד אינסופי, מה שמוביל לאובייקטים הנקראים אופרטורים היפו־נורמליים, שסביבם פותחה תאוריה עניפה.

20 נניח ש־ $P \in \mathcal{L}(V)$ כך ש־ $P^2 = P$. הוכיחו כי התנאים הבאים שקולים.

א. P צמוד לעצמו.

ב. P נורמלי.

ג. קיים תת־מרחב U של V כך ש־ $P = P_U$.

21 נניח ש־ $D: \mathcal{P}_8(\mathbf{R}) \to \mathcal{P}_8(\mathbf{R})$ הוא אופרטור הגזירה המוגדר על ידי $Dp = p'$. הוכיחו כי לא קיימת מכפלה פנימית מעל $\mathcal{P}_8(\mathbf{R})$ שתגרום ל־ D להיות אופרטור נורמלי.

22 הביאו דוגמה של אופרטור $T \in \mathcal{L}(\mathbf{R}^3)$ כך ש־ T נורמלי אך אינו צמוד לעצמו.

23 נניח ש־ T אופרטור נורמלי מעל V. בנוסף, נניח ש־ $v, w \in V$ מקיימים את המשוואות

$$\|v\| = \|w\| = 2, \quad Tv = 3v, \quad Tw = 4w.$$

הראו כי $\|T(v + w)\| = 10$.

24 נניח ש־ $T \in \mathcal{L}(V)$ ו־

$$a_0 + a_1 z + a_2 z^2 + \cdots + a_{m-1} z^{m-1} + z^m$$

הוא הפולינום המינימלי של T. הוכיחו כי הפולינום המינימלי של T^* הוא

$$\overline{a_0} + \overline{a_1}\, z + \overline{a_2}\, z^2 + \cdots + \overline{a_{m-1}}\, z^{m-1} + z^m.$$

תרגיל זה מראה כי הפולינום המינימלי של T^* שווה לפולינום המינימלי של T אם $\mathbf{F} = \mathbf{R}$.

25 נניח ש־ $T \in \mathcal{L}(V)$. הוכיחו כי T לכסין אם־ורק־אם T^* לכסין.

26 נְקַבֵּעַ $u, x \in V$. נגדיר $T \in \mathcal{L}(V)$ על ידי $Tv = \langle v, u \rangle x$ לכל $v \in V$.

א. הוכיחו כי אם V מרחב וקטורי ממשי, אזי T צמוד לעצמו אם־ורק־אם הרשימה u, x תלויה ליניארית.

ב. הוכיחו כי T נורמלי אם־ורק־אם הרשימה u, x תלויה ליניארית.

27 נניח ש־ $T \in \mathcal{L}(V)$ נורמלי. הוכיחו כי

$$\operatorname{range} T^k = \operatorname{range} T \quad \text{וגם} \quad \operatorname{null} T^k = \operatorname{null} T$$

לכל מספר שלם חיובי k.

28 נניח ש־ $T \in \mathcal{L}(V)$ נורמלי. הוכיחו כי אם $\lambda \in \mathbf{F}$, אזי הפולינום המינימלי של T אינו כפולת פולינום של $(x - \lambda)^2$.

29 Prove or give a counterexample: If $T \in \mathcal{L}(V)$ and there is an orthonormal basis $e_1, \dots, e_n$ of V such that $\|Te_k\| = \|T^*e_k\|$ for each $k = 1, \dots, n$, then T is normal.

30 Suppose that $T \in \mathcal{L}(\mathbf{F}^3)$ is normal and $T(1,1,1) = (2,2,2)$. Suppose $(z_1, z_2, z_3) \in \operatorname{null} T$. Prove that $z_1 + z_2 + z_3 = 0$.

31 Fix a positive integer n. In the inner product space of continuous real-valued functions on $[-\pi, \pi]$ with inner product $\langle f, g \rangle = \int_{-\pi}^{\pi} fg$, let

$$V = \operatorname{span}(1, \cos x, \cos 2x, \dots, \cos nx, \sin x, \sin 2x, \dots, \sin nx).$$

(a) Define $D \in \mathcal{L}(V)$ by $Df = f'$. Show that $D^* = -D$. Conclude that D is normal but not self-adjoint.

(b) Define $T \in \mathcal{L}(V)$ by $Tf = f''$. Show that T is self-adjoint.

32 Suppose $T : V \to W$ is a linear map. Show that under the standard identification of V with V' (see 6.58) and the corresponding identification of W with W', the adjoint map $T^* : W \to V$ corresponds to the dual map $T' : W' \to V'$. More precisely, show that

$$T'(\varphi_w) = \varphi_{T^*w}$$

for all $w \in W$, where φ_w and φ_{T^*w} are defined as in 6.58.

29 הוכיחו או הביאו דוגמה נגדית: אם $T \in \mathcal{L}(V)$ וישנו בסיס אורתונורמלי $e_1, \dots, e_n$ של V כך ש־ $\|Te_k\| = \|T^* e_k\|$ לכל $k = 1, \dots, n$, אזי T נורמלי.

30 נניח ש־ $T \in \mathcal{L}(\mathbf{F}^3)$ נורמלי ו־ $T(1,1,1) = (2,2,2)$. נניח ש־ $(z_1, z_2, z_3) \in \operatorname{null} T$. הוכיחו כי $z_1 + z_2 + z_3 = 0$.

31 נְקַבֵּעַ מספר שלם חיובי n. במרחב המכפלה הפנימית של הפונקציות הממשיות הרציפות מעל $[-\pi, \pi]$ עם מכפלה פנימית $\langle f, g \rangle = \int_{-\pi}^{\pi} fg$, יהי

$$V = \operatorname{span}(1, \cos x, \cos 2x, \dots, \cos nx, \sin x, \sin 2x, \dots, \sin nx).$$

א. הגדירו $D \in \mathcal{L}(V)$ על ידי $Df = f'$. הראו כי $D^* = -D$. הסיקו כי D נורמלי אך אינו צמוד לעצמו.

ב. הגדירו $T \in \mathcal{L}(V)$ על ידי $Tf = f''$. הראו כי T צמוד לעצמו.

32 נניח ש־ $T : V \to W$ העתקה ליניארית. הראו כי תחת הזיהוי הסטנדרטי של V עם V' (ראו 6.58) והזיהוי התואם של W עם W', ההעתקה הצמודה $T^* : W \to V$ מתאימה להעתקה הדואלית $T' : W' \to V'$. ביתר דיוק, הראו כי

$$T'(\varphi_w) = \varphi_{T^* w}$$

לכל $w \in W$, כאשר φ_w ו־ $\varphi_{T^* w}$ מוגדרים כמו ב־ 6.58.

7B Spectral Theorem

Recall that a diagonal matrix is a square matrix that is 0 everywhere except possibly on the diagonal. Recall that an operator on V is called diagonalizable if the operator has a diagonal matrix with respect to some basis of V. Recall also that this happens if and only if there is a basis of V consisting of eigenvectors of the operator (see 5.55).

The nicest operators on V are those for which there is an *orthonormal* basis of V with respect to which the operator has a diagonal matrix. These are precisely the operators $T \in \mathcal{L}(V)$ such that there is an orthonormal basis of V consisting of eigenvectors of T. Our goal in this section is to prove the spectral theorem, which characterizes these operators as the self-adjoint operators when $\mathbf{F} = \mathbf{R}$ and as the normal operators when $\mathbf{F} = \mathbf{C}$.

The spectral theorem is probably the most useful tool in the study of operators on inner product spaces. Its extension to certain infinite-dimensional inner product spaces (see, for example, Section 10D of the author's book *Measure, Integration & Real Analysis*) plays a key role in functional analysis.

Because the conclusion of the spectral theorem depends on $\mathbf{F}$, we will break the spectral theorem into two pieces, called the real spectral theorem and the complex spectral theorem.

Real Spectral Theorem

To prove the real spectral theorem, we will need two preliminary results. These preliminary results hold on both real and complex inner product spaces, but they are not needed for the proof of the complex spectral theorem.

This completing-the-square technique can be used to derive the quadratic formula.

You could guess that the next result is true and even discover its proof by thinking about quadratic polynomials with real coefficients. Specifically, suppose $b, c \in \mathbf{R}$ and $b^2 < 4c$. Let x be a real number. Then

$$x^2 + bx + c = \left(x + \frac{b}{2}\right)^2 + \left(c - \frac{b^2}{4}\right) > 0.$$

In particular, $x^2 + bx + c$ is an invertible real number (a convoluted way of saying that it is not 0). Replacing the real number x with a self-adjoint operator (recall the analogy between real numbers and self-adjoint operators) leads to the next result.

7.26 *invertible quadratic expressions*

Suppose $T \in \mathcal{L}(V)$ is self-adjoint and $b, c \in \mathbf{R}$ are such that $b^2 < 4c$. Then

$$T^2 + bT + cI$$

is an invertible operator.

7ג משפט הפירוק הספקטרלי

נזכיר כי מטריצה אלכסונית היא מטריצה ריבועית שאיבריה 0 בכל מקום פרט לאלכסון שבו ייתכנו איברים השונים מ־ 0. נזכיר כי אופרטור מעל V נקרא לכסין אם לאופרטור יש מטריצה אלכסונית ביחס לבסיס כלשהו של V. נזכיר בנוסף שהדבר קורה אם־ורק־אם קיים בסיס של V המורכב מווקטורים עצמיים של האופרטור (ראו 5.55).

האופרטורים הידידותיים ביותר מעל V הם אלו אשר עבורם קיים בסיס אורתונורמלי של V אשר ביחס אליו יש לאופרטור מטריצה אלכסונית. אלו הם בדיוק האופרטורים $T \in \mathcal{L}(V)$ כך שקיים בסיס אורתונורמלי של V המורכב מווקטורים עצמיים של T. מטרתנו ביחידה זו היא להוכיח את משפט הפירוק הספקטרלי, המאפיין אופרטורים אלו כאופרטורים צמודים לעצמם כאשר $\mathbf{F} = \mathbf{R}$ וכאופרטורים נורמליים כאשר $\mathbf{F} = \mathbf{C}$.

משפט הפירוק הספקטרלי הוא ככל הנראה הכלי השימושי ביותר בחקר אופרטורים מעל מרחבי מכפלה פנימית. להרחבתו למרחבי מכפלה פנימית מסוימים מממד אינסופי (ראו, לדוגמה, חלק D10 בספרו של המחבר Measure, Integration & Real Analysis) תפקיד מרכזי באנליזה פונקציונלית.

מאחר שהמסקנה של משפט הפירוק הספקטרלי תלויה ב־ $\mathbf{F}$, אנו נפצל את משפט הפירוק הספקטרלי לשני חלקים, שלהם נקרא משפט הפירוק הספקטרלי הממשי ומשפט הפירוק הספקטרלי המרוכב.

משפט הפירוק הספקטרלי הממשי

להוכחת משפט הפירוק הספקטרלי הממשי, נזדקק לשתי מסקנות מקדימות. מסקנות מקדימות אלו מתקיימות עבור שני מרחבי המכפלה הפנימית הממשי והמרוכב, אך איננו זקוקים להן להוכחת משפט הפירוק הספקטרלי המרוכב.

טכניקת השלמה לריבוע זו היא כלי שימושי לגזירת נוסחת השורשים.

תוכלו לנחש שהמסקנה הבאה נכונה ואפילו למצוא את ההוכחה עבורה בכך שתתבוננו בפולינום ממעלה שנייה עם מקדמים ממשיים. ליתר דיוק, נניח ש־ $b, c \in \mathbf{R}$ ו־ $b^2 < 4c$. יהי x מספר ממשי. אזי

$$x^2 + bx + c = \left(x + \frac{b}{2}\right)^2 + \left(c - \frac{b^2}{4}\right) > 0.$$

בפרט, $x^2 + bx + c$ הוא מספר ממשי הפיך (דרך מפותלת לומר שהוא שונה מ־ 0). החלפת המספר הממשי x באופרטור צמוד לעצמו (נזכיר את האנלוגיה בין מספרים ממשיים ואופרטורים צמודים לעצמם) מובילה אל המסקנה הבאה.

7.26 ביטויים ריבועיים הפיכים

נניח ש־ $T \in \mathcal{L}(V)$ צמוד לעצמו ו־ $b, c \in \mathbf{R}$ כך ש־ $b^2 < 4c$. אזי

$$T^2 + bT + cI$$

אופרטור הפיך.

Proof Let v be a nonzero vector in V. Then

$$\begin{aligned}\langle (T^2 + bT + cI)v, v\rangle &= \langle T^2 v, v\rangle + b\langle Tv, v\rangle + c\langle v, v\rangle \\ &= \langle Tv, Tv\rangle + b\langle Tv, v\rangle + c\|v\|^2 \\ &\geq \|Tv\|^2 - |b|\,\|Tv\|\,\|v\| + c\|v\|^2 \\ &= \left(\|Tv\| - \frac{|b|\,\|v\|}{2}\right)^2 + \left(c - \frac{b^2}{4}\right)\|v\|^2 \\ &> 0,\end{aligned}$$

where the third line above holds by the Cauchy–Schwarz inequality (6.14). The last inequality implies that $(T^2 + bT + cI)v \neq 0$. Thus $T^2 + bT + cI$ is injective, which implies that it is invertible (see 3.65). ∎

The next result will be a key tool in our proof of the real spectral theorem.

7.27 *minimal polynomial of self-adjoint operator*

Suppose $T \in \mathcal{L}(V)$ is self-adjoint. Then the minimal polynomial of T equals $(z - \lambda_1) \cdots (z - \lambda_m)$ for some $\lambda_1, \ldots, \lambda_m \in \mathbf{R}$.

Proof First suppose $\mathbf{F} = \mathbf{C}$. The zeros of the minimal polynomial of T are the eigenvalues of T [by 5.27(a)]. All eigenvalues of T are real (by 7.12). Thus the second version of the fundamental theorem of algebra (see 4.13) tells us that the minimal polynomial of T has the desired form.

Now suppose $\mathbf{F} = \mathbf{R}$. By the factorization of a polynomial over $\mathbf{R}$ (see 4.16) there exist $\lambda_1, \ldots, \lambda_m \in \mathbf{R}$ and $b_1, \ldots, b_N, c_1, \ldots, c_N \in \mathbf{R}$ with $b_k^2 < 4c_k$ for each k such that the minimal polynomial of T equals

$$(z - \lambda_1) \cdots (z - \lambda_m)(z^2 + b_1 z + c_1) \cdots (z^2 + b_N z + c_N); \tag{7.28}$$

here either m or N might equal 0, meaning that there are no terms of the corresponding form. Now

$$(T - \lambda_1 I) \cdots (T - \lambda_m I)(T^2 + b_1 T + c_1 I) \cdots (T^2 + b_N T + c_N I) = 0.$$

If $N > 0$, then we could multiply both sides of the equation above on the right by the inverse of $T^2 + b_N T + c_N I$ (which is an invertible operator by 7.26) to obtain a polynomial expression of T that equals 0. The corresponding polynomial would have degree two less than the degree of 7.28, violating the minimality of the degree of the polynomial with this property. Thus we must have $N = 0$, which means that the minimal polynomial in 7.28 has the form $(z - \lambda_1) \cdots (z - \lambda_m)$, as desired. ∎

The result above along with 5.27(a) implies that every self-adjoint operator has an eigenvalue. In fact, as we will see in the next result, self-adjoint operators have enough eigenvectors to form a basis.

הוכחה יהי v וקטור שונה מאפס ב־ V. אזי

$$\begin{aligned}\langle (T^2+bT+cI)v, v\rangle &= \langle T^2v, v\rangle + b\langle Tv, v\rangle + c\langle v, v\rangle\\ &= \langle Tv, Tv\rangle + b\langle Tv, v\rangle + c\|v\|^2\\ &\geq \|Tv\|^2 - |b|\,\|Tv\|\,\|v\| + c\|v\|^2\\ &= \left(\|Tv\| - \frac{|b|\,\|v\|}{2}\right)^2 + \left(c - \frac{b^2}{4}\right)\|v\|^2\\ &> 0,\end{aligned}$$

כאשר השורה השלישית לעיל מתקיימת לפי אי־שוויון קושי־שוורץ (6.14). האי־שוויון האחרון גורר ש־ $(T^2+bT+cI)v \neq 0$. לפיכך $T^2+bT+cI$ חד־חד־ערכי, מה שגורר שהוא הפיך (ראו 3.65). ■

המסקנה הבאה תשמש ככלי מרכזי בהוכחתנו את משפט הפירוק הספקטרלי הממשי.

7.27 פולינום מינימלי של אופרטור צמוד לעצמו

נניח ש־ $T \in \mathcal{L}(V)$ צמוד לעצמו. אזי הפולינום המינימלי של T שווה ל־ $(z-\lambda_1)\cdots(z-\lambda_m)$ עבור $\lambda_1, \ldots, \lambda_m \in \mathbf{R}$ כלשהם.

הוכחה תחילה נניח כי $\mathbf{F} = \mathbf{C}$. שורשי הפולינום המינימלי של T הם הערכים העצמיים של T [לפי 5.27(א)]. כל הערכים העצמיים של T ממשיים (לפי 7.12). לפיכך המשפט היסודי של האלגברה בגרסתו השנייה (ראו 4.13) מורה לנו כי לפולינום המינימלי של T יש את הצורה הרצויה.

כעת נניח ש־ $\mathbf{F} = \mathbf{R}$. לפי הפירוק לגורמים של פולינום מעל $\mathbf{R}$ (ראו 4.16), קיימים $\lambda_1, \ldots, \lambda_m \in \mathbf{R}$ ו־ $b_1, \ldots, b_N, c_1, \ldots, c_N \in \mathbf{R}$ עם $b_k^2 < 4c_k$ לכל k כך שהפולינום המינימלי של T שווה ל־

$$(z-\lambda_1)\cdots(z-\lambda_m)(z^2+b_1z+c_1)\cdots(z^2+b_Nz+c_N); \tag{7.28}$$

כאשר ייתכן ש־ m או N שווים 0, כלומר אין איברים מהצורה המתאימה. כעת

$$(T-\lambda_1 I)\cdots(T-\lambda_m I)(T^2+b_1T+c_1I)\cdots(T^2+b_NT+c_NI) = 0.$$

אם $N > 0$, אזי נוכל להכפיל כל אגף של המשוואה לעיל בהופכי של $T^2+b_NT+c_NI$ מימין (שהוא אופרטור הפיך לפי 7.26) לקבלת ביטוי פולינומי של T השווה 0. מעלת הפולינום המתאים תהיה פחותה בשתיים ממעלת 7.28, שזוהי הפרת מינימליות המעלה של הפולינום המינימלי. לפיכך האפשרות היחידה היא ש־ $N = 0$, ומשמעות הדבר היא שלפולינום המינימלי ב־ 7.28 יש את התבנית $(z-\lambda_1)\cdots(z-\lambda_m)$, כפי שרצינו. ■

המסקנה לעיל יחד עם 5.27(א) גוררת שלכל אופרטור צמוד לעצמו יש ערך עצמי. למעשה, כפי שנראה במסקנה הבאה, לאופרטורים צמודים לעצמם יש מספיק וקטורים עצמיים כדי ליצור בסיס.

The next result, which gives a complete description of the self-adjoint operators on a real inner product space, is one of the major theorems in linear algebra.

7.29 *real spectral theorem*

Suppose $\mathbf{F} = \mathbf{R}$ and $T \in \mathcal{L}(V)$. Then the following are equivalent.

(a) T is self-adjoint.

(b) T has a diagonal matrix with respect to some orthonormal basis of V.

(c) V has an orthonormal basis consisting of eigenvectors of T.

Proof First suppose (a) holds, so T is self-adjoint. Our results on minimal polynomials, specifically 6.37 and 7.27, imply that T has an upper-triangular matrix with respect to some orthonormal basis of V. With respect to this orthonormal basis, the matrix of T^* is the transpose of the matrix of T. However, $T^* = T$. Thus the transpose of the matrix of T equals the matrix of T. Because the matrix of T is upper-triangular, this means that all entries of the matrix above and below the diagonal are 0. Hence the matrix of T is a diagonal matrix with respect to the orthonormal basis. Thus (a) implies (b).

Conversely, now suppose (b) holds, so T has a diagonal matrix with respect to some orthonormal basis of V. That diagonal matrix equals its transpose. Thus with respect to that basis, the matrix of T^* equals the matrix of T. Hence $T^* = T$, proving that (b) implies (a).

The equivalence of (b) and (c) follows from the definitions [or see the proof that (a) and (b) are equivalent in 5.55]. ∎

7.30 example: *an orthonormal basis of eigenvectors for an operator*

Consider the operator T on $\mathbf{R}^3$ whose matrix (with respect to the standard basis) is

$$\begin{pmatrix} 14 & -13 & 8 \\ -13 & 14 & 8 \\ 8 & 8 & -7 \end{pmatrix}.$$

This matrix with real entries equals its transpose; thus T is self-adjoint. As you can verify,

$$\frac{(1,-1,0)}{\sqrt{2}}, \frac{(1,1,1)}{\sqrt{3}}, \frac{(1,1,-2)}{\sqrt{6}}$$

is an orthonormal basis of $\mathbf{R}^3$ consisting of eigenvectors of T. With respect to this basis, the matrix of T is the diagonal matrix

$$\begin{pmatrix} 27 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & -15 \end{pmatrix}.$$

See Exercise 17 for a version of the real spectral theorem that applies simultaneously to more than one operator.

המסקנה הבאה, המספקת תיאור מלא של אופרטורים הצמודים לעצמם מעל מרחב מכפלה פנימית ממשי, היא אחד המשפטים המרכזיים באלגברה ליניארית.

7.29 משפט הפירוק הספקטרלי הממשי

נניח ש־ $\mathbf{F} = \mathbf{R}$ ו־ $T \in \mathcal{L}(V)$. אזי התנאים הבאים שקולים.

א. T צמוד לעצמו.

ב. יש ל־ T מטריצה אלכסונית ביחס לבסיס אורתונורמלי כלשהו של V.

ג. יש ל־ V בסיס אורתונורמלי המורכב מווקטורים עצמיים של T.

הוכחה תחילה נניח כי (א) מתקיים, כלומר T צמוד לעצמו. המסקנות שלנו הנוגעות לפולינום המינימלי, בפרט 6.37 ו־ 7.27, גוררות שיש ל־ T מטריצה משולשית עליונה ביחס לבסיס אורתונורמלי כלשהו של V. ביחס לבסיס אורתונורמלי זה, המטריצה של T^* היא המשוחלפת של המטריצה של T. אולם, $T^* = T$. לפיכך המשוחלפת של המטריצה של T שווה למטריצה של T. מאחר שהמטריצה של T משולשית עליונה, משמעות הדבר כי כל איברי המטריצה מעל ומתחת האלכסון הם 0. לכן המטריצה של T מטריצה אלכסונית ביחס אל הבסיס האורתונורמלי. לפיכך (א) גורר את (ב).

בכיוון ההפוך, כעת נניח ש־ (ב) מתקיים, כלומר יש ל־ T מטריצה אלכסונית ביחס לבסיס אורתונורמלי כלשהו של V. מטריצה אלכסונית זו שווה למשוחלפת שלה. לפיכך ביחס לבסיס זה, המטריצה של T^* שווה למטריצה של T. לכן $T^* = T$, ובכך הוכחנו כי (ב) גורר את (א).

השקילות בין (ב) ל־ (ג) נובעת מההגדרות [לחילופין, ראו את ההוכחה ש־ (א) ו־ (ב) שקולים ב־ 5.55].

7.30 דוגמה: בסיס אורתונורמלי של וקטוריס עצמיים עבור אופרטור

נתבונן באופרטור T מעל $\mathbf{R}^3$ שהמטריצה שלו (ביחס לבסיס הסטנדרטי) היא

$$\begin{pmatrix} 14 & -13 & 8 \\ -13 & 14 & 8 \\ 8 & 8 & -7 \end{pmatrix}.$$

מטריצה זו עם איברים ממשיים שווה למשוחלפת שלה; לפיכך T צמוד לעצמו. כפי שאתם יכולים לוודא,

$$\frac{(1,-1,0)}{\sqrt{2}}, \frac{(1,1,1)}{\sqrt{3}}, \frac{(1,1,-2)}{\sqrt{6}}$$

הוא בסיס אורתונורמלי של $\mathbf{R}^3$ המורכב מווקטורים עצמיים של T. ביחס לבסיס זה, המטריצה של T היא המטריצה האלכסונית

$$\begin{pmatrix} 27 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & -15 \end{pmatrix}.$$

ראו תרגיל 17 לגרסה של משפט הפירוק הספקטרלי הממשי המופעל בו זמנית על יותר מאופרטור אחד.

Complex Spectral Theorem

The next result gives a complete description of the normal operators on a complex inner product space.

7.31 *complex spectral theorem*

Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$. Then the following are equivalent.

(a) T is normal.

(b) T has a diagonal matrix with respect to some orthonormal basis of V.

(c) V has an orthonormal basis consisting of eigenvectors of T.

Proof First suppose (a) holds, so T is normal. By Schur's theorem (6.38), there is an orthonormal basis $e_1, \dots, e_n$ of V with respect to which T has an upper-triangular matrix. Thus we can write

7.32
$$\mathcal{M}(T, (e_1, \dots, e_n)) = \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ & \ddots & \vdots \\ 0 & & a_{n,n} \end{pmatrix}.$$

We will show that this matrix is actually a diagonal matrix.

We see from the matrix above that

$$\|Te_1\|^2 = |a_{1,1}|^2,$$
$$\|T^*e_1\|^2 = |a_{1,1}|^2 + |a_{1,2}|^2 + \cdots + |a_{1,n}|^2.$$

Because T is normal, $\|Te_1\| = \|T^*e_1\|$ (see 7.20). Thus the two equations above imply that all entries in the first row of the matrix in 7.32, except possibly the first entry $a_{1,1}$, equal 0.

Now 7.32 implies

$$\|Te_2\|^2 = |a_{2,2}|^2$$

(because $a_{1,2} = 0$, as we showed in the paragraph above) and

$$\|T^*e_2\|^2 = |a_{2,2}|^2 + |a_{2,3}|^2 + \cdots + |a_{2,n}|^2.$$

Because T is normal, $\|Te_2\| = \|T^*e_2\|$. Thus the two equations above imply that all entries in the second row of the matrix in 7.32, except possibly the diagonal entry $a_{2,2}$, equal 0.

Continuing in this fashion, we see that all nondiagonal entries in the matrix 7.32 equal 0. Thus (b) holds, completing the proof that (a) implies (b).

Now suppose (b) holds, so T has a diagonal matrix with respect to some orthonormal basis of V. The matrix of T^* (with respect to the same basis) is obtained by taking the conjugate transpose of the matrix of T; hence T^* also has a diagonal matrix. Any two diagonal matrices commute; thus T commutes with T^*, which means that T is normal. In other words, (a) holds, completing the proof that (b) implies (a).

The equivalence of (b) and (c) follows from the definitions (also see 5.55). ∎

משפט הפירוק הספקטרלי המרוכב

המסקנה הבאה מספקת תיאור מלא של אופרטורים נורמליים מעל מרחב מכפלה פנימית מרוכב.

7.31 משפט הפירוק הספקטרלי המרוכב

נניח ש־ $\mathbf{F} = \mathbf{C}$ ו־ $T \in \mathcal{L}(V)$. אזי התנאים הבאים שקולים.

א. T נורמלי.

ב. יש ל־ T מטריצה אלכסונית ביחס לבסיס אורתונורמלי כלשהו של V.

ג. יש ל־ V בסיס אורתונורמלי המורכב מווקטורים עצמיים של T.

הוכחה תחילה נניח כי (א) מתקיים, כלומר T נורמלי. לפי משפט הפירוק של שור (6.38), קיים בסיס אורתונורמלי $e_1, \dots, e_n$ של V אשר ביחס אליו יש ל־ T מטריצה משולשית עליונה. לפיכך נוכל לכתוב

$$\mathcal{M}(T, (e_1, \dots, e_n)) = \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ & \ddots & \vdots \\ 0 & & a_{n,n} \end{pmatrix}. \tag{7.32}$$

אנו נַרְאֶה כי מטריצה זו היא למעשה מטריצה אלכסונית.

מהמטריצה לעיל אנו רואים ש־

$$\|Te_1\|^2 = |a_{1,1}|^2,$$
$$\|T^*e_1\|^2 = |a_{1,1}|^2 + |a_{1,2}|^2 + \cdots + |a_{1,n}|^2.$$

מאחר ש־ T אופרטור נורמלי, $\|Te_1\| = \|T^*e_1\|$ (ראו 7.20). לפיכך שתי המשוואות לעיל גוררות שכל האיברים בשורה הראשונה של המטריצה ב 7.32, למעט אולי האיבר הראשון $a_{1,1}$, שווים 0.

כעת 7.32 גורר ש־

$$\|Te_2\|^2 = |a_{2,2}|^2$$

(מאחר ש־ $a_{1,2} = 0$, כפי שהראינו בפסקה הקודמת) כמו גם

$$\|T^*e_2\|^2 = |a_{2,2}|^2 + |a_{2,3}|^2 + \cdots + |a_{2,n}|^2.$$

מאחר ש־ T אופרטור נורמלי, $\|Te_2\| = \|T^*e_2\|$. לפיכך שתי המשוואות לעיל גוררות שכל האיברים בשורה השנייה של המטריצה ב 7.32, למעט אולי איבר האלכסון $a_{2,2}$, שווים 0.

נמשיך בדרך זו, ונראה כי כל האיברים שאינם באלכסון המטריצה 7.32 שווים 0. לפיכך (ב) מתקיים, ובכך השלמנו את ההוכחה כי (א) גורר את (ב).

כעת נניח כי (ב) מתקיים, כלומר יש ל־ T מטריצה אלכסונית ביחס לבסיס אורתונורמלי כלשהו של V. המטריצה של T^* (ביחס אל אותו הבסיס) מתקבלת על ידי לקיחת הצמודה־משוחלפת של המטריצה של T; לכן גם ל־ T^* יש מטריצה אלכסונית. כל שתי מטריצות אלכסוניות מתחלפות; לפיכך T מתחלף עם T^*, משמעות הדבר ש־ T נורמלי. במילים אחרות, (א) מתקיים, ובכך הושלמה ההוכחה כי (ב) גורר את (א).

השקילות של (ב) ו־ (ג) נובעת מההגדרות (ראו בנוסף 5.55). ■

See Exercises 13 and 20 for alternative proofs that (a) implies (b) in the previous result.

Exercises 14 and 15 interpret the real spectral theorem and the complex spectral theorem by expressing the domain space as an orthogonal direct sum of eigenspaces.

See Exercise 16 for a version of the complex spectral theorem that applies simultaneously to more than one operator.

The main conclusion of the complex spectral theorem is that every normal operator on a complex finite-dimensional inner product space is diagonalizable by an orthonormal basis, as illustrated by the next example.

7.33 example: *an orthonormal basis of eigenvectors for an operator*

Consider the operator $T \in \mathcal{L}(\mathbf{C}^2)$ defined by $T(w, z) = (2w - 3z, 3w + 2z)$. The matrix of T (with respect to the standard basis) is

$$\begin{pmatrix} 2 & -3 \\ 3 & 2 \end{pmatrix}.$$

As we saw in Example 7.19, T is a normal operator.

As you can verify,

$$\tfrac{1}{\sqrt{2}}(i, 1), \tfrac{1}{\sqrt{2}}(-i, 1)$$

is an orthonormal basis of $\mathbf{C}^2$ consisting of eigenvectors of T, and with respect to this basis the matrix of T is the diagonal matrix

$$\begin{pmatrix} 2 + 3i & 0 \\ 0 & 2 - 3i \end{pmatrix}.$$

Exercises 7B

1 Prove that a normal operator on a complex inner product space is self-adjoint if and only if all its eigenvalues are real.

This exercise strengthens the analogy (for normal operators) between self-adjoint operators and real numbers.

2 Suppose $\mathbf{F} = \mathbf{C}$. Suppose $T \in \mathcal{L}(V)$ is normal and has only one eigenvalue. Prove that T is a scalar multiple of the identity operator.

3 Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$ is normal. Prove that the set of eigenvalues of T is contained in $\{0, 1\}$ if and only if there is a subspace U of V such that $T = P_U$.

4 Prove that a normal operator on a complex inner product space is skew (meaning it equals the negative of its adjoint) if and only if all its eigenvalues are purely imaginary (meaning that they have real part equal to 0).

ראו תרגילים 13 ו־ 20 להוכחות חלופיות לכך ש־ (א) גורר את (ב) במסקנה הקודמת.

תרגילים 14 ו־ 15 מפרשים את משפט הפירוק הספקטרלי הממשי ומשפט הפירוק הספקטרלי המרוכב בכך שהם מבטאים את מרחב התחום כסכום ישר אורתוגונלי של המרחבים העצמיים.

ראו תרגיל 16 לגרסה של משפט הפירוק הספקטרלי המרוכב המופעל בו זמנית על יותר מאופרטור אחד.

המסקנה המרכזית של משפט הפירוק הספקטרלי המרוכב היא שכל אופרטור נורמלי מעל מרחב מכפלה פנימית מרוכב מממד סופי לכסין באמצעות בסיס אורתונורמלי, כפי שממחישה הדוגמה הבאה.

דוגמה: בסיס אורתונורמלי של וקטורים עצמיים עבור אופרטור 7.33

נתבונן באופרטור $T \in \mathcal{L}(\mathbf{C}^2)$ המוגדר על ידי $T(w,z) = (2w - 3z, 3w + 2z)$. המטריצה של T (ביחס אל הבסיס הסטנדרטי) היא

$$\begin{pmatrix} 2 & -3 \\ 3 & 2 \end{pmatrix}.$$

כפי שראינו בדוגמה 7.19, T אופרטור נורמלי.

כפי שבאפשרותכם לוודא,

$$\tfrac{1}{\sqrt{2}}(i,1), \tfrac{1}{\sqrt{2}}(-i,1)$$

היא בסיס אורתונורמלי של $\mathbf{C}^2$ המורכב מווקטורים עצמיים של T, וביחס לבסיס זה המטריצה של T היא המטריצה האלכסונית

$$\begin{pmatrix} 2+3i & 0 \\ 0 & 2-3i \end{pmatrix}.$$

תרגילים 7ז

1 הוכיחו כי אופרטור נורמלי מעל מרחב מכפלה פנימית מרוכב צמוד לעצמו אם־ורק־אם כל הערכים העצמיים שלו ממשיים.

תרגיל זה מחזק את האנלוגיה (עבור אופרטורים נורמליים) בין אופרטורים צמודים לעצמם ומספרים ממשיים.

2 נניח ש־ $\mathbf{F} = \mathbf{C}$. נניח ש־ $T \in \mathcal{L}(V)$ נורמלי ויש לו ערך עצמי אחד בלבד. הוכיחו כי T כפולת סקלר של אופרטור הזהות.

3 נניח ש־ $\mathbf{F} = \mathbf{C}$ ו־ $T \in \mathcal{L}(V)$ נורמלי. הוכיחו כי קבוצת הערכים העצמיים של T מוכלת ב־ $\{0,1\}$ אם־ורק־אם ישנו תת־מרחב U של V כך ש־ $T = P_U$.

4 הוכיחו כי אופרטור נורמלי מעל מרחב מכפלה פנימית מרוכב הוא אנטי־צמוד (כלומר הוא שווה למינוס של האופרטור הצמוד שלו) אם־ורק־אם כל הערכים העצמיים שלו מדומים טהורים (כלומר החלק הממשי שבהם שווה 0).

אלגברה ליניארית בדרך הישר, מהדורה רביעית, מחבר: שלדון אקסלר
Linear Algebra Done Right, fourth edition, by Sheldon Axler

5 Prove or give a counterexample: If $T \in \mathcal{L}(\mathbf{C}^3)$ is a diagonalizable operator, then T is normal (with respect to the usual inner product).

6 Suppose V is a complex inner product space and $T \in \mathcal{L}(V)$ is a normal operator such that $T^9 = T^8$. Prove that T is self-adjoint and $T^2 = T$.

7 Give an example of an operator T on a complex vector space such that $T^9 = T^8$ but $T^2 \neq T$.

8 Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$. Prove that T is normal if and only if every eigenvector of T is also an eigenvector of T^*.

9 Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$. Prove that T is normal if and only if there exists a polynomial $p \in \mathcal{P}(\mathbf{C})$ such that $T^* = p(T)$.

10 Suppose V is a complex inner product space. Prove that every normal operator on V has a square root.

*An operator $S \in \mathcal{L}(V)$ is called a **square root** of $T \in \mathcal{L}(V)$ if $S^2 = T$. We will discuss more about square roots of operators in Sections 7C and 8C.*

11 Prove that every self-adjoint operator on V has a cube root.

*An operator $S \in \mathcal{L}(V)$ is called a **cube root** of $T \in \mathcal{L}(V)$ if $S^3 = T$.*

12 Suppose V is a complex vector space and $T \in \mathcal{L}(V)$ is normal. Prove that if S is an operator on V that commutes with T, then S commutes with T^*.

The result in this exercise is called Fuglede's theorem.

13 Without using the complex spectral theorem, use the version of Schur's theorem that applies to two commuting operators (take $\mathcal{E} = \{T, T^*\}$ in Exercise 20 in Section 6B) to give a different proof that if $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$ is normal, then T has a diagonal matrix with respect to some orthonormal basis of V.

14 Suppose $\mathbf{F} = \mathbf{R}$ and $T \in \mathcal{L}(V)$. Prove that T is self-adjoint if and only if all pairs of eigenvectors corresponding to distinct eigenvalues of T are orthogonal and $V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T)$, where $\lambda_1, \ldots, \lambda_m$ denote the distinct eigenvalues of T.

15 Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$. Prove that T is normal if and only if all pairs of eigenvectors corresponding to distinct eigenvalues of T are orthogonal and $V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T)$, where $\lambda_1, \ldots, \lambda_m$ denote the distinct eigenvalues of T.

16 Suppose $\mathbf{F} = \mathbf{C}$ and $\mathcal{E} \subseteq \mathcal{L}(V)$. Prove that there is an orthonormal basis of V with respect to which every element of $\mathcal{E}$ has a diagonal matrix if and only if S and T are commuting normal operators for all $S, T \in \mathcal{E}$.

This exercise extends the complex spectral theorem to the context of a collection of commuting normal operators.

5 הוכיחו או הביאו דוגמה נגדית: אם $T \in \mathcal{L}(\mathbf{C}^3)$ אופרטור לכסין, אזי T נורמלי (ביחס אל המכפלה הפנימית הרגילה).

6 נניח ש־ V מרחב מכפלה פנימית מרוכב ו־ $T \in \mathcal{L}(V)$ אופרטור נורמלי כך ש־ $T^9 = T^8$. הוכיחו כי T צמוד לעצמו וכי $T^2 = T$.

7 הביאו דוגמה של אופרטור T מעל מרחב וקטורי מרוכב כך ש־ $T^9 = T^8$ אבל $T^2 \neq T$.

8 נניח ש־ $\mathbf{F} = \mathbf{C}$ ו־ $T \in \mathcal{L}(V)$. הוכיחו כי T נורמלי אם־ורק־אם כל וקטור עצמי של T הוא גם וקטור עצמי של T^*.

9 נניח ש־ $\mathbf{F} = \mathbf{C}$ ו־ $T \in \mathcal{L}(V)$. הוכיחו כי T נורמלי אם־ורק־אם קיים פולינום $p \in \mathcal{P}(\mathbf{C})$ כך ש־ $T^* = p(T)$.

10 נניח ש־ V מרחב מכפלה פנימית מרוכב. הוכיחו כי לכל אופרטור נורמלי מעל V יש שורש ריבועי.

אופרטור $S \in \mathcal{L}(V)$ נקרא **שורש ריבועי** של $T \in \mathcal{L}(V)$ אם $S^2 = T$. אנו נרחיב על שורשים ריבועיים של אופרטורים ביחידות 3ז ו־ 3ח.

11 הוכיחו כי לכל אופרטור צמוד לעצמו מעל V יש שורש שלישי.

אופרטור $S \in \mathcal{L}(V)$ נקרא **שורש שלישי** של $T \in \mathcal{L}(V)$ אם $S^3 = T$.

12 נניח ש־ V מרחב וקטורי מרוכב ו־ $T \in \mathcal{L}(V)$ נורמלי. הוכיחו כי אם S אופרטור מעל V המתחלף עם T, אזי S מתחלף עם T^*.

המסקנה בתרגיל זה נקראת משפט פוּגְלֶדֶה (Fuglede's theorem).

13 מבלי להשתמש במשפט הפירוק הספקטרלי המרוכב, השתמשו בגרסה של משפט הפירוק של שור המיושם על שני אופרטורים מתחלפים (קחו את $\mathcal{E} = \{T, T^*\}$ בתרגיל 20 ביחידה 2ו) כדי לתת הוכחה שונה לכך שאם $\mathbf{F} = \mathbf{C}$ ו־ $T \in \mathcal{L}(V)$ נורמלי, אזי יש ל־ T מטריצה אלכסונית ביחס לבסיס אורתונורמלי כלשהו של V.

14 נניח ש־ $\mathbf{F} = \mathbf{R}$ ו־ $T \in \mathcal{L}(V)$. הוכיחו כי T צמוד לעצמו אם־ורק־אם כל זוגות הווקטורים העצמיים השייכים לערכים עצמיים שונים של T הם אורתוגונליים וגם $V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T)$, כאשר $\lambda_1, \ldots, \lambda_m$ מסמנים את הערכים העצמיים השונים של T.

15 נניח ש־ $\mathbf{F} = \mathbf{C}$ ו־ $T \in \mathcal{L}(V)$. הוכיחו כי T נורמלי אם־ורק־אם כל הזוגות של וקטורים עצמיים השייכים לערכים עצמיים שונים של T אורתוגונליים ו־ $V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T)$, כאשר $\lambda_1, \ldots, \lambda_m$ מסמנים את הערכים העצמיים השונים של T.

16 נניח ש־ $\mathbf{F} = \mathbf{C}$ ו־ $\mathcal{E} \subseteq \mathcal{L}(V)$. הוכיחו כי קיים בסיס אורתונורמלי של V אשר ביחס אליו לכל איבר של $\mathcal{E}$ יש מטריצה אלכסונית אם־ורק־אם S ו־ T אופרטורים נורמליים מתחלפים לכל $S, T \in \mathcal{E}$.

תרגיל זה מרחיב את משפט הפירוק הספקטרלי המרוכב להקשר של אוסף אופרטורים נורמליים מתחלפים.

17 Suppose $\mathbf{F} = \mathbf{R}$ and $\mathcal{E} \subseteq \mathcal{L}(V)$. Prove that there is an orthonormal basis of V with respect to which every element of $\mathcal{E}$ has a diagonal matrix if and only if S and T are commuting self-adjoint operators for all $S, T \in \mathcal{E}$.

This exercise extends the real spectral theorem to the context of a collection of commuting self-adjoint operators.

18 Give an example of a real inner product space V, an operator $T \in \mathcal{L}(V)$, and real numbers b, c with $b^2 < 4c$ such that

$$T^2 + bT + cI$$

is not invertible.

This exercise shows that the hypothesis that T is self-adjoint cannot be deleted in 7.26, even for real vector spaces.

19 Suppose $T \in \mathcal{L}(V)$ is self-adjoint and U is a subspace of V that is invariant under T.

(a) Prove that $U^\perp$ is invariant under T.
(b) Prove that $T|_U \in \mathcal{L}(U)$ is self-adjoint.
(c) Prove that $T|_{U^\perp} \in \mathcal{L}(U^\perp)$ is self-adjoint.

20 Suppose $T \in \mathcal{L}(V)$ is normal and U is a subspace of V that is invariant under T.

(a) Prove that $U^\perp$ is invariant under T.
(b) Prove that U is invariant under T^*.
(c) Prove that $(T|_U)^* = (T^*)|_U$.
(d) Prove that $T|_U \in \mathcal{L}(U)$ and $T|_{U^\perp} \in \mathcal{L}(U^\perp)$ are normal operators.

This exercise can be used to give yet another proof of the complex spectral theorem (use induction on $\dim V$ and the result that T has an eigenvector).

21 Suppose that T is a self-adjoint operator on a finite-dimensional inner product space and that 2 and 3 are the only eigenvalues of T. Prove that

$$T^2 - 5T + 6I = 0.$$

22 Give an example of an operator $T \in \mathcal{L}(\mathbf{C}^3)$ such that 2 and 3 are the only eigenvalues of T and $T^2 - 5T + 6I \neq 0$.

23 Suppose $T \in \mathcal{L}(V)$ is self-adjoint, $\lambda \in \mathbf{F}$, and $\epsilon > 0$. Suppose there exists $v \in V$ such that $\|v\| = 1$ and

$$\|Tv - \lambda v\| < \epsilon.$$

Prove that T has an eigenvalue λ' such that $|\lambda - \lambda'| < \epsilon$.

This exercise shows that for a self-adjoint operator, a number that is close to satisfying an equation that would make it an eigenvalue is close to an eigenvalue.

17 נניח ש־ $\mathbf{F} = \mathbf{R}$ ו־ $\mathcal{E} \subseteq \mathcal{L}(V)$. הוכיחו כי קיים בסיס אורתונורמלי של V אשר ביחס אליו יש לכל איבר של $\mathcal{E}$ מטריצה אלכסונית אם־ורק־אם S ו־ T אופרטורים מתחלפים הצמודים לעצמם לכל $S, T \in \mathcal{E}$.

תרגיל זה מרחיב את משפט הפירוק הספקטרלי המשי להקשר של אוסף אופרטורים מתחלפים הצמודים לעצמם.

18 הביאו דוגמה של מרחב מכפלה פנימית ממשי V, אופרטור $T \in \mathcal{L}(V)$, ומספרים ממשיים b, c עם $b^2 < 4c$ כך ש־

$$T^2 + bT + cI$$

אינו הפיך.

תרגיל זה מראה כי ההנחה ש־ T צמוד לעצמו הכרחית ב־ 7.26, אפילו עבור מרחבים וקטוריים ממשיים.

19 נניח ש־ $T \in \mathcal{L}(V)$ צמוד לעצמו ו־ U תת־מרחב של V השמור תחת T.

א. הוכיחו כי $U^\perp$ שמור תחת T.

ב. הוכיחו כי $T|_U \in \mathcal{L}(U)$ צמוד לעצמו.

ג. הוכיחו כי $T|_{U^\perp} \in \mathcal{L}(U^\perp)$ צמוד לעצמו.

20 נניח ש־ $T \in \mathcal{L}(V)$ נורמלי ו־ U תת־מרחב של V השמור תחת T.

א. הוכיחו כי $U^\perp$ שמור תחת T.

ב. הוכיחו כי U שמור תחת T^*.

ג. הוכיחו כי $(T|_U)^* = (T^*)|_U$.

ד. הוכיחו כי $T|_U \in \mathcal{L}(U)$ ו־ $T|_{U^\perp} \in \mathcal{L}(U^\perp)$ אופרטורים נורמליים.

תרגיל זה יכול לשמש כהוכחה נוספת של משפט הפירוק הספקטרלי המרוכב (השתמשו באינדוקציה על $\dim V$ ובמסקנה כי יש ל־ T וקטור עצמי).

21 נניח ש־ T אופרטור צמוד לעצמו מעל מרחב מכפלה פנימית מממד סופי וכי 2 ו־ 3 הם הערכים העצמיים היחידים של T. הוכיחו כי

$$T^2 - 5T + 6I = 0.$$

22 הביאו דוגמה של אופרטור $T \in \mathcal{L}(\mathbf{C}^3)$ כך ש־ 2 ו־ 3 הם הערכים העצמיים היחידים של T ו־ $T^2 - 5T + 6I \neq 0$.

23 נניח ש־ $T \in \mathcal{L}(V)$ צמוד לעצמו, $\lambda \in \mathbf{F}$, ו־ $\epsilon > 0$. נניח שקיים $v \in V$ כך ש־ $\|v\| = 1$ וגם

$$\|Tv - \lambda v\| < \epsilon.$$

הוכיחו כי יש ל־ T ערך עצמי λ' כך ש־ $|\lambda - \lambda'| < \epsilon$.

תרגיל זה מראה כי עבור אופרטור צמוד לעצמו, מספר שכמעט מקיים את המשוואה ההופכת אותו לערך עצמי, קרוב לערך עצמי.

24 Suppose U is a finite-dimensional vector space and $T \in \mathcal{L}(U)$.

(a) Suppose $\mathbf{F} = \mathbf{R}$. Prove that T is diagonalizable if and only if there is a basis of U such that the matrix of T with respect to this basis equals its transpose.

(b) Suppose $\mathbf{F} = \mathbf{C}$. Prove that T is diagonalizable if and only if there is a basis of U such that the matrix of T with respect to this basis commutes with its conjugate transpose.

This exercise adds another equivalence to the list of conditions equivalent to diagonalizability in 5.55.

25 Suppose that $T \in \mathcal{L}(V)$ and there is an orthonormal basis $e_1, \dots, e_n$ of V consisting of eigenvectors of T, with corresponding eigenvalues $\lambda_1, \dots, \lambda_n$. Show that if $k \in \{1, \dots, n\}$, then the pseudoinverse $T^\dagger$ satisfies the equation

$$T^\dagger e_k = \begin{cases} \frac{1}{\lambda_k} e_k & \text{if } \lambda_k \neq 0, \\ 0 & \text{if } \lambda_k = 0. \end{cases}$$

24 נניח ש־ U מרחב וקטורי מממד סופי ו־ $T \in \mathcal{L}(U)$.

א. נניח ש־ $\mathbf{F} = \mathbf{R}$. הוכיחו כי T לכסין אם־ורק־אם קיים בסיס של U כך שהמטריצה של T ביחס לבסיס זה שווה למשוחלפת שלה.

ב. נניח ש־ $\mathbf{F} = \mathbf{C}$. הוכיחו כי T לכסין אם־ורק־אם קיים בסיס של U כך שהמטריצה של T ביחס לבסיס זה מתחלפת עם הצמודה־משוחלפת שלה.

תרגיל זה פוסיף עוד תנאי שקול לרשימת התנאים השקולים של לכסינות ב־ 5.55.

25 נניח ש־ $T \in \mathcal{L}(V)$ וקיים בסיס אורתונורמלי $e_1, \dots, e_n$ של V המורכב מווקטורים עצמיים של T, עם ערכים עצמיים מתאימים $\lambda_1, \dots, \lambda_n$. הראו כי אם $k \in \{1, \dots, n\}$, אזי הפסידו־הופכי $T^\dagger$ מקיים את המשוואה

$$T^\dagger e_k = \begin{cases} \frac{1}{\lambda_k} e_k & \text{if } \lambda_k \neq 0, \\ 0 & \text{if } \lambda_k = 0. \end{cases}$$

7C Positive Operators

7.34 definition: *positive operator*

An operator $T \in \mathcal{L}(V)$ is called *positive* if T is self-adjoint and

$$\langle Tv, v\rangle \geq 0$$

for all $v \in V$.

If V is a complex vector space, then the requirement that T be self-adjoint can be dropped from the definition above (by 7.14).

7.35 example: *positive operators*

(a) Let $T \in \mathcal{L}(\mathbf{F}^2)$ be the operator whose matrix (using the standard basis) is $\begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}$. Then T is self-adjoint and $\langle T(w,z), (w,z)\rangle = 2|w|^2 - 2\operatorname{Re}(w\bar{z}) + |z|^2 = |w-z|^2 + |w|^2 \geq 0$ for all $(w,z) \in \mathbf{F}^2$. Thus T is a positive operator.

(b) If U is a subspace of V, then the orthogonal projection P_U is a positive operator, as you should verify.

(c) If $T \in \mathcal{L}(V)$ is self-adjoint and $b, c \in \mathbf{R}$ are such that $b^2 < 4c$, then $T^2 + bT + cI$ is a positive operator, as shown by the proof of 7.26.

7.36 definition: *square root*

An operator R is called a *square root* of an operator T if $R^2 = T$.

7.37 example: *square root of an operator*

If $T \in \mathcal{L}(\mathbf{F}^3)$ is defined by $T(z_1, z_2, z_3) = (z_3, 0, 0)$, then the operator $R \in \mathcal{L}(\mathbf{F}^3)$ defined by $R(z_1, z_2, z_3) = (z_2, z_3, 0)$ is a square root of T because $R^2 = T$, as you can verify.

Because positive operators correspond to nonnegative numbers, better terminology would use the term nonnegative operators. However, operator theorists consistently call these positive operators, so we follow that custom. Some mathematicians use the term ***positive semidefinite operator****, which means the same as positive operator.*

The characterizations of the positive operators in the next result correspond to characterizations of the nonnegative numbers among $\mathbf{C}$. Specifically, a number $z \in \mathbf{C}$ is nonnegative if and only if it has a nonnegative square root, corresponding to condition (d). Also, z is nonnegative if and only if it has a real square root, corresponding to condition (e). Finally, z is nonnegative if and only if there exists $w \in \mathbf{C}$ such that $z = \overline{w}w$, corresponding to condition (f). See Exercise 20 for another condition that is equivalent to being a positive operator.

7ג אופרטורים חיוביים

7.34 הגדרה: אופרטור חיובי

אופרטור $T \in \mathcal{L}(V)$ נקרא חיובי אם T צמוד לעצמו וגם

$$\langle Tv, v\rangle \geq 0$$

לכל $v \in V$.

אם V מרחב וקטורי מרוכב, אזי אפשר להסיר מההגדרה לעיל את הדרישה ש־ T יהיה צמוד לעצמו (לפי 7.14).

7.35 דוגמה: אופרטורים חיוביים

א. יהי $T \in \mathcal{L}(\mathbf{F}^2)$ האופרטור אשר המטריצה שלו (ביחס לבסיס הסטנדרטי) היא $\begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}$. אזי T צמוד לעצמו וגם

$$\begin{aligned}\langle T(w,z),(w,z)\rangle &= 2|w|^2 - 2\operatorname{Re}(w\bar{z}) + |z|^2 \\ &= |w-z|^2 + |w|^2 \\ &\geq 0\end{aligned}$$

לכל $(w,z) \in \mathbf{F}^2$. לפיכך T אופרטור חיובי.

ב. אם U תת־מרחב של V, אזי ההיטל האורתוגונלי P_U הוא אופרטור חיובי, כפי שעליכם לוודא.

ג. אם $T \in \mathcal{L}(V)$ צמוד לעצמו ו־ $b, c \in \mathbf{R}$ הם כך ש־ $b^2 < 4c$, אזי $T^2 + bT + cI$ אופרטור חיובי, כפי שראינו בהוכחת 7.26.

7.36 הגדרה: שורש ריבועי

אופרטור R נקרא שורש ריבועי של אופרטור T אם $R^2 = T$.

7.37 דוגמה: שורש ריבועי של אופרטור

אם $T \in \mathcal{L}(\mathbf{F}^3)$ אופרטור המוגדר על ידי $T(z_1, z_2, z_3) = (z_3, 0, 0)$, אזי האופרטור $R \in \mathcal{L}(\mathbf{F}^3)$ המוגדר על ידי $R(z_1, z_2, z_3) = (z_2, z_3, 0)$ הוא שורש ריבועי של T מאחר ש־ $R^2 = T$, כפי שבאפשרותכם לוודא.

האפיונים של האופרטורים החיוביים במסקנה הבאה תואמים לאפיוני המספרים האי־שליליים ב־ $\mathbf{C}$. בפרט, מספר $z \in \mathbf{C}$ הוא אי־שלילי אם־ורק־אם יש לו שורש ריבועי אי־שלילי, בהתאמה לתנאי (ד). בנוסף, z אי־שלילי אם־ורק־אם יש לו שורש ריבועי ממשי, בהתאמה לתנאי (ה). לבסוף, z אי־שלילי אם־ורק־אם קיים $w \in \mathbf{C}$ כך ש־ $z = \overline{w}w$, בהתאמה לתנאי (ו). ראו תרגיל 20 לתנאי נוסף השקול לכך שאופרטור יהיה חיובי.

מאחר שאופרטורים חיוביים מתאימים למספרים אי־שליליים, נכון יותר היה להשתמש במונח אופרטורים אי־שליליים. אולם, מאחר שהשימוש במונח אופרטור חיובי התקבע בקרב המתמטיקאים, נשתמש בו גם אנו.

7.38 *characterizations of positive operators*

Let $T \in \mathcal{L}(V)$. Then the following are equivalent.

(a) T is a positive operator.

(b) T is self-adjoint and all eigenvalues of T are nonnegative.

(c) With respect to some orthonormal basis of V, the matrix of T is a diagonal matrix with only nonnegative numbers on the diagonal.

(d) T has a positive square root.

(e) T has a self-adjoint square root.

(f) $T = R^*R$ for some $R \in \mathcal{L}(V)$.

Proof We will prove that (a) $\Rightarrow$ (b) $\Rightarrow$ (c) $\Rightarrow$ (d) $\Rightarrow$ (e) $\Rightarrow$ (f) $\Rightarrow$ (a).

First suppose (a) holds, so that T is positive, which implies that T is self-adjoint (by definition of positive operator). To prove the other condition in (b), suppose λ is an eigenvalue of T. Let v be an eigenvector of T corresponding to λ. Then

$$0 \le \langle Tv, v\rangle = \langle \lambda v, v\rangle = \lambda\langle v, v\rangle.$$

Thus λ is a nonnegative number. Hence (b) holds, showing that (a) implies (b).

Now suppose (b) holds, so that T is self-adjoint and all eigenvalues of T are nonnegative. By the spectral theorem (7.29 and 7.31), there is an orthonormal basis $e_1, \dots, e_n$ of V consisting of eigenvectors of T. Let $\lambda_1, \dots, \lambda_n$ be the eigenvalues of T corresponding to $e_1, \dots, e_n$; thus each λ_k is a nonnegative number. The matrix of T with respect to $e_1, \dots, e_n$ is the diagonal matrix with $\lambda_1, \dots, \lambda_n$ on the diagonal, which shows that (b) implies (c).

Now suppose (c) holds. Suppose $e_1, \dots, e_n$ is an orthonormal basis of V such that the matrix of T with respect to this basis is a diagonal matrix with nonnegative numbers $\lambda_1, \dots, \lambda_n$ on the diagonal. The linear map lemma (3.4) implies that there exists $R \in \mathcal{L}(V)$ such that

$$Re_k = \sqrt{\lambda_k}e_k$$

for each $k = 1, \dots, n$. As you should verify, R is a positive operator. Furthermore, $R^2e_k = \lambda_k e_k = Te_k$ for each k, which implies that $R^2 = T$. Thus R is a positive square root of T. Hence (d) holds, which shows that (c) implies (d).

Every positive operator is self-adjoint (by definition of positive operator). Thus (d) implies (e).

Now suppose (e) holds, meaning that there exists a self-adjoint operator R on V such that $T = R^2$. Then $T = R^*R$ (because $R^* = R$). Hence (e) implies (f).

Finally, suppose (f) holds. Let $R \in \mathcal{L}(V)$ be such that $T = R^*R$. Then $T^* = (R^*R)^* = R^*(R^*)^* = R^*R = T$. Hence T is self-adjoint. To complete the proof that (a) holds, note that

$$\langle Tv, v\rangle = \langle R^*Rv, v\rangle = \langle Rv, Rv\rangle \ge 0$$

for every $v \in V$. Thus T is positive, showing that (f) implies (a). ∎

7.38 אפיונים של אופרטורים חיוביים

יהי $T \in \mathcal{L}(V)$. אזי התנאים הבאים שקולים.

א. T אופרטור חיובי.

ב. T צמוד לעצמו וכל הערכים העצמיים של T אי־שליליים.

ג. ביחס לבסיס אורתונורמלי כלשהו של V, המטריצה של T מטריצה אלכסונית עם מספרים אי־שליליים בלבד על האלכסון.

ד. ל־ T יש שורש ריבועי חיובי.

ה. ל־ T יש שורש ריבועי הצמוד לעצמו.

ו. $T = R^*R$ עבור $R \in \mathcal{L}(V)$ כלשהו.

הוכחה אנו נוכיח כי (א) $\Rightarrow$ (ב) $\Rightarrow$ (ג) $\Rightarrow$ (ד) $\Rightarrow$ (ה) $\Rightarrow$ (ו) $\Rightarrow$ (א).

תחילה נניח כי (א) מתקיים, כך ש־ T חיובי, מה שגורר ש־ T צמוד לעצמו (לפי ההגדרה של אופרטור חיובי). להוכחת התנאי הנוסף בסעיף (ב), נניח ש־ λ ערך עצמי של T. יהי v וקטור עצמי של T השייך ל־ λ. אזי

$$0 \le \langle Tv, v\rangle = \langle \lambda v, v\rangle = \lambda\langle v, v\rangle.$$

לפיכך λ מספר אי־שלילי. לכן (ב) מתקיים, ובכך הראינו כי (א) גורר את (ב).

כעת נניח כי (ב) מתקיים, כך ש־ T צמוד לעצמו וכל הערכים העצמיים של T הם אי־שליליים. לפי משפט הפירוק הספקטרלי (7.29 ו־ 7.31), קיים בסיס אורתונורמלי $e_1, \dots, e_n$ של V המורכב מווקטורים עצמיים של T. יהיו $\lambda_1, \dots, \lambda_n$ הערכים העצמיים של T השייכים ל־ $e_1, \dots, e_n$; לפיכך כל λ_k הוא מספר אי־שלילי. המטריצה של T ביחס אל $e_1, \dots, e_n$ היא המטריצה האלכסונית עם $\lambda_1, \dots, \lambda_n$ על האלכסון, מה שמראה ש־ (ב) גורר את (ג).

כעת נניח ש־ (ג) מתקיים. נניח ש־ $e_1, \dots, e_n$ בסיס אורתונורמלי של V כך שהמטריצה של T ביחס לבסיס זה היא מטריצה אלכסונית עם $\lambda_1, \dots, \lambda_n$ מספרים אי־שליליים על האלכסון. מלמת ההעתקה הליניארית (3.4) נובע כי קיים $R \in \mathcal{L}(V)$ כך ש־

$$Re_k = \sqrt{\lambda_k}e_k$$

לכל $k = 1, \dots, n$. כפי שעליכם לוודא, R אופרטור חיובי. יתר על כן, $R^2e_k = \lambda_k e_k = Te_k$ לכל k, מה שגורר ש־ $R^2 = T$. לפיכך R שורש ריבועי חיובי של T. לכן (ד) מתקיים, מה שמראה ש־ (ג) גורר את (ד).

כל אופרטור חיובי צמוד לעצמו (לפי ההגדרה של אופרטור חיובי). לפיכך (ד) גורר את (ה).

כעת נניח ש־ (ה) מתקיים, משמעות הדבר כי קיים אופרטור צמוד לעצמו R מעל V כך ש־ $T = R^2$. אזי $T = R^*R$ (מאחר ש־ $R = R^*$). לכן (ה) גורר את (ו).

לבסוף, נניח ש־ (ו) מתקיים. נגדיר $R \in \mathcal{L}(V)$ כך ש־ $T = R^*R$. אזי נקבל $T^* = (R^*R)^* = R^*(R^*)^* = R^*R = T$. לכן T צמוד לעצמו. להשלמת ההוכחה כי (א) מתקיים, שימו לב ש־

$$\langle Tv, v\rangle = \langle R^*Rv, v\rangle = \langle Rv, Rv\rangle \ge 0$$

לכל $v \in V$. לפיכך T חיובי, ובכך הראינו כי (ו) גורר את (א). ■

Every nonnegative number has a unique nonnegative square root. The next result shows that positive operators enjoy a similar property.

7.39 *each positive operator has only one positive square root*

Every positive operator on V has a unique positive square root.

A positive operator can have infinitely many square roots (although only one of them can be positive). For example, the identity operator on V has infinitely many square roots if $\dim V > 1$.

Proof Suppose $T \in \mathcal{L}(V)$ is positive. Suppose $v \in V$ is an eigenvector of T. Hence there exists a real number $\lambda \geq 0$ such that $Tv = \lambda v$.

Let R be a positive square root of T. We will prove that $Rv = \sqrt{\lambda}v$. This will imply that the behavior of R on the eigenvectors of T is uniquely determined. Because there is a basis of V consisting of eigenvectors of T (by the spectral theorem), this will imply that R is uniquely determined.

To prove that $Rv = \sqrt{\lambda}v$, note that the spectral theorem asserts that there is an orthonormal basis $e_1, \dots, e_n$ of V consisting of eigenvectors of R. Because R is a positive operator, all its eigenvalues are nonnegative. Thus there exist nonnegative numbers $\lambda_1, \dots, \lambda_n$ such that $Re_k = \sqrt{\lambda_k}e_k$ for each $k = 1, \dots, n$.

Because $e_1, \dots, e_n$ is a basis of V, we can write

$$v = a_1 e_1 + \cdots + a_n e_n$$

for some numbers $a_1, \dots, a_n \in \mathbf{F}$. Thus

$$Rv = a_1\sqrt{\lambda_1}e_1 + \cdots + a_n\sqrt{\lambda_n}e_n.$$

Hence

$$\lambda v = Tv = R^2 v = a_1\lambda_1 e_1 + \cdots + a_n\lambda_n e_n.$$

The equation above implies that

$$a_1\lambda e_1 + \cdots + a_n\lambda e_n = a_1\lambda_1 e_1 + \cdots + a_n\lambda_n e_n.$$

Thus $a_k(\lambda - \lambda_k) = 0$ for each $k = 1, \dots, n$. Hence

$$v = \sum_{\{k:\lambda_k = \lambda\}} a_k e_k.$$

Thus

$$Rv = \sum_{\{k:\lambda_k = \lambda\}} a_k\sqrt{\lambda}e_k = \sqrt{\lambda}v,$$

as desired. ∎

The notation defined below makes sense thanks to the result above.

7.40 notation: $\sqrt{T}$

For T a positive operator, $\sqrt{T}$ denotes the unique positive square root of T.

לכל מספר אי־שלילי יש שורש ריבועי אי־שלילי יחיד. המסקנה הבאה מראה כי לאופרטורים חיוביים יש תכונה דומה.

7.39 לכל אופרטור חיובי יש שורש ריבועי חיובי אחד בלבד

לכל אופרטור חיובי מעל V יש שורש ריבועי חיובי יחיד.

לאופרטור חיובי יכולים להיות אינסוף שורשים ריבועיים (אף שרק אחד מהם יכול להיות חיובי). לדוגמה, לאופרטור הזהות מעל V יש אינסוף שורשים ריבועיים אם $\dim V > 1$.

הוכחה נניח ש־ $T \in \mathcal{L}(V)$ חיובי. נניח ש־ $v \in V$ וקטור עצמי של T. לכן קיים מספר ממשי $\lambda \geq 0$ כך ש־ $Tv = \lambda v$.

יהי R שורש ריבועי חיובי של T. אנו נוכיח כי $Rv = \sqrt{\lambda} v$. מכאן נסיק שההתנהגות של R מעל הווקטורים העצמיים של T נקבעת באופן יחיד. מאחר שקיים בסיס של V המורכב מהווקטורים העצמיים של T (לפי משפט הפירוק הספקטרלי), מזה נגרר ש־ R נקבע באופן יחיד.

כדי להוכיח כי $Rv = \sqrt{\lambda} v$, נשים לב שמשפט הפירוק הספקטרלי קובע כי קיים בסיס אורתונורמלי $e_1, \dots, e_n$ של V המורכב מווקטורים עצמיים של R. מאחר ש־ R אופרטור חיובי, כל הערכים העצמיים אי־שליליים. לפיכך קיימים מספרים אי־שליליים $\lambda_1, \dots, \lambda_n$ כך ש־ $Re_k = \sqrt{\lambda_k} e_k$ לכל $k = 1, \dots, n$.

מאחר ש־ $e_1, \dots, e_n$ בסיס של V, נוכל לכתוב

$$v = a_1 e_1 + \cdots + a_n e_n$$

עבור $a_1, \dots, a_n \in \mathbf{F}$ כלשהם. לפיכך

$$Rv = a_1 \sqrt{\lambda_1} e_1 + \cdots + a_n \sqrt{\lambda_n} e_n.$$

לכן

$$\lambda v = Tv = R^2 v = a_1 \lambda_1 e_1 + \cdots + a_n \lambda_n e_n.$$

המשוואה לעיל גוררת ש־

$$a_1 \lambda e_1 + \cdots + a_n \lambda e_n = a_1 \lambda_1 e_1 + \cdots + a_n \lambda_n e_n.$$

לפיכך $a_k(\lambda - \lambda_k) = 0$ לכל $k = 1, \dots, n$. לכן

$$v = \sum_{\{k : \lambda_k = \lambda\}} a_k e_k.$$

לפיכך

$$Rv = \sum_{\{k : \lambda_k = \lambda\}} a_k \sqrt{\lambda} e_k = \sqrt{\lambda} v,$$

כפי שרצינו. ∎

הסימון המוגדר להלן הגיוני ובעל משמעות לאור המסקנה לעיל.

7.40 סימון: $\sqrt{T}$

עבור T אופרטור חיובי, $\sqrt{T}$ מסמן את השורש הריבועי החיובי היחיד של T.

7.41 example: *square root of positive operators*

Define operators S, T on $\mathbf{R}^2$ (with the usual Euclidean inner product) by

$$S(x, y) = (x, 2y) \quad \text{and} \quad T(x, y) = (x + y, x + y).$$

Then with respect to the standard basis of $\mathbf{R}^2$ we have

7.42
$$\mathcal{M}(S) = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \quad \text{and} \quad \mathcal{M}(T) = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

Each of these matrices equals its transpose; thus S and T are self-adjoint.

If $(x, y) \in \mathbf{R}^2$, then

$$\langle S(x, y), (x, y) \rangle = x^2 + 2y^2 \geq 0$$

and

$$\langle T(x, y), (x, y) \rangle = x^2 + 2xy + y^2 = (x + y)^2 \geq 0.$$

Thus S and T are positive operators.

The standard basis of $\mathbf{R}^2$ is an orthonormal basis consisting of eigenvectors of S. Note that

$$\left(\tfrac{1}{\sqrt{2}}, \tfrac{1}{\sqrt{2}}\right), \left(\tfrac{1}{\sqrt{2}}, -\tfrac{1}{\sqrt{2}}\right)$$

is an orthonormal basis of eigenvectors of T, with eigenvalue 2 for the first eigenvector and eigenvalue 0 for the second eigenvector. Thus $\sqrt{T}$ has the same eigenvectors, with eigenvalues $\sqrt{2}$ and 0.

You can verify that

$$\mathcal{M}\left(\sqrt{S}\right) = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{2} \end{pmatrix} \quad \text{and} \quad \mathcal{M}\left(\sqrt{T}\right) = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$$

with respect to the standard basis by showing that the squares of the matrices above are the matrices in 7.42 and that each matrix above is the matrix of a positive operator.

The statement of the next result does not involve a square root, but the clean proof makes nice use of the square root of a positive operator.

7.43 T *positive and* $\langle Tv, v \rangle = 0 \implies Tv = 0$

Suppose T is a positive operator on V and $v \in V$ is such that $\langle Tv, v \rangle = 0$. Then $Tv = 0$.

Proof We have

$$0 = \langle Tv, v \rangle = \left\langle \sqrt{T}\sqrt{T}v, v \right\rangle = \left\langle \sqrt{T}v, \sqrt{T}v \right\rangle = \left\| \sqrt{T}v \right\|^2.$$

Hence $\sqrt{T}v = 0$. Thus $Tv = \sqrt{T}\left(\sqrt{T}v\right) = 0$, as desired. ∎

7.41 דוגמה: שורש ריבועי של אופרטורים חיוביים

נגדיר אופרטורים S, T מעל $\mathbf{R}^2$ (עם המכפלה הפנימית האוקלידית הרגילה) על ידי

$$S(x,y) = (x, 2y) \quad \text{וגם} \quad T(x,y) = (x+y, x+y).$$

אזי ביחס לבסיס הסטנדרטי של $\mathbf{R}^2$ נקבל

$$\mathcal{M}(S) = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}, \mathcal{M}(T) = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}. \tag{7.42}$$

כל אחת מהמטריצות הללו שווה למשוחלפת שלה; לפיכך S ו־ T צמודים לעצמם.

אם $(x,y) \in \mathbf{R}^2$, אזי

$$\langle S(x,y), (x,y) \rangle = x^2 + 2y^2 \geq 0$$

וגם

$$\langle T(x,y), (x,y) \rangle = x^2 + 2xy + y^2 = (x+y)^2 \geq 0.$$

לפיכך S ו־ T אופרטורים חיוביים.

הבסיס הסטנדרטי של $\mathbf{R}^2$ הוא בסיס אורתונורמלי המורכב מווקטורים עצמיים של S. נשים לב כי

$$\left(\tfrac{1}{\sqrt{2}}, \tfrac{1}{\sqrt{2}}\right), \left(\tfrac{1}{\sqrt{2}}, -\tfrac{1}{\sqrt{2}}\right)$$

הוא בסיס אורתונורמלי של וקטורים עצמיים של T, עם הערך העצמי 2 עבור הווקטור העצמי הראשון והערך העצמי 0 עבור הווקטור העצמי השני. לפיכך ל־ $\sqrt{T}$ אותם וקטורים עצמיים, עם ערכים עצמיים $\sqrt{2}$ ו־ 0.

באפשרותכם לוודא כי

$$\mathcal{M}\left(\sqrt{S}\right) = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{2} \end{pmatrix} \quad \text{וגם} \quad \mathcal{M}\left(\sqrt{T}\right) = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$$

ביחס לבסיס הסטנדרטי על ידי כך שתראו כי ריבועי המטריצות לעיל הם המטריצות ב־ 7.42 ובנוסף כל מטריצה לעיל היא מטריצה של אופרטור חיובי.

הניסוח של המסקנה הבאה אינו עוסק בשורש ריבועי, אך ההוכחה הנקייה עושה שימוש יפה בשורש ריבועי של אופרטור חיובי.

7.43 T חיובי ו־ $\langle Tv, v \rangle = 0 \implies Tv = 0$

נניח ש־ T אופרטור חיובי מעל V ו־ $v \in V$ כך ש־ $\langle Tv, v \rangle = 0$. אזי $Tv = 0$.

הוכחה מתקיים

$$0 = \langle Tv, v \rangle = \left\langle \sqrt{T}\sqrt{T}v, v \right\rangle = \left\langle \sqrt{T}v, \sqrt{T}v \right\rangle = \left\| \sqrt{T}v \right\|^2.$$

לכן $\sqrt{T}v = 0$. לפיכך $Tv = \sqrt{T}\left(\sqrt{T}v\right) = 0$, כפי שרצינו. ■

Exercises 7C

1 Suppose $T \in \mathcal{L}(V)$. Prove that if both T and $-T$ are positive operators, then $T = 0$.

2 Suppose $T \in \mathcal{L}(\mathbf{F}^4)$ is the operator whose matrix (with respect to the standard basis) is

$$\begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{pmatrix}.$$

Show that T is an invertible positive operator.

3 Suppose n is a positive integer and $T \in \mathcal{L}(\mathbf{F}^n)$ is the operator whose matrix (with respect to the standard basis) consists of all 1's. Show that T is a positive operator.

4 Suppose n is an integer with $n > 1$. Show that there exists an n-by-n matrix A such that all of the entries of A are positive numbers and $A = A^*$, but the operator on $\mathbf{F}^n$ whose matrix (with respect to the standard basis) equals A is not a positive operator.

5 Suppose $T \in \mathcal{L}(V)$ is self-adjoint. Prove that T is a positive operator if and only if for every orthonormal basis $e_1, \dots, e_n$ of V, all entries on the diagonal of $\mathcal{M}(T, (e_1, \dots, e_n))$ are nonnegative numbers.

6 Prove that the sum of two positive operators on V is a positive operator.

7 Suppose $S \in \mathcal{L}(V)$ is an invertible positive operator and $T \in \mathcal{L}(V)$ is a positive operator. Prove that $S + T$ is invertible.

8 Suppose $T \in \mathcal{L}(V)$. Prove that T is a positive operator if and only if the pseudoinverse $T^\dagger$ is a positive operator.

9 Suppose $T \in \mathcal{L}(V)$ is a positive operator and $S \in \mathcal{L}(W, V)$. Prove that S^*TS is a positive operator on W.

10 Suppose T is a positive operator on V. Suppose $v, w \in V$ are such that

$$Tv = w \quad \text{and} \quad Tw = v.$$

Prove that $v = w$.

11 Suppose T is a positive operator on V and U is a subspace of V invariant under T. Prove that $T|_U \in \mathcal{L}(U)$ is a positive operator on U.

12 Suppose $T \in \mathcal{L}(V)$ is a positive operator. Prove that T^k is a positive operator for every positive integer k.

תרגילים 7ג

1 נניח ש־ $T \in \mathcal{L}(V)$. הוכיחו כי אם T ו־ $-T$ שניהם אופרטורים חיוביים, אזי $T = 0$.

2 נניח ש־ $T \in \mathcal{L}(\mathbf{F}^4)$ הוא האופרטור אשר המטריצה שלו (ביחס לבסיס הסטנדרטי) היא

$$\begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{pmatrix}.$$

הראו כי T אופרטור חיובי הפיך.

3 נניח ש־ n מספר שלם חיובי ו־ $T \in \mathcal{L}(\mathbf{F}^n)$ הוא האופרטור שהמטריצה שלו (ביחס לבסיס הסטנדרטי) מורכבת כולה מאיברים עם המספר 1. הראו כי T אופרטור חיובי.

4 נניח ש־ n מספר שלם עם $n > 1$. הראו כי קיימת A מטריצה $n \times n$ כך שכל איברי A מספרים חיוביים ו־ $A = A^*$, אך האופרטור מעל $\mathbf{F}^n$ אשר המטריצה שלו (ביחס לבסיס הסטנדרטי) שווה ל־ A אינו אופרטור חיובי.

5 נניח ש־ $T \in \mathcal{L}(V)$ צמוד לעצמו. הוכיחו כי T אופרטור חיובי אם־ורק־אם לכל בסיס אורתונורמלי $e_1, \dots, e_n$ של V, כל האיברים על האלכסון של המטריצה $\mathcal{M}(T, (e_1, \dots, e_n))$ הם מספרים אי־שליליים.

6 הוכיחו כי הסכום של שני אופרטורים חיוביים מעל V הוא אופרטור חיובי.

7 נניח ש־ $S \in \mathcal{L}(V)$ אופרטור חיובי הפיך ו־ $T \in \mathcal{L}(V)$ אופרטור חיובי. הוכיחו כי $S + T$ אופרטור הפיך.

8 נניח ש־ $T \in \mathcal{L}(V)$. הוכיחו כי T אופרטור חיובי אם־ורק־אם הפסידו־הופכי $T^\dagger$ הוא אופרטור חיובי.

9 נניח ש־ $T \in \mathcal{L}(V)$ אופרטור חיובי ו־ $S \in \mathcal{L}(W, V)$. הוכיחו כי S^*TS אופרטור חיובי מעל W.

10 נניח ש־ T אופרטור חיובי מעל V. נניח ש־ $v, w \in V$ כך ש־

$$Tv = w \quad \text{וגם} \quad Tw = v.$$

הוכיחו כי $v = w$.

11 נניח ש־ T אופרטור חיובי מעל V ו־ U תת־מרחב של V שמור תחת T. הוכיחו כי $T|_U \in \mathcal{L}(U)$ אופרטור חיובי מעל U.

12 נניח ש־ $T \in \mathcal{L}(V)$ אופרטור חיובי. הוכיחו כי T^k אופרטור חיובי לכל k מספר שלם חיובי.

13 Suppose $T \in \mathcal{L}(V)$ is self-adjoint and $\alpha \in \mathbf{R}$.

(a) Prove that $T - \alpha I$ is a positive operator if and only if α is less than or equal to every eigenvalue of T.

(b) Prove that $\alpha I - T$ is a positive operator if and only if α is greater than or equal to every eigenvalue of T.

14 Suppose T is a positive operator on V and $v_1, \dots, v_m \in V$. Prove that

$$\sum_{j=1}^{m} \sum_{k=1}^{m} \langle T v_k, v_j \rangle \geq 0.$$

15 Suppose $T \in \mathcal{L}(V)$ is self-adjoint. Prove that there exist positive operators $A, B \in \mathcal{L}(V)$ such that

$$T = A - B \quad \text{and} \quad \sqrt{T^*T} = A + B \quad \text{and} \quad AB = BA = 0.$$

16 Suppose T is a positive operator on V. Prove that

$$\operatorname{null} \sqrt{T} = \operatorname{null} T \quad \text{and} \quad \operatorname{range} \sqrt{T} = \operatorname{range} T.$$

17 Suppose that $T \in \mathcal{L}(V)$ is a positive operator. Prove that there exists a polynomial p with real coefficients such that $\sqrt{T} = p(T)$.

18 Suppose S and T are positive operators on V. Prove that ST is a positive operator if and only if S and T commute.

19 Show that the identity operator on $\mathbf{F}^2$ has infinitely many self-adjoint square roots.

20 Suppose $T \in \mathcal{L}(V)$ and $e_1, \dots, e_n$ is an orthonormal basis of V. Prove that T is a positive operator if and only if there exist $v_1, \dots, v_n \in V$ such that

$$\langle T e_k, e_j \rangle = \langle v_k, v_j \rangle$$

for all $j, k = 1, \dots, n$.

The numbers $\{\langle T e_k, e_j \rangle\}_{j,k=1,\dots,n}$ are the entries in the matrix of T with respect to the orthonormal basis $e_1, \dots, e_n$.

21 Suppose n is a positive integer. The n-by-n *Hilbert matrix* is the n-by-n matrix whose entry in row j, column k is $\frac{1}{j+k-1}$. Suppose $T \in \mathcal{L}(V)$ is an operator whose matrix with respect to some orthonormal basis of V is the n-by-n Hilbert matrix. Prove that T is a positive invertible operator.

Example: The 4-by-4 Hilbert matrix is

$$\begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \frac{1}{6} \\ \frac{1}{4} & \frac{1}{5} & \frac{1}{6} & \frac{1}{7} \end{pmatrix}.$$

13 נניח ש־ $T \in \mathcal{L}(V)$ צמוד לעצמו ו־ $\alpha \in \mathbf{R}$.

א. הוכיחו כי $T - \alpha I$ אופרטור חיובי אם־ורק־אם α קטן או שווה לכל ערך עצמי של T.

ב. הוכיחו כי $\alpha I - T$ אופרטור חיובי אם־ורק־אם α גדול או שווה לכל ערך עצמי של T.

14 נניח ש־ T אופרטור חיובי מעל V ו־ $v_1, \dots, v_m \in V$. הוכיחו כי

$$\sum_{j=1}^{m} \sum_{k=1}^{m} \langle T v_k, v_j \rangle \geq 0.$$

15 נניח ש־ $T \in \mathcal{L}(V)$ צמוד לעצמו. הוכיחו כי קיימים אופרטורים חיוביים $A, B \in \mathcal{L}(V)$ כך ש־

$$T = A - B \quad \text{וגם} \quad \sqrt{T^*T} = A + B \quad \text{וגם} \quad AB = BA = 0.$$

16 נניח ש־ T אופרטור חיובי מעל V. הוכיחו כי

$$\operatorname{null} \sqrt{T} = \operatorname{null} T \quad \text{וגם} \quad \operatorname{range} \sqrt{T} = \operatorname{range} T.$$

17 נניח ש־ $T \in \mathcal{L}(V)$ אופרטור חיובי. הוכיחו כי קיים פולינום p עם מקדמים ממשיים כך ש־ $\sqrt{T} = p(T)$.

18 נניח ש־ S ו־ T אופרטורים חיוביים מעל V. הוכיחו כי ST אופרטור חיובי אם־ורק־אם S ו־ T מתחלפים.

19 הראו כי לאופרטור הזהות מעל $\mathbf{F}^2$ יש אינסוף שורשים ריבועיים הצמודים לעצמם.

20 נניח ש־ $T \in \mathcal{L}(V)$ ו־ $e_1, \dots, e_n$ בסיס אורתונורמלי של V. הוכיחו כי T אופרטור חיובי אם־ורק־אם קיימים $v_1, \dots, v_n \in V$ כך ש־

$$\langle T e_k, e_j \rangle = \langle v_k, v_j \rangle$$

לכל $j, k = 1, \dots, n$.

המספרים $\{\langle T e_k, e_j \rangle\}_{j,k=1,\dots,n}$ הם איברי המטריצה של T ביחס אל הבסיס האורתונורמלי $e_1, \dots, e_n$.

21 נניח ש־ n מספר שלם חיובי. מטריצת הילברט היא מטריצה $n \times n$ שהאיבר בשורה j, עמודה k הוא $\frac{1}{j+k-1}$. נניח ש־ $T \in \mathcal{L}(V)$ הוא אופרטור שהמטריצה שלו ביחס לבסיס אורתונורמלי כלשהו של V היא מטריצת הילברט $n \times n$. הוכיחו כי T אופרטור חיובי הפיך.

דוגמה: מטריצת הילברט 4×4 היא

$$\begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \frac{1}{6} \\ \frac{1}{4} & \frac{1}{5} & \frac{1}{6} & \frac{1}{7} \end{pmatrix}.$$

אלגברה ליניארית בדרך הישר, מהדורה רביעית, מחבר: שלדון אקסלר
Linear Algebra Done Right, fourth edition, by Sheldon Axler

22 Suppose $T \in \mathcal{L}(V)$ is a positive operator and $u \in V$ is such that $\|u\| = 1$ and $\|Tu\| \geq \|Tv\|$ for all $v \in V$ with $\|v\| = 1$. Show that u is an eigenvector of T corresponding to the largest eigenvalue of T.

23 For $T \in \mathcal{L}(V)$ and $u, v \in V$, define $\langle u, v\rangle_T$ by $\langle u, v\rangle_T = \langle Tu, v\rangle$.

(a) Suppose $T \in \mathcal{L}(V)$. Prove that $\langle \cdot, \cdot \rangle_T$ is an inner product on V if and only if T is an invertible positive operator (with respect to the original inner product $\langle \cdot, \cdot \rangle$).

(b) Prove that every inner product on V is of the form $\langle \cdot, \cdot \rangle_T$ for some positive invertible operator $T \in \mathcal{L}(V)$.

24 Suppose S and T are positive operators on V. Prove that

$$\operatorname{null}(S + T) = \operatorname{null} S \cap \operatorname{null} T.$$

25 Let T be the second derivative operator in Exercise 31(b) in Section 7A. Show that $-T$ is a positive operator.

22 נניח ש־ $T \in \mathcal{L}(V)$ אופרטור חיובי ו־ $u \in V$ כך ש־ $\|u\| = 1$ ו־ $\|Tu\| \geq \|Tv\|$ לכל $v \in V$ עם $\|v\| = 1$. הראו כי u וקטור עצמי של T השייך לערך העצמי הגדול ביותר של T.

23 עבור $T \in \mathcal{L}(V)$ ו־ $u, v \in V$, נגדיר $\langle u, v\rangle_T$ על ידי $\langle u, v\rangle_T = \langle Tu, v\rangle$.

א. נניח ש־ $T \in \mathcal{L}(V)$. הוכיחו כי $\langle \cdot, \cdot\rangle_T$ הוא מכפלה פנימית מעל V אם־ורק־אם T הוא אופרטור חיובי הפיך (ביחס אל המכפלה הפנימית המקורית $\langle \cdot, \cdot\rangle$).

ב. הוכיחו כי כל מכפלה פנימית מעל V היא מהצורה $\langle \cdot, \cdot\rangle_T$ עבור $T \in \mathcal{L}(V)$ אופרטור חיובי הפיך כלשהו.

24 נניח ש־ S ו־ T אופרטורים חיוביים מעל V. הוכיחו כי

$$\operatorname{null}(S + T) = \operatorname{null} S \cap \operatorname{null} T.$$

25 יהי T אופרטור הנגזרת השנייה בתרגיל 31(ב) ביחידה 1ז. הראו כי $-T$ אופרטור חיובי.

אלגברה ליניארית בדרך הישר, מהדורה רביעית, מחבר: שלדון אקסלר
Linear Algebra Done Right, fourth edition, by Sheldon Axler

7D Isometries, Unitary Operators, and Matrix Factorization

Isometries

Linear maps that preserve norms are sufficiently important to deserve a name.

7.44 definition: *isometry*

A linear map $S \in \mathcal{L}(V, W)$ is called an *isometry* if

$$\|Sv\| = \|v\|$$

for every $v \in V$. In other words, a linear map is an isometry if it preserves norms.

If $S \in \mathcal{L}(V, W)$ is an isometry and $v \in V$ is such that $Sv = 0$, then

$$\|v\| = \|Sv\| = \|0\| = 0,$$

which implies that $v = 0$. Thus every isometry is injective.

*The Greek word **isos** means equal; the Greek word **metron** means measure. Thus **isometry** literally means equal measure.*

7.45 example: *orthonormal basis maps to orthonormal list $\implies$ isometry*

Suppose $e_1, \dots, e_n$ is an orthonormal basis of V and $g_1, \dots, g_n$ is an orthonormal list in W. Let $S \in \mathcal{L}(V, W)$ be the linear map such that $Se_k = g_k$ for each $k = 1, \dots, n$. To show that S is an isometry, suppose $v \in V$. Then

7.46 $$v = \langle v, e_1\rangle e_1 + \cdots + \langle v, e_n\rangle e_n$$

and

7.47 $$\|v\|^2 = |\langle v, e_1\rangle|^2 + \cdots + |\langle v, e_n\rangle|^2,$$

where we have used 6.30(b). Applying S to both sides of 7.46 gives

$$Sv = \langle v, e_1\rangle Se_1 + \cdots + \langle v, e_n\rangle Se_n = \langle v, e_1\rangle g_1 + \cdots + \langle v, e_n\rangle g_n.$$

Thus

7.48 $$\|Sv\|^2 = |\langle v, e_1\rangle|^2 + \cdots + |\langle v, e_n\rangle|^2.$$

Comparing 7.47 and 7.48 shows that $\|v\| = \|Sv\|$. Thus S is an isometry.

The next result gives conditions equivalent to being an isometry. The equivalence of (a) and (c) shows that a linear map is an isometry if and only if it preserves inner products. The equivalence of (a) and (d) shows that a linear map is an isometry if and only if it maps some orthonormal basis to an orthonormal list. Thus the isometries given by Example 7.45 include all isometries. Furthermore, a linear map is an isometry if and only if it maps every orthonormal basis to an orthonormal list [because whether or not (a) holds does not depend on the basis $e_1, \dots, e_n$].

7ג איזומטריות, אופרטורים אוניטריים, ופירוק מטריצה לגורמים

איזומטריות

העתקות ליניאריות המשמרות את הנורמות חשובות עד כדי הענקת שם ייעודי.

7.44 הגדרה: איזומטריה

העתקה ליניארית $S \in \mathcal{L}(V,W)$ נקראת **איזומטריה** אם

$$\|Sv\| = \|v\|$$

לכל $v \in V$. במילים אחרות, העתקה ליניארית היא איזומטריה אם היא משמרת את הנורמות.

ביווניית עתיקה המילה **איזוס** משמעותה שָׁוֶה; המילה **מטרון** משמעותה מִדָּה. לפיכך **איזומטריה** משמעותה שָׁוֶת מִדָּה.

אם $S \in \mathcal{L}(V,W)$ איזומטריה ו־ $v \in V$ כך ש־ $Sv = 0$, אזי

$$\|v\| = \|Sv\| = \|0\| = 0,$$

מכאן נובע ש־ $v = 0$. לפיכך כל איזומטריה היא חד־חד־ערכית.

7.45 דוגמה: בסיס אורתונורמלי המועתק לרשימה אורתונורמלית ⇐ איזומטריה

נניח ש־ $e_1, \dots, e_n$ בסיס אורתונורמלי של V ו־ $g_1, \dots, g_n$ רשימה אורתונורמלית ב־ W. תהי $S \in \mathcal{L}(V,W)$ ההעתקה ליניארית כך ש־ $Se_k = g_k$ לכל $k = 1, \dots, n$. כדי להראות כי S איזומטריה, נניח ש־ $v \in V$. אזי

$$v = \langle v, e_1\rangle e_1 + \cdots + \langle v, e_n\rangle e_n \tag{7.46}$$

וגם

$$\|v\|^2 = |\langle v, e_1\rangle|^2 + \cdots + |\langle v, e_n\rangle|^2, \tag{7.47}$$

כאשר השתמשנו ב־ 6.30(ב). הפעלת S על שני האגפים של 7.46 נותנת

$$Sv = \langle v, e_1\rangle Se_1 + \cdots + \langle v, e_n\rangle Se_n = \langle v, e_1\rangle g_1 + \cdots + \langle v, e_n\rangle g_n.$$

לפיכך

$$\|Sv\|^2 = |\langle v, e_1\rangle|^2 + \cdots + |\langle v, e_n\rangle|^2. \tag{7.48}$$

השוואת 7.47 ו־ 7.48 מראה לנו ש־ $\|Sv\| = \|v\|$. לפיכך S איזומטריה.

המסקנה הבאה מספקת תנאים שקולים לאיזומטריה. השקילות בין (א) ו־ (ג) מורה לנו כי העתקה ליניארית היא איזומטריה אם־ורק־אם היא משמרת את המכפלות הפנימיות. השקילות בין (א) ו־ (ד) מורה לנו כי העתקה ליניארית היא איזומטריה אם־ורק־אם היא מעתיקה בסיס אורתונורמלי כלשהו לרשימה אורתונורמלית. לפיכך האיזומטריות הנתונות בדוגמה 7.45 כוללות את כל האיזומטריות. יתר על כן, העתקה ליניארית היא איזומטריה אם־ורק־אם היא מעתיקה כל בסיס אורתונורמלי לרשימה אורתונורמלית [מאחר שקיומו של (א) אינו תלוי בבסיס $e_1, \dots, e_n$].

The equivalence of (a) and (e) in the next result shows that a linear map is an isometry if and only if the columns of its matrix (with respect to any orthonormal bases) form an orthonormal list. Here we are identifying the columns of an m-by-n matrix with elements of $\mathbf{F}^m$ and then using the Euclidean inner product on $\mathbf{F}^m$.

7.49 *characterizations of isometries*

Suppose $S \in \mathcal{L}(V, W)$. Suppose $e_1, \dots, e_n$ is an orthonormal basis of V and $f_1, \dots, f_m$ is an orthonormal basis of W. Then the following are equivalent.

(a) S is an isometry.

(b) $S^*S = I$.

(c) $\langle Su, Sv\rangle = \langle u, v\rangle$ for all $u, v \in V$.

(d) $Se_1, \dots, Se_n$ is an orthonormal list in W.

(e) The columns of $\mathcal{M}(S, (e_1, \dots, e_n), (f_1, \dots, f_m))$ form an orthonormal list in $\mathbf{F}^m$ with respect to the Euclidean inner product.

Proof First suppose (a) holds, so S is an isometry. If $v \in V$ then

$$\big\langle (I - S^*S)v, v\big\rangle = \langle v, v\rangle - \big\langle S^*Sv, v\big\rangle = \|v\|^2 - \langle Sv, Sv\rangle = \|v\|^2 - \|Sv\|^2 = 0.$$

Hence the self-adjoint operator $I - S^*S$ equals 0 (by 7.16). Thus $S^*S = I$, proving that (a) implies (b).

Now suppose (b) holds, so $S^*S = I$. If $u, v \in V$ then

$$\langle Su, Sv\rangle = \big\langle S^*Su, v\big\rangle = \langle Iu, v\rangle = \langle u, v\rangle,$$

proving that (b) implies (c).

Now suppose that (c) holds, so $\langle Su, Sv\rangle = \langle u, v\rangle$ for all $u, v \in V$. Thus if $j, k \in \{1, \dots, n\}$, then

$$\langle Se_j, Se_k\rangle = \langle e_j, e_k\rangle.$$

Hence $Se_1, \dots, Se_n$ is an orthonormal list in W, proving that (c) implies (d).

Now suppose that (d) holds, so $Se_1, \dots, Se_n$ is an orthonormal list in W. Let $A = \mathcal{M}(S, (e_1, \dots, e_n), (f_1, \dots, f_m))$. If $k, r \in \{1, \dots, n\}$, then

$$\text{7.50} \qquad \sum_{j=1}^{m} A_{j,k}\overline{A_{j,r}} = \Big\langle \sum_{j=1}^{m} A_{j,k} f_j, \sum_{j=1}^{m} A_{j,r} f_j \Big\rangle = \langle Se_k, Se_r\rangle = \begin{cases} 1 & \text{if } k = r, \\ 0 & \text{if } k \neq r. \end{cases}$$

The left side of 7.50 is the inner product in $\mathbf{F}^m$ of columns k and r of A. Thus the columns of A form an orthonormal list in $\mathbf{F}^m$, proving that (d) implies (e).

Now suppose (e) holds, so the columns of the matrix A defined in the paragraph above form an orthonormal list in $\mathbf{F}^m$. Then 7.50 shows that $Se_1, \dots, Se_n$ is an orthonormal list in W. Thus Example 7.45, with $Se_1, \dots, Se_n$ playing the role of $g_1, \dots, g_n$, shows that S is an isometry, proving that (e) implies (a). ∎

See Exercises 1 and 11 for additional conditions that are equivalent to being an isometry.

השקילות בין (א) ו־ (ה) במסקנה הבאה מורה לנו כי העתקה ליניארית היא איזומטריה אם־ורק־אם עמודות המטריצה שלה (ביחס לכל בסיס אורתונורמלי) יוצרות רשימה אורתונורמלית. כאן אנו מזהים את עמודות המטריצה $m \times n$ עם איברים של $\mathbf{F}^m$ ואז משתמשים במכפלה הפנימית האוקלידית מעל $\mathbf{F}^m$.

7.49 אפיונים של איזומטריות

נניח ש־ $S \in \mathcal{L}(V, W)$. נניח ש־ $e_1, \dots, e_n$ בסיס אורתונורמלי של V ו־ $f_1, \dots, f_m$ בסיס אורתונורמלי של W. אזי התנאים הבאים שקולים.

א. S איזומטריה.

ב. $S^*S = I$.

ג. $\langle Su, Sv\rangle = \langle u, v\rangle$ לכל $u, v \in V$.

ד. $Se_1, \dots, Se_n$ רשימה אורתונורמלית ב־ W.

ה. העמודות של $\mathcal{M}(S, (e_1, \dots, e_n), (f_1, \dots, f_m))$ יוצרות רשימה אורתונורמלית ב־ $\mathbf{F}^m$ ביחס למכפלה הפנימית האוקלידית.

הוכחה תחילה נניח כי (א) מתקיים, אזי S איזומטריה. אם $v \in V$ אזי

$$\langle (I - S^*S)v, v\rangle = \langle v, v\rangle - \langle S^*Sv, v\rangle = \|v\|^2 - \langle Sv, Sv\rangle = \|v\|^2 - \|Sv\|^2 = 0.$$

לכן האופרטור הצמוד לעצמו $I - S^*S$ שווה 0 (לפי 7.16). לפיכך $S^*S = I$, ובכך הוכחנו כי (א) גורר את (ב).

כעת נניח כי (ב) מתקיים, אזי $S^*S = I$. אם $u, v \in V$ אזי

$$\langle Su, Sv\rangle = \langle S^*Su, v\rangle = \langle Iu, v\rangle = \langle u, v\rangle,$$

בכך הוכחנו כי (ב) גורר את (ג).

כעת נניח כי (ג) מתקיים, אזי $\langle Su, Sv\rangle = \langle u, v\rangle$ לכל $u, v \in V$. לפיכך אם $j, k \in \{1, \dots, n\}$, אזי

$$\langle Se_j, Se_k\rangle = \langle e_j, e_k\rangle.$$

לכן $Se_1, \dots, Se_n$ רשימה אורתונורמלית ב־ W, ובכך הוכחנו כי (ג) גורר את (ד).

כעת נניח ש־ (ד) מתקיים, אזי $Se_1, \dots, Se_n$ רשימה אורתונורמלית ב־ W. תהי $A = \mathcal{M}(S, (e_1, \dots, e_n), (f_1, \dots, f_m))$. אם $k, r \in \{1, \dots, n\}$, אזי

$$\sum_{j=1}^{m} A_{j,k}\overline{A_{j,r}} = \left\langle \sum_{j=1}^{m} A_{j,k} f_j, \sum_{j=1}^{m} A_{j,r} f_j \right\rangle = \langle Se_k, Se_r\rangle = \begin{cases} 1 & \text{if } k = r, \\ 0 & \text{if } k \neq r. \end{cases} \tag{7.50}$$

האגף השמאלי של 7.50 הוא המכפלה הפנימית ב־ $\mathbf{F}^m$ של עמודות k ו־ r של A. לפיכך העמודות של A יוצרות רשימה אורתונורמלית ב־ $\mathbf{F}^m$, ובכך הוכחנו כי (ד) גורר את (ה).

כעת נניח כי (ה) מתקיים, כלומר העמודות של המטריצה A שהוגדרו בפסקה הקודמת יוצרות רשימה אורתונורמלית ב־ $\mathbf{F}^m$. אזי לפי 7.50 הרשימה $Se_1, \dots, Se_n$ היא אורתונורמלית ב־ W. לפיכך דוגמה 7.45, יחד עם $Se_1, \dots, Se_n$ הממלאים את התפקיד של $g_1, \dots, g_n$, מראה ש־ S איזומטריה, ובכך הוכחנו כי (ה) גורר את (א). ■

ראו תרגילים 1 ו־ 11 לתנאים שקולים נוספים לאיזומטריה.

Unitary Operators

In this subsection, we confine our attention to linear maps from a vector space to itself. In other words, we will be working with operators.

7.51 definition: *unitary operator*

An operator $S \in \mathcal{L}(V)$ is called *unitary* if S is an invertible isometry.

Although the words "unitary" and "isometry" mean the same thing for operators on finite-dimensional inner product spaces, remember that a unitary operator maps a vector space to itself, while an isometry maps a vector space to another (possibly different) vector space.

As previously noted, every isometry is injective. Every injective operator on a finite-dimensional vector space is invertible (see 3.65). A standing assumption for this chapter is that V is a finite-dimensional inner product space. Thus we could delete the word "invertible" from the definition above without changing the meaning. The unnecessary word "invertible" has been retained in the definition above for consistency with the definition readers may encounter when learning about inner product spaces that are not necessarily finite-dimensional.

7.52 example: *rotation of* $\mathbf{R}^2$

Suppose $\theta \in \mathbf{R}$ and S is the operator on $\mathbf{F}^2$ whose matrix with respect to the standard basis of $\mathbf{F}^2$ is

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

The two columns of this matrix form an orthonormal list in $\mathbf{F}^2$; hence S is an isometry [by the equivalence of (a) and (e) in 7.49]. Thus S is a unitary operator.

If $\mathbf{F} = \mathbf{R}$, then S is the operator of counterclockwise rotation by θ radians around the origin of $\mathbf{R}^2$. This observation gives us another way to think about why S is an isometry, because each rotation around the origin of $\mathbf{R}^2$ preserves norms.

The next result (7.53) lists several conditions that are equivalent to being a unitary operator. All the conditions equivalent to being an isometry in 7.49 should be added to this list. The extra conditions in 7.53 arise because of limiting the context to linear maps from a vector space to itself. For example, 7.49 shows that a linear map $S \in \mathcal{L}(V, W)$ is an isometry if and only if $S^*S = I$, while 7.53 shows that an operator $S \in \mathcal{L}(V)$ is a unitary operator if and only if $S^*S = SS^* = I$.

Another difference is that 7.49(d) mentions an orthonormal list, while 7.53(d) mentions an orthonormal basis. Also, 7.49(e) mentions the columns of $\mathcal{M}(T)$, while 7.53(e) mentions the rows of $\mathcal{M}(T)$. Furthermore, $\mathcal{M}(T)$ in 7.49(e) is with respect to an orthonormal basis of V and an orthonormal basis of W, while $\mathcal{M}(T)$ in 7.53(e) is with respect to a single basis of V doing double duty.

אופרטורים אוניטריים

בתת־יחידה זו של הפרק, נמקד את תשומת ליבנו בהעתקות ליניאריות ממרחב וקטורי לעצמו. במילים אחרות, אנו נעבוד עם אופרטורים.

7.51 הגדרה: אופרטור אוניטרי

אופרטור $S \in \mathcal{L}(V)$ נקרא אוניטרי אם S הוא איזומטריה הפיכה.

אף שהמילים "אוניטרי" ו־ "איזומטריה" משמעותן זהה עבור אופרטורים מעל מרחבי מכפלה פנימית מממד סופי, זכרו כי אופרטור אוניטרי מעתיק מרחב וקטורי לעצמו, בעוד איזומטריה מעתיקה מרחב וקטורי אל מרחב וקטורי (שייתכן שהוא אחר).

כפי שצוין קודם, כל איזומטריה היא חד־חד־ערכית. כל אופרטור חד־חד־ערכי מעל מרחב וקטורי מממד סופי הפיך (ראו 3.65). הנחת המוצא לפרק זה היא ש־ V מרחב מכפלה פנימית מממד סופי. לפיכך אנו יכולים להשמיט את המילה "הפיך" מההגדרה לעיל בלי שהמשמעות תשתנה. המילה "הפיך" הבלתי נדרשת נשמרה בטקסט ההגדרה לעיל לשמירה על עקביות עם ההגדרה שבה עשויים הקוראים להיתקל בלימוד מרחבי מכפלה פנימית שאינם בהכרח מממד סופי.

7.52 דוגמה: סיבוב של $\mathbf{R}^2$

נניח ש־ $\theta \in \mathbf{R}$ ו־ S הוא האופרטור מעל $\mathbf{F}^2$ אשר המטריצה שלו ביחס לבסיס הסטנדרטי של $\mathbf{F}^2$ היא

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

שתי העמודות של מטריצה זו יוצרות רשימה אורתונורמלית ב־ $\mathbf{F}^2$; לכן S איזומטריה [לפי השקילות בין (א) ו־ (ה) ב־ 7.49]. לפיכך S אופרטור אוניטרי.

אם $\mathbf{F} = \mathbf{R}$, אזי S הוא אופרטור המבצע סיבוב נגד כיוון השעון של θ רדיאנים סביב ראשית הצירים של $\mathbf{R}^2$. התבוננות זו מעניקה לנו דרך נוספת לחשוב על הסיבה מדוע S איזומטריה, מאחר שכל סיבוב סביב ראשית הצירים של $\mathbf{R}^2$ משמר את הנורמות.

במסקנה הבאה (7.53) מפורטים מספר תנאים השקולים לכך שאופרטור יהיה אוניטרי. כל התנאים השקולים כדי להיות איזומטריה המפורטים ב־ 7.49 צריכים להתווסף לרשימה זו. התנאים הנוספים ב־ 7.53 נוצרו לאור העובדה כי הגבלנו את ההקשר להעתקות ממרחב וקטורי לעצמו. לדוגמה, 7.49 מורה לנו כי העתקה ליניארית $S \in \mathcal{L}(V, W)$ איזומטריה אם־ורק־אם $S^*S = I$, בעוד 7.53 מורה לנו כי אופרטור $S \in \mathcal{L}(V)$ אופרטור אוניטרי אם־ורק־אם $S^*S = SS^* = I$.

הבדל נוסף הוא שב־ 7.49(ד) אנו מתייחסים לרשימה אורתונורמלית, בעוד ב־ 7.53(ד) אנו מתייחסים לבסיס אורתונורמלי. בנוסף, ב־ 7.49(ה) אנו מתייחסים לעמודות של $\mathcal{M}(T)$, בעוד ב־ 7.53(ה) אנו מתייחסים לשורות של $\mathcal{M}(T)$. יתר על כן, $\mathcal{M}(T)$ ב־ 7.49(ה) היא ביחס לבסיס אורתונורמלי של V ובסיס אורתונורמלי של W, בעוד $\mathcal{M}(T)$ ב־ 7.53(ה) היא ביחס לבסיס יחיד של V הממלא תפקיד כפול.

7.53 *characterizations of unitary operators*

Suppose $S \in \mathcal{L}(V)$. Suppose $e_1, \dots, e_n$ is an orthonormal basis of V. Then the following are equivalent.

(a) S is a unitary operator.

(b) $S^*S = SS^* = I$.

(c) S is invertible and $S^{-1} = S^*$.

(d) $Se_1, \dots, Se_n$ is an orthonormal basis of V.

(e) The rows of $\mathcal{M}(S, (e_1, \dots, e_n))$ form an orthonormal basis of $\mathbf{F}^n$ with respect to the Euclidean inner product.

(f) S^* is a unitary operator.

Proof First suppose (a) holds, so S is a unitary operator. Hence

$$S^*S = I$$

by the equivalence of (a) and (b) in 7.49. Multiply both sides of this equation by S^{-1} on the right, getting $S^* = S^{-1}$. Thus $SS^* = SS^{-1} = I$, as desired, proving that (a) implies (b).

The definitions of invertible and inverse show that (b) implies (c).

Now suppose (c) holds, so S is invertible and $S^{-1} = S^*$. Thus $S^*S = I$. Hence $Se_1, \dots, Se_n$ is an orthonormal list in V, by the equivalence of (b) and (d) in 7.49. The length of this list equals $\dim V$. Thus $Se_1, \dots, Se_n$ is an orthonormal basis of V, proving that (c) implies (d).

Now suppose (d) holds, so $Se_1, \dots, Se_n$ is an orthonormal basis of V. The equivalence of (a) and (d) in 7.49 shows that S is a unitary operator. Thus

$$(S^*)^*S^* = SS^* = I,$$

where the last equation holds because we already showed that (a) implies (b) in this result. The equation above and the equivalence of (a) and (b) in 7.49 show that S^* is an isometry. Thus the columns of $\mathcal{M}(S^*, (e_1, \dots, e_n))$ form an orthonormal basis of $\mathbf{F}^n$ [by the equivalence of (a) and (e) of 7.49]. The rows of $\mathcal{M}(S, (e_1, \dots, e_n))$ are the complex conjugates of the columns of $\mathcal{M}(S^*, (e_1, \dots, e_n))$. Thus the rows of $\mathcal{M}(S, (e_1, \dots, e_n))$ form an orthonormal basis of $\mathbf{F}^n$, proving that (d) implies (e).

Now suppose (e) holds. Thus the columns of $\mathcal{M}(S^*, (e_1, \dots, e_n))$ form an orthonormal basis of $\mathbf{F}^n$. The equivalence of (a) and (e) in 7.49 shows that S^* is an isometry, proving that (e) implies (f).

Now suppose (f) holds, so S^* is a unitary operator. The chain of implications we have already proved in this result shows that (a) implies (f). Applying this result to S^* shows that $(S^*)^*$ is a unitary operator, proving that (f) implies (a).

We have shown that (a) $\Rightarrow$ (b) $\Rightarrow$ (c) $\Rightarrow$ (d) $\Rightarrow$ (e) $\Rightarrow$ (f) $\Rightarrow$ (a), completing the proof. ∎

7.53 אפיונים של אופרטורים אוניטריים

נניח ש־ $S \in \mathcal{L}(V)$. נניח ש־ $e_1, \dots, e_n$ בסיס אורתונורמלי של V. אזי התנאים הבאים שקולים.

א. S אופרטור אוניטרי.

ב. $S^*S = SS^* = I$.

ג. S הפיך ו־ $S^{-1} = S^*$.

ד. $Se_1, \dots, Se_n$ בסיס אורתונורמלי של V.

ה. השורות של $\mathcal{M}(S, (e_1, \dots, e_n))$ יוצרות בסיס אורתונורמלי של $\mathbf{F}^n$ ביחס אל המכפלה הפנימית האוקלידית.

ו. S^* אופרטור אוניטרי.

הוכחה תחילה נניח כי (א) מתקיים, אזי S אופרטור אוניטרי. לכן

$$S^*S = I$$

לפי השקילות של (א) ו־ (ב) ב־ 7.49. כשנכפיל כל אגף של משוואה זו מימין ב־ S^{-1}, נקבל $S^* = S^{-1}$. לפיכך $SS^* = SS^{-1} = I$, כפי שרצינו, ובכך הוכחנו כי (א) גורר את (ב).

ההגדרות של הפיך ושל הופכי מראות לנו כי (ב) גורר את (ג).

כעת נניח ש־ (ג) מתקיים, אזי S הפיך ו־ $S^{-1} = S^*$. לפיכך $S^*S = I$. לכן $Se_1, \dots, Se_n$ היא רשימה אורתונורמלית ב־ V, לפי השקילות של (ב) ו־ (ד) ב־ 7.49. מספר האיברים של רשימה זו הוא $\dim V$. לפיכך $Se_1, \dots, Se_n$ בסיס אורתונורמלי של V, ובכך הוכחנו כי (ג) גורר את (ד).

כעת נניח כי (ד) מתקיים, אזי $Se_1, \dots, Se_n$ בסיס אורתונורמלי של V. מהשקילות בין (א) ל־ (ד) ב־ 7.49 אנו רואים כי S אופרטור אוניטרי. לפיכך

$$(S^*)^*S^* = SS^* = I,$$

כאשר המשוואה האחרונה מתקיימת מאחר שכבר הראינו כי (א) גורר את (ב) במסקנה זו. המשוואה לעיל והשקילות של (א) ו־ (ב) ב־ 7.49 מראות כי S^* איזומטריה. לפיכך העמודות של $\mathcal{M}(S^*, (e_1, \dots, e_n))$ יוצרות בסיס אורתונורמלי של $\mathbf{F}^n$ [משקילות (א) ו־ (ה) ב־ 7.49]. השורות של $\mathcal{M}(S, (e_1, \dots, e_n))$ הן הצמוד המרוכב של העמודות של $\mathcal{M}(S^*, (e_1, \dots, e_n))$. לפיכך השורות של $\mathcal{M}(S, (e_1, \dots, e_n))$ יוצרות בסיס אורתונורמלי של $\mathbf{F}^n$, ובכך הוכחנו כי (ד) גורר את (ה).

כעת נניח כי (ה) מתקיים. לפיכך העמודות של $\mathcal{M}(S^*, (e_1, \dots, e_n))$ יוצרות בסיס אורתונורמלי של $\mathbf{F}^n$. השקילות של (א) ו־ (ה) ב־ 7.49 מראה ש־ S^* איזומטריה, ובכך הוכחנו כי (ה) גורר את (ו).

כעת נניח ש־ (ו) מתקיים, אזי S^* אופרטור אוניטרי. שרשרת הגרירות שכבר הוכחנו במסקנה זו מראה ש־ (א) גורר את (ו). ביישום מסקנה זו עבור S^* נראה כי $(S^*)^*$ אופרטור אוניטרי, ובכך הוכחנו כי (ו) גורר את (א).

הראינו כיצד (א) $\Rightarrow$ (ב) $\Rightarrow$ (ג) $\Rightarrow$ (ד) $\Rightarrow$ (ה) $\Rightarrow$ (ו) $\Rightarrow$ (א), ובכך השלמנו את ההוכחה. ■

Recall our analogy between $\mathbf{C}$ and $\mathcal{L}(V)$. Under this analogy, a complex number z corresponds to an operator $S \in \mathcal{L}(V)$, and $\overline{z}$ corresponds to S^*. The real numbers ($z = \overline{z}$) correspond to the self-adjoint operators ($S = S^*$), and the nonnegative numbers correspond to the (badly named) positive operators.

Another distinguished subset of $\mathbf{C}$ is the unit circle, which consists of the complex numbers z such that $|z| = 1$. The condition $|z| = 1$ is equivalent to the condition $\overline{z}z = 1$. Under our analogy, this corresponds to the condition $S^*S = I$, which is equivalent to S being a unitary operator. Hence the analogy shows that the unit circle in $\mathbf{C}$ corresponds to the set of unitary operators. In the next two results, this analogy appears in the eigenvalues of unitary operators. Also see Exercise 15 for another example of this analogy.

7.54 *eigenvalues of unitary operators have absolute value* 1

Suppose λ is an eigenvalue of a unitary operator. Then $|\lambda| = 1$.

Proof Suppose $S \in \mathcal{L}(V)$ is a unitary operator and λ is an eigenvalue of S. Let $v \in V$ be such that $v \neq 0$ and $Sv = \lambda v$. Then

$$|\lambda| \, \|v\| = \|\lambda v\| = \|Sv\| = \|v\|.$$

Thus $|\lambda| = 1$, as desired. ∎

The next result characterizes unitary operators on finite-dimensional complex inner product spaces, using the complex spectral theorem as the main tool.

7.55 *description of unitary operators on complex inner product spaces*

Suppose $\mathbf{F} = \mathbf{C}$ and $S \in \mathcal{L}(V)$. Then the following are equivalent.

(a) S is a unitary operator.

(b) There is an orthonormal basis of V consisting of eigenvectors of S whose corresponding eigenvalues all have absolute value 1.

Proof Suppose (a) holds, so S is a unitary operator. The equivalence of (a) and (b) in 7.53 shows that S is normal. Thus the complex spectral theorem (7.31) shows that there is an orthonormal basis $e_1, \ldots, e_n$ of V consisting of eigenvectors of S. Every eigenvalue of S has absolute value 1 (by 7.54), completing the proof that (a) implies (b).

Now suppose (b) holds. Let $e_1, \ldots, e_n$ be an orthonormal basis of V consisting of eigenvectors of S whose corresponding eigenvalues $\lambda_1, \ldots, \lambda_n$ all have absolute value 1. Then $Se_1, \ldots, Se_n$ is also an orthonormal basis of V because

$$\langle Se_j, Se_k \rangle = \langle \lambda_j e_j, \lambda_k e_k \rangle = \lambda_j \overline{\lambda_k} \langle e_j, e_k \rangle = \begin{cases} 0 & \text{if } j \neq k, \\ 1 & \text{if } j = k \end{cases}$$

for all $j, k = 1, \ldots, n$. Thus the equivalence of (a) and (d) in 7.53 shows that S is unitary, proving that (b) implies (a). ∎

נזכיר את האנלוגיה שהצגנו בין $\mathbf{C}$ ל־ $\mathcal{L}(V)$. תחת אנלוגיה זו, מספר מרוכב z מתאים לאופרטור $S \in \mathcal{L}(V)$, ו־ $\bar{z}$ מתאים ל־ S^*. המספרים הממשיים ($z = \bar{z}$) מתאימים לאופרטורים הצמודים לעצמם ($S = S^*$), והמספרים האי־שליליים מתאימים לאופרטורים החיוביים (נושאי השם הבעייתי).

תת־קבוצה נוספת מובחנת של $\mathbf{C}$ היא מעגל היחידה, המורכב מהמספרים המרוכבים z כך ש־ $|z| = 1$. התנאי $|z| = 1$ שקול לתנאי $\bar{z}z = 1$. תחת האנלוגיה שלנו, הדבר מתאים לתנאי ש־ $S^*S = I$, השקול לכך ש־ S אופרטור אוניטרי. לכן האנלוגיה מורה לנו כי מעגל היחידה ב־ $\mathbf{C}$ מתאים לקבוצת האופרטורים האוניטריים. בשתי המסקנות הבאות, אנלוגיה זו מופיעה בערכים העצמיים של האופרטורים האוניטריים. ראו בנוסף תרגיל 15 לדוגמה נוספת של אנלוגיה זו.

7.54 הערך המוחלט של ערכים עצמיים של אופרטורים אוניטריים הוא 1

נניח ש־ λ ערך עצמי של אופרטור אוניטרי. אזי $|\lambda| = 1$.

הוכחה נניח ש־ $S \in \mathcal{L}(V)$ אופרטור אוניטרי ו־ λ ערך עצמי של S. יהי $v \in V$ כך ש־ $v \neq 0$ ו־ $Sv = \lambda v$. אזי

$$|\lambda| \, \|v\| = \|\lambda v\| = \|Sv\| = \|v\|.$$

לפיכך $|\lambda| = 1$, כפי שרצינו. ∎

המסקנה הבאה מאפיינת אופרטורים אוניטריים מעל מרחבי מכפלה פנימית מרוכבים מממד סופי, תוך שימוש במשפט הפירוק הספקטרלי המרוכב ככלי מרכזי.

7.55 תיאור של אופרטורים אוניטריים מעל מרחבי מכפלה פנימית מרוכבים

נניח ש־ $\mathbf{F} = \mathbf{C}$ ו־ $S \in \mathcal{L}(V)$. אזי התנאים הבאים שקולים.

א. S אופרטור אוניטרי.

ב. קיים בסיס אורתונורמלי של V המורכב מווקטורים עצמיים של S שהערך המוחלט של כל הערכים העצמיים השייכים להם הוא 1.

הוכחה נניח ש־ (א) מתקיים, אזי S אופרטור אוניטרי. השקילות של (א) ו־ (ב) ב־ 7.53 מראה ש־ S נורמלי. לפיכך משפט הפירוק הספקטרלי המרוכב (7.31) מראה שקיים בסיס אורתונורמלי $e_1, \dots, e_n$ של V המורכב מווקטורים עצמיים של S. לכל ערך עצמי של S יש ערך מוחלט 1 (לפי 7.54), ובכך השלמנו את ההוכחה כי (א) גורר את (ב).

כעת נניח ש־ (ב) מתקיים. יהי $e_1, \dots, e_n$ בסיס אורתונורמלי של V המורכב מווקטורים עצמיים של S אשר הערך המוחלט של כל הערכים העצמיים השייכים להם הוא 1. אזי $\lambda_1, \dots, \lambda_n$ מאחר ש־ $Se_1, \dots, Se_n$ גם כן בסיס אורתונורמלי של V

$$\langle Se_j, Se_k \rangle = \langle \lambda_j e_j, \lambda_k e_k \rangle = \lambda_j \overline{\lambda_k} \langle e_j, e_k \rangle = \begin{cases} 0 & \text{if } j \neq k, \\ 1 & \text{if } j = k \end{cases}$$

לכל $j, k = 1, \dots, n$. לפיכך השקילות של (א) ו־ (ד) ב־ 7.53 מראה ש־ S אוניטרי, ובכך הוכחנו כי (ב) גורר את (א). ∎

QR Factorization

In this subsection, we shift our attention from operators to matrices. This switch should give you good practice in identifying an operator with a square matrix (after picking a basis of the vector space on which the operator is defined). You should also become more comfortable with translating concepts and results back and forth between the context of operators and the context of square matrices.

When starting with n-by-n matrices instead of operators, unless otherwise specified assume that the associated operators live on $\mathbf{F}^n$ (with the Euclidean inner product) and that their matrices are computed with respect to the standard basis of $\mathbf{F}^n$.

We begin by making the following definition, transferring the notion of a unitary operator to a unitary matrix.

7.56 definition: *unitary matrix*

An n-by-n matrix is called *unitary* if its columns form an orthonormal list in $\mathbf{F}^n$.

In the definition above, we could have replaced "orthonormal list in $\mathbf{F}^n$" with "orthonormal basis of $\mathbf{F}^n$" because every orthonormal list of length n in an n-dimensional inner product space is an orthonormal basis. If $S \in \mathcal{L}(V)$ and $e_1, \dots, e_n$ and $f_1, \dots, f_n$ are orthonormal bases of V, then S is a unitary operator if and only if $\mathcal{M}(S, (e_1, \dots, e_n), (f_1, \dots, f_n))$ is a unitary matrix, as shown by the equivalence of (a) and (e) in 7.49. Also note that we could also have replaced "columns" in the definition above with "rows" by using the equivalence between conditions (a) and (e) in 7.53.

The next result, whose proof will be left as an exercise for the reader, gives some equivalent conditions for a square matrix to be unitary. In (c), Qv denotes the matrix product of Q and v, identifying elements of $\mathbf{F}^n$ with n-by-1 matrices (sometimes called column vectors). The norm in (c) below is the usual Euclidean norm on $\mathbf{F}^n$ that comes from the Euclidean inner product. In (d), Q^* denotes the conjugate transpose of the matrix Q, which corresponds to the adjoint of the associated operator.

7.57 *characterizations of unitary matrices*

Suppose Q is an n-by-n matrix. Then the following are equivalent.

(a) Q is a unitary matrix.

(b) The rows of Q form an orthonormal list in $\mathbf{F}^n$.

(c) $\|Qv\| = \|v\|$ for every $v \in \mathbf{F}^n$.

(d) $Q^*Q = QQ^* = I$, the n-by-n matrix with 1's on the diagonal and 0's elsewhere.

פירוק QR

בתת־יחידה זו של הפרק, אנו מסיטים את תשומת הלב שלנו מאופרטורים למטריצות. מעבר זה אמור לספק תרגול טוב לזיהוי אופרטור עם מטריצה ריבועית (אחרי שבחרנו בסיס של המרחב הווקטורי שעליו האופרטור מוגדר). בנוסף, עליכם להרגיש בנוח יותר עם תרגום רעיונות ומסקנות הלוך וחזור בין ההקשר של אופרטורים וההקשר של מטריצות ריבועיות.

כשאנו פותחים עם מטריצות $n \times n$ במקום עם אופרטורים, אם לא צוין אחרת עליכם להניח כי האופרטורים המשויכים למטריצות חיים מעל $\mathbf{F}^n$ (עם המכפלה הפנימית האוקלידית) וכי המטריצות מחושבות ביחס לבסיס הסטנדרטי של $\mathbf{F}^n$.

אנו פותחים בהגדרה הבאה, המעבירה את רעיון האופרטור האוניטרי אל מטריצה אוניטרית.

7.56 הגדרה: מטריצה אוניטרית

מטריצה $n \times n$ נקראת אוניטרית אם עמודותיה יוצרות רשימה אורתונורמלית ב־ $\mathbf{F}^n$.

מאחר שכל רשימה אורתונורמלית בת n איברים במרחב מכפלה פנימית מממד n היא בסיס אורתונורמלי, יכולנו לכתוב, בהגדרה לעיל, "בסיס אורתונורמלי של $\mathbf{F}^n$" במקום "רשימה אורתונורמלית ב־ $\mathbf{F}^n$". אם $S \in \mathcal{L}(V)$ ו־ $e_1, \dots, e_n$ ו־ $f_1, \dots, f_n$ הם בסיסים אורתונורמליים של V, אזי S אופרטור אוניטרי אם־ורק־אם $\mathcal{M}(S, (e_1, \dots, e_n), (f_1, \dots, f_n))$ מטריצה אוניטרית, כפי שהראתה לנו השקילות בין (א) ו־ (ה) ב־ 7.49. בנוסף, נשים לב כי בהגדרה לעיל יכולנו להחליף "עמודות" ב"שורות" על ידי שימוש בשקילות הקיימת בין (א) ו־ (ה) ב־ 7.53.

המסקנה הבאה, שהוכחתה הושארה לקוראים כתרגיל, מספקת מספר תנאים שקולים כדי שמטריצה ריבועית תהיה אוניטרית. ב־ (ג), Qv מסמן כפל מטריצות של Q ו־ v, תוך זיהוי איברים של $\mathbf{F}^n$ עם מטריצות $n \times 1$ (לפעמים נקראים וקטורי עמודה). הנורמה ב־ (ג) להלן היא הנורמה האוקלידית הרגילה מעל $\mathbf{F}^n$ שמקורה במכפלה הפנימית האוקלידית. ב־ (ד), Q^* מסמנת את המטריצה הצמודה־משוחלפת של Q, המתאימה לאופרטור הצמוד של האופרטור שאליו משויכת המטריצה.

7.57 אפיונים של מטריצות אוניטריות

נניח ש־ Q מטריצה $n \times n$. אזי התנאים הבאים שקולים.

א. Q מטריצה אוניטרית.

ב. השורות של Q יוצרות רשימה אורתונורמלית ב־ $\mathbf{F}^n$.

ג. $\|Qv\| = \|v\|$ לכל $v \in \mathbf{F}^n$.

ד. $Q^*Q = QQ^* = I$, המטריצה $n \times n$ עם אחדים על האלכסון ואפסים בכל מקום אחר.

The QR factorization stated and proved below is the main tool in the widely used QR algorithm (not discussed here) for finding good approximations to eigenvalues and eigenvectors of square matrices. In the result below, if the matrix A is in $\mathbf{F}^{n,n}$, then the matrices Q and R are also in $\mathbf{F}^{n,n}$.

7.58 *QR factorization*

Suppose A is a square matrix with linearly independent columns. Then there exist unique matrices Q and R such that Q is unitary, R is upper triangular with only positive numbers on its diagonal, and

$$A = QR.$$

Proof Let $v_1, \dots, v_n$ denote the columns of A, thought of as elements of $\mathbf{F}^n$. Apply the Gram–Schmidt procedure (6.32) to the list $v_1, \dots, v_n$, getting an orthonormal basis $e_1, \dots, e_n$ of $\mathbf{F}^n$ such that

$$\text{span}(v_1, \dots, v_k) = \text{span}(e_1, \dots, e_k) \tag{7.59}$$

for each $k = 1, \dots, n$. Let R be the n-by-n matrix defined by

$$R_{j,k} = \langle v_k, e_j \rangle,$$

where $R_{j,k}$ denotes the entry in row j, column k of R. If $j > k$, then e_j is orthogonal to $\text{span}(e_1, \dots, e_k)$ and hence e_j is orthogonal to v_k (by 7.59). In other words, if $j > k$ then $\langle v_k, e_j \rangle = 0$. Thus R is an upper-triangular matrix.

Let Q be the unitary matrix whose columns are $e_1, \dots, e_n$. If $k \in \{1, \dots, n\}$, then the kth column of QR equals a linear combination of the columns of Q, with the coefficients for the linear combination coming from the kth column of R—see 3.51(a). Hence the kth column of QR equals

$$\langle v_k, e_1 \rangle e_1 + \cdots + \langle v_k, e_k \rangle e_k,$$

which equals v_k [by 6.30(a)], the kth column of A. Thus $A = QR$, as desired.

The equations defining the Gram–Schmidt procedure (see 6.32) show that each v_k equals a positive multiple of e_k plus a linear combination of $e_1, \dots, e_{k-1}$. Thus each $\langle v_k, e_k \rangle$ is a positive number. Hence all entries on the diagonal of R are positive numbers, as desired.

Finally, to show that Q and R are unique, suppose we also have $A = \widehat{Q}\,\widehat{R}$, where $\widehat{Q}$ is unitary and $\widehat{R}$ is upper triangular with only positive numbers on its diagonal. Let $q_1, \dots, q_n$ denote the columns of $\widehat{Q}$. Thinking of matrix multiplication as above, we see that each v_k is a linear combination of $q_1, \dots, q_k$, with the coefficients coming from the kth column of $\widehat{R}$. This implies that $\text{span}(v_1, \dots, v_k) = \text{span}(q_1, \dots, q_k)$ and $\langle v_k, q_k \rangle > 0$. The uniqueness of the orthonormal lists satisfying these conditions (see Exercise 10 in Section 6B) now shows that $q_k = e_k$ for each $k = 1, \dots, n$. Hence $\widehat{Q} = Q$, which then implies that $\widehat{R} = R$, completing the proof of uniqueness. ∎

הפירוק QR המנוסח ומוכח להלן הוא הכלי המרכזי באלגוריתם QR הנפוץ בשימוש (לא נעסוק בו כאן) למציאת קירובים טובים לערכים עצמיים ווקטורים עצמיים של מטריצות ריבועיות. במסקנה להלן, אם המטריצה A היא ב־ $\mathbf{F}^{n,n}$, אזי המטריצות Q ו־ R גם הן ב־ $\mathbf{F}^{n,n}$.

7.58 פירוק QR

נניח ש־ A מטריצה ריבועית עם עמודות בלתי תלויות ליניארית. אזי קיימות מטריצות Q ו־ R יחידות כך ש־ Q אוניטרית, R משולשית עליונה עם מספרים חיוביים בלבד על האלכסון, ו־

$$A = QR.$$

הוכחה יהיו $v_1, \dots, v_n$ עמודותיה של A, כשאנו חושבים עליהן כאיברים של $\mathbf{F}^n$. הפעלת תהליך גרם־שמידט (6.32) על הרשימה $v_1, \dots, v_n$, מניבה בסיס אורתונורמלי $e_1, \dots, e_n$ של $\mathbf{F}^n$ כך ש־

$$\operatorname{span}(v_1, \dots, v_k) = \operatorname{span}(e_1, \dots, e_k) \tag{7.59}$$

לכל $k = 1, \dots, n$. תהי R המטריצה $n \times n$ המוגדרת על ידי

$$R_{j,k} = \langle v_k, e_j \rangle,$$

כאשר $R_{j,k}$ מסמן את האיבר בשורה j, עמודה k של R. אם $j > k$, אזי e_j אורתוגונלי ל־ $\operatorname{span}(e_1, \dots, e_k)$ ולכן e_j אורתוגונלי ל־ v_k (לפי 7.59). במילים אחרות, אם $j > k$ אזי $\langle v_k, e_j \rangle = 0$. לפיכך R מטריצה משולשית עליונה.

תהי Q המטריצה האוניטרית שעמודותיה הן $e_1, \dots, e_n$. אם $k \in \{1, \dots, n\}$, אזי העמודה ה־ k^{th} של QR שווה לצירוף ליניארי של העמודות של Q, עם מקדמי הצירוף הליניארי המגיעים מהעמודה ה־ k^{th} של R—ראו 3.51(א). לכן העמודה ה־ k^{th} של QR שווה ל־

$$\langle v_k, e_1 \rangle e_1 + \cdots + \langle v_k, e_k \rangle e_k,$$

השווה ל־ v_k [לפי 6.30(א)], העמודה ה־ k^{th} של A. לפיכך $A = QR$, כפי שרצינו.

המשוואות המגדירות את תהליך גרם־שמידט (ראו 6.32) מראות כי כל v_k שווה לכפולה חיובית של e_k בתוספת צירוף ליניארי של $e_1, \dots, e_{k-1}$. לפיכך כל $\langle v_k, e_k \rangle$ מספר חיובי. לכן איברי האלכסון של R הם מספרים חיוביים, כפי שרצינו.

לסיום, כדי להראות כי Q ו־ R יחידות, נניח שיש לנו בנוסף גם את $A = \widehat{Q}\widehat{R}$, כאשר $\widehat{Q}$ אוניטרית ו־ $\widehat{R}$ משולשית עליונה עם מספרים חיוביים בלבד על האלכסון. יהיו $q_1, \dots, q_n$ עמודות $\widehat{Q}$. כשחושבים על כפל המטריצות כמובא לעיל, אנו רואים כי כל v_k הוא צירוף ליניארי של $q_1, \dots, q_k$, עם מקדמים מהעמודה ה־ k^{th} של $\widehat{R}$. מכאן נובע ש־ $\operatorname{span}(v_1, \dots, v_k) = \operatorname{span}(q_1, \dots, q_k)$ ו־ $\langle v_k, q_k \rangle > 0$. מהיְחִידוּת של רשימות אורתונורמליות המקיימות את התנאים הללו (ראו תרגיל 10 ביחידה 21) אנו מסיקים כי $q_k = e_k$ לכל $k = 1, \dots, n$. לכן $\widehat{Q} = Q$, מה שמוביל למסקנה כי $\widehat{R} = R$, ובזאת השלמנו את הוכחת היְחִידוּת.

The proof of the QR factorization shows that the columns of the unitary matrix can be computed by applying the Gram–Schmidt procedure to the columns of the matrix to be factored. The next example illustrates the computation of the QR factorization based on the proof that we just completed.

7.60 example: *QR factorization of a 3-by-3 matrix*

To find the QR factorization of the matrix

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & -4 \\ 0 & 3 & 2 \end{pmatrix},$$

follow the proof of 7.58. Thus set v_1, v_2, v_3 equal to the columns of A:

$$v_1 = (1, 0, 0), \quad v_2 = (2, 1, 3), \quad v_3 = (1, -4, 2).$$

Apply the Gram–Schmidt procedure to v_1, v_2, v_3, producing the orthonormal list

$$e_1 = (1, 0, 0), \quad e_2 = \Big(0, \tfrac{1}{\sqrt{10}}, \tfrac{3}{\sqrt{10}}\Big), \quad e_3 = \Big(0, -\tfrac{3}{\sqrt{10}}, \tfrac{1}{\sqrt{10}}\Big).$$

Still following the proof of 7.58, let Q be the unitary matrix whose columns are e_1, e_2, e_3:

$$Q = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{10}} & -\frac{3}{\sqrt{10}} \\ 0 & \frac{3}{\sqrt{10}} & \frac{1}{\sqrt{10}} \end{pmatrix}.$$

As in the proof of 7.58, let R be the 3-by-3 matrix whose entry in row j, column k is $\langle v_k, e_j \rangle$, which gives

$$R = \begin{pmatrix} 1 & 2 & 1 \\ 0 & \sqrt{10} & \frac{\sqrt{10}}{5} \\ 0 & 0 & \frac{7\sqrt{10}}{5} \end{pmatrix}.$$

Note that R is indeed an upper-triangular matrix with only positive numbers on the diagonal, as required by the QR factorization.

Now matrix multiplication can verify that $A = QR$ is the desired factorization of A:

$$QR = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{10}} & -\frac{3}{\sqrt{10}} \\ 0 & \frac{3}{\sqrt{10}} & \frac{1}{\sqrt{10}} \end{pmatrix} \begin{pmatrix} 1 & 2 & 1 \\ 0 & \sqrt{10} & \frac{\sqrt{10}}{5} \\ 0 & 0 & \frac{7\sqrt{10}}{5} \end{pmatrix} = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & -4 \\ 0 & 3 & 2 \end{pmatrix} = A.$$

Thus $A = QR$, as expected.

The QR factorization will be the major tool used in the proof of the Cholesky factorization (7.63) in the next subsection. For another nice application of the QR factorization, see the proof of Hadamard's inequality (9.66).

ההוכחה של הפירוק QR מראה שעמודות המטריצה האוניטרית ניתנות לחישוב באמצעות הפעלת תהליך גרם־שמידט על עמודות המטריצה המועמדת לפירוק. הדוגמה הבאה ממחישה את החישוב של הפירוק QR בהתבסס על ההוכחה שזה עתה השלמנו.

7.60 דוגמה: פירוק QR של מטריצה 3×3

למציאת הפירוק QR של המטריצה

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & -4 \\ 0 & 3 & 2 \end{pmatrix},$$

נלך בעקבות ההוכחה של 7.58. לפיכך נגדיר v_1, v_2, v_3 השווים לעמודות של A:

$$v_1 = (1,0,0), \quad v_2 = (2,1,3), \quad v_3 = (1,-4,2).$$

יישום תהליך גרם־שמידט על v_1, v_2, v_3 מפיק את הרשימה האורתונורמלית

$$e_1 = (1,0,0), \quad e_2 = \Big(0, \tfrac{1}{\sqrt{10}}, \tfrac{3}{\sqrt{10}}\Big), \quad e_3 = \Big(0, -\tfrac{3}{\sqrt{10}}, \tfrac{1}{\sqrt{10}}\Big).$$

נמשיך בעקבות ההוכחה של 7.58, ונגדיר Q להיות המטריצה האוניטרית שעמודותיה הן e_1, e_2, e_3:

$$Q = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{10}} & -\frac{3}{\sqrt{10}} \\ 0 & \frac{3}{\sqrt{10}} & \frac{1}{\sqrt{10}} \end{pmatrix}.$$

כמו בהוכחה של 7.58, תהי R המטריצה 3×3 שהאיבר בשורה j, עמודה k הוא $\langle v_k, e_j \rangle$, מכאן נקבל

$$R = \begin{pmatrix} 1 & 2 & 1 \\ 0 & \sqrt{10} & \frac{\sqrt{10}}{5} \\ 0 & 0 & \frac{7\sqrt{10}}{5} \end{pmatrix}.$$

שימו לב ש־ R בהחלט מטריצה משולשית עליונה עם מספרים חיוביים בלבד על האלכסון, כפי שדורש פירוק QR.

כעת באמצעות כפל מטריצות נוכל לוודא כי $A = QR$ הוא הפירוק לגורמים של A שרצינו:

$$QR = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{10}} & -\frac{3}{\sqrt{10}} \\ 0 & \frac{3}{\sqrt{10}} & \frac{1}{\sqrt{10}} \end{pmatrix} \begin{pmatrix} 1 & 2 & 1 \\ 0 & \sqrt{10} & \frac{\sqrt{10}}{5} \\ 0 & 0 & \frac{7\sqrt{10}}{5} \end{pmatrix} = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & -4 \\ 0 & 3 & 2 \end{pmatrix} = A.$$

לפיכך $A = QR$, כפי שציפינו.

הפירוק QR יהיה הכלי המרכזי בהוכחה של פירוק שולסקי (7.63) בתת־החלק הבא. ליישום נחמד נוסף של פירוק QR, ראו ההוכחה של אי־שוויון הדמר (9.66).

If a QR factorization is available, then it can be used to solve a corresponding system of linear equations without using Gaussian elimination. Specifically, suppose A is an n-by-n square matrix with linearly independent columns. Suppose that $b \in \mathbf{F}^n$ and we want to solve the equation $Ax = b$ for $x = (x_1, \dots, x_n) \in \mathbf{F}^n$ (as usual, we are identifying elements of $\mathbf{F}^n$ with n-by-1 column vectors).

Suppose $A = QR$, where Q is unitary and R is upper triangular with only positive numbers on its diagonal (Q and R are computable from A using just the Gram–Schmidt procedure, as shown in the proof of 7.58). The equation $Ax = b$ is equivalent to the equation $QRx = b$. Multiplying both sides of this last equation by Q^* on the left and using 7.57(d) gives the equation

$$Rx = Q^*b.$$

The matrix Q^* is the conjugate transpose of the matrix Q. Thus computing Q^*b is straightforward. Because R is an upper-triangular matrix with positive numbers on its diagonal, the system of linear equations represented by the equation above can quickly be solved by first solving for x_n, then for x_{n-1}, and so on.

Cholesky Factorization

We begin this subsection with a characterization of positive invertible operators in terms of inner products.

7.61 *positive invertible operator*

A self-adjoint operator $T \in \mathcal{L}(V)$ is a positive invertible operator if and only if $\langle Tv, v\rangle > 0$ for every nonzero $v \in V$.

Proof First suppose T is a positive invertible operator. If $v \in V$ and $v \neq 0$, then because T is invertible we have $Tv \neq 0$. This implies that $\langle Tv, v\rangle \neq 0$ (by 7.43). Hence $\langle Tv, v\rangle > 0$.

To prove the implication in the other direction, suppose now that $\langle Tv, v\rangle > 0$ for every nonzero $v \in V$. Thus $Tv \neq 0$ for every nonzero $v \in V$. Hence T is injective. Thus T is invertible, as desired. ∎

The next definition transfers the result above to the language of matrices. Here we are using the usual Euclidean inner product on $\mathbf{F}^n$ and identifying elements of $\mathbf{F}^n$ with n-by-1 column vectors.

7.62 definition: *positive definite*

A matrix $B \in \mathbf{F}^{n,n}$ is called *positive definite* if $B^* = B$ and

$$\langle Bx, x\rangle > 0$$

for every nonzero $x \in \mathbf{F}^n$.

אם פירוק QR זמין, אזי ניתן להשתמש בו לפתרון מערכת משוואות ליניאריות ללא שימוש בשיטת החילוץ של גאוס. ליתר דיוק, נניח ש־ A מטריצה ריבועית $n \times n$ עם עמודות בלתי תלויות ליניארית. נניח ש־ $b \in \mathbf{F}^n$ ואנו מעוניינים לפתור את המשוואה $Ax = b$ עבור $x = (x_1, \dots, x_n) \in \mathbf{F}^n$ (כהרגלנו, אנו מזהים את האיברים של $\mathbf{F}^n$ עם וקטורי עמודה $n \times 1$).

נניח ש־ $A = QR$, כאשר Q אוניטרית ו־ R משולשית עליונה עם מספרים חיוביים בלבד על האלכסון (Q ו־ R ניתנות לחישוב מ־ A בשימוש בתהליך גרם־שמידט בלבד, כפי שמוצג בהוכחה של 7.58). המשוואה $Ax = b$ שקולה אפוא למשוואה $QRx = b$. בהכפלת שני האגפים של משוואה אחרונה זו ב־ Q^* מצד שמאל תוך שימוש ב־ 7.57(ד) נקבל את המשוואה

$$Rx = Q^*b.$$

המטריצה Q^* היא הצמודה־משוחלפת של המטריצה Q. לפיכך חישוב Q^*b הוא פשוט וישיר. מאחר ש־ R מטריצה משולשית עליונה עם מספרים חיוביים בלבד על האלכסון, מערכת המשוואות הליניאריות המיוצגת על ידי המשוואה לעיל יכולה להיפתר במהירות בכך שנפתור תחילה עבור x_n, ואז עבור x_{n-1}, וכן הלאה.

פירוק שולסקי

אנו פותחים תת־יחידה זו של הפרק עם אפיון של אופרטורים חיוביים הפיכים במונחים של מכפלות פנימיות.

7.61 אופרטור חיובי הפיך

אופרטור צמוד לעצמו $T \in \mathcal{L}(V)$ הוא אופרטור חיובי הפיך אם־ורק־אם $\langle Tv, v\rangle > 0$ לכל $v \in V$ השונה מאפס.

הוכחה תחילה נניח כי T אופרטור חיובי הפיך. אם $v \in V$ ו־ $v \neq 0$, אזי מאחר ש־ T הפיך נקבל ש־ $Tv \neq 0$. מכאן נובע ש־ $\langle Tv, v\rangle \neq 0$ (לפי 7.43). לכן $\langle Tv, v\rangle > 0$.

להוכחת הגרירה בכיוון ההפוך, נניח כעת ש־ $\langle Tv, v\rangle > 0$ לכל $v \in V$ השונה מאפס. לפיכך $Tv \neq 0$ לכל $v \in V$ השונה מאפס. לכן T חד־חד־ערכי. לפיכך T הפיך, כפי שרצינו. ∎

ההגדרה הבאה מעבירה את המסקנה לעיל לשפה של מטריצות. כאן אנו משתמשים במכפלה הפנימית האוקלידית הרגילה מעל $\mathbf{F}^n$ ומזהים את איברי $\mathbf{F}^n$ עם וקטורי עמודה $n \times 1$.

7.62 הגדרה: חיובית לחלוטין

מטריצה $B \in \mathbf{F}^{n,n}$ נקראת חיובית לחלוטין אם $B^* = B$ וגם

$$\langle Bx, x\rangle > 0$$

לכל $x \in \mathbf{F}^n$ השונה מאפס.

A matrix is upper triangular if and only if its conjugate transpose is lower triangular (meaning that all entries above the diagonal are 0). The factorization below, which has important consequences in computational linear algebra, writes a positive definite matrix as the product of a lower triangular matrix and its conjugate transpose.

Our next result is solely about matrices, although the proof makes use of the identification of results about operators with results about square matrices. In the result below, if the matrix B is in $\mathbf{F}^{n,n}$, then the matrix R is also in $\mathbf{F}^{n,n}$.

7.63 *Cholesky factorization*

Suppose B is a positive definite matrix. Then there exists a unique upper-triangular matrix R with only positive numbers on its diagonal such that

$$B = R^*R.$$

Proof Because B is positive definite, there exists an invertible square matrix A of the same size as B such that $B = A^*A$ [by the equivalence of (a) and (f) in 7.38].

Let $A = QR$ be the QR factorization of A (see 7.58), where Q is unitary and R is upper triangular with only positive numbers on its diagonal. Then $A^* = R^*Q^*$. Thus

$$B = A^*A = R^*Q^*QR = R^*R,$$

André-Louis Cholesky (1875–1918) discovered this factorization, which was published posthumously in 1924.

as desired.

To prove the uniqueness part of this result, suppose S is an upper-triangular matrix with only positive numbers on its diagonal and $B = S^*S$. The matrix S is invertible because B is invertible (see Exercise 11 in Section 3D). Multiplying both sides of the equation $B = S^*S$ by S^{-1} on the right gives the equation $BS^{-1} = S^*$.

Let A be the matrix from the first paragraph of this proof. Then

$$\begin{aligned}(AS^{-1})^*(AS^{-1}) &= (S^*)^{-1}A^*AS^{-1}\\ &= (S^*)^{-1}BS^{-1}\\ &= (S^*)^{-1}S^*\\ &= I.\end{aligned}$$

Thus AS^{-1} is unitary.

Hence $A = (AS^{-1})S$ is a factorization of A as the product of a unitary matrix and an upper-triangular matrix with only positive numbers on its diagonal. The uniqueness of the QR factorization, as stated in 7.58, now implies that $S = R$. ∎

In the first paragraph of the proof above, we could have chosen A to be the unique positive definite matrix that is a square root of B (see 7.39). However, the proof was presented with the more general choice of A because for specific positive definite matrices B, it may be easier to find a different choice of A.

מטריצה היא משולשית עליונה אם־ורק־אם הצמודה־משוחלפת שלה היא משולשית תחתונה (כלומר כל האיברים מעל לאלכסון הם 0). הפירוק לגורמים להלן, שהוא בעל השלכות חשובות באלגברה ליניארית חישובית, כותב מטריצה חיובית לחלוטין כמכפלה של מטריצה משולשית תחתונה והצמודה־משוחלפת שלה.

מסקנתנו הבאה היא על מטריצות בלבד, אף שההוכחה משתמשת בזיהוי תוצאות עבור אופרטורים עם תוצאות עבור מטריצות ריבועיות. במסקנה להלן, אם המטריצה B היא ב־ $\mathbf{F}^{n,n}$, אזי המטריצה R גם היא ב־ $\mathbf{F}^{n,n}$.

7.63 פירוק שולסקי

נניח ש־ B מטריצה חיובית לחלוטין. אזי קיימת מטריצה משולשית עליונה יחידה R עם מספרים חיוביים בלבד על האלכסון כך ש־

$$B = R^*R.$$

הוכחה מאחר ש־ B חיובית לחלוטין, קיימת מטריצה ריבועית הפיכה A מאותו הגודל של B כך ש־ $B = A^*A$ [לפי השקילות של (א) ו־ (ו) ב־ 7.38].

יהי $A = QR$ הפירוק QR של A (ראו 7.58), כאשר Q אוניטרית ו־ R משולשית עליונה עם מספרים חיוביים בלבד על האלכסון. אזי $A^* = R^*Q^*$.

אנדרה־לואי שולסקי (1875–1918) גילה פירוק זה, שפורסם לאחר מותו בשנת 1924.

לפיכך

$$B = A^*A = R^*Q^*QR = R^*R,$$

כפי שרצינו.

להוכחת חלק היחידות של המסקנה, נניח ש־ S מטריצה משולשית עליונה עם מספרים חיוביים בלבד על האלכסון ו־ $B = S^*S$. המטריצה S הפיכה מאחר ש־ B הפיכה (ראו תרגיל 11 ביחידה 4ג). הכפלת שני אגפי המשוואה $B = S^*S$ ב־ S^{-1} מימין נותנת את המשוואה $BS^{-1} = S^*$.

תהי A המטריצה מהפסקה הראשונה של ההוכחה. אזי

$$\begin{aligned}(AS^{-1})^*(AS^{-1}) &= (S^*)^{-1}A^*AS^{-1}\\ &= (S^*)^{-1}BS^{-1}\\ &= (S^*)^{-1}S^*\\ &= I.\end{aligned}$$

לפיכך AS^{-1} אוניטרית.

לכן $A = (AS^{-1})S$ הוא פירוק לגורמים של A כמכפלה של מטריצה אוניטרית ומטריצה משולשית עליונה עם מספרים חיוביים בלבד על האלכסון. היחידות של פירוק QR, כפי שנוסחה ב־ 7.58, גוררת כעת ש־ $S = R$. ■

בפסקה הראשונה של ההוכחה לעיל, יכולנו לבחור את A כמטריצה החיובית לחלוטין היחידה שהיא שורש ריבועי של B (ראו 7.39). בכל זאת, ההוכחה הוצגה עם הבחירה הכללית יותר של A מכיוון שעבור מטריצות ספציפיות חיוביות לחלוטין B, ייתכן שיהיה קל יותר למצוא בחירה אחרת של A.

Exercises 7D

1 Suppose $\dim V \geq 2$ and $S \in \mathcal{L}(V, W)$. Prove that S is an isometry if and only if Se_1, Se_2 is an orthonormal list in W for every orthonormal list e_1, e_2 of length two in V.

2 Suppose $T \in \mathcal{L}(V, W)$ and $T \neq 0$. Prove that T is a scalar multiple of an isometry if and only if T preserves orthogonality.

The phrase "T preserves orthogonality" means that $\langle Tu, Tv\rangle = 0$ for all $u, v \in V$ such that $\langle u, v\rangle = 0$.

3 (a) Show that the product of two unitary operators on V is a unitary operator.
(b) Show that the inverse of a unitary operator on V is a unitary operator.

This exercise shows that the set of unitary operators on V is a group, where the group operation is the usual product of two operators.

4 Suppose $\mathbf{F} = \mathbf{C}$ and $A, B \in \mathcal{L}(V)$ are self-adjoint. Show that $A + iB$ is unitary if and only if $AB = BA$ and $A^2 + B^2 = I$.

5 Suppose $S \in \mathcal{L}(V)$. Prove that the following are equivalent.
(a) S is a self-adjoint unitary operator.
(b) $S = 2P - I$ for some orthogonal projection P on V.
(c) There exists a subspace U of V such that $Su = u$ for every $u \in U$ and $Sw = -w$ for every $w \in U^{\perp}$.

6 Suppose T_1, T_2 are both normal operators on $\mathbf{F}^3$ with $2, 5, 7$ as eigenvalues. Prove that there exists a unitary operator $S \in \mathcal{L}(\mathbf{F}^3)$ such that $T_1 = S^* T_2 S$.

7 Give an example of two self-adjoint operators $T_1, T_2 \in \mathcal{L}(\mathbf{F}^4)$ such that the eigenvalues of both operators are $2, 5, 7$ but there does not exist a unitary operator $S \in \mathcal{L}(\mathbf{F}^4)$ such that $T_1 = S^* T_2 S$. Be sure to explain why there is no unitary operator with the required property.

8 Prove or give a counterexample: If $S \in \mathcal{L}(V)$ and there exists an orthonormal basis $e_1, \ldots, e_n$ of V such that $\|Se_k\| = 1$ for each e_k, then S is a unitary operator.

9 Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$. Suppose every eigenvalue of T has absolute value 1 and $\|Tv\| \leq \|v\|$ for every $v \in V$. Prove that T is a unitary operator.

10 Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$ is a self-adjoint operator such that $\|Tv\| \leq \|v\|$ for all $v \in V$.
(a) Show that $I - T^2$ is a positive operator.
(b) Show that $T + i\sqrt{I - T^2}$ is a unitary operator.

11 Suppose $S \in \mathcal{L}(V)$. Prove that S is a unitary operator if and only if

$$\{Sv : v \in V \text{ and } \|v\| \leq 1\} = \{v \in V : \|v\| \leq 1\}.$$

12 Prove or give a counterexample: If $S \in \mathcal{L}(V)$ is invertible and $\|S^{-1}v\| = \|Sv\|$ for every $v \in V$, then S is unitary.

תרגילים 7ג

1 נניח ש־ $\dim V \geq 2$ ו־ $S \in \mathcal{L}(V, W)$. הוכיחו כי S איזומטריה אם־ורק־אם Se_1, Se_2 רשימה אורתונורמלית ב־ W לכל רשימה אורתונורמלית e_1, e_2 של שני איברים ב־ V.

2 נניח ש־ $T \in \mathcal{L}(V, W)$ ו־ $T \neq 0$. הוכיחו כי T כפולת סקלר של איזומטריה אם־ורק־אם T משמרת אורתוגונליות.

הביטוי "T משמרת אורתוגונליות" משמעותו ש־ $\langle Tu, Tv \rangle = 0$ לכל $u, v \in V$ כך ש־ $\langle u, v \rangle = 0$.

3 א. הראו כי המכפלה של שני אופרטורים אוניטריים מעל V היא אופרטור אוניטרי.

ב. הראו כי ההופכי של אופרטור אוניטרי מעל V הוא אופרטור אוניטרי.

תרגיל זה מראה שהקבוצה של האופרטורים האוניטריים מעל V היא חבורה, כאשר פעולת החבורה היא המכפלה הרגילה של שני אופרטורים.

4 נניח ש־ $\mathbf{F} = \mathbf{C}$ ו־ $A, B \in \mathcal{L}(V)$ צמודים לעצמם. הראו כי $A + iB$ אוניטרי אם־ורק־אם $AB = BA$ וגם $A^2 + B^2 = I$.

5 נניח ש־ $S \in \mathcal{L}(V)$. הוכיחו כי התנאים הבאים שקולים.

א. S אופרטור אוניטרי צמוד לעצמו.

ב. $S = 2P - I$ עבור היטל אורתוגונלי P כלשהו מעל V.

ג. קיים תת־מרחב U של V כך ש־ $Su = u$ לכל $u \in U$ ו־ $Sw = -w$ לכל $w \in U^\perp$.

6 נניח ש־ T_1, T_2 שניהם אופרטורים נורמליים מעל $\mathbf{F}^3$ עם $2, 5, 7$ כערכים עצמיים. הוכיחו כי קיים אופרטור אוניטרי $S \in \mathcal{L}(\mathbf{F}^3)$ כך ש־ $T_1 = S^* T_2 S$.

7 הביאו דוגמה של שני אופרטורים צמודים לעצמם $T_1, T_2 \in \mathcal{L}(\mathbf{F}^4)$ כך שהערכים העצמיים של שני האופרטורים הם $2, 5, 7$ אך לא קיים אופרטור אוניטרי $S \in \mathcal{L}(\mathbf{F}^4)$ כך ש־ $T_1 = S^* T_2 S$. ודאו כי אתם מסבירים מדוע אין בנמצא אופרטור אוניטרי עם התכונה הנדרשת.

8 הוכיחו או הביאו דוגמה נגדית: אם $S \in \mathcal{L}(V)$ וקיים בסיס אורתונורמלי $e_1, \dots, e_n$ של V כך ש־ $\|Se_k\| = 1$ לכל e_k, אזי S אופרטור אוניטרי.

9 נניח ש־ $\mathbf{F} = \mathbf{C}$ ו־ $T \in \mathcal{L}(V)$. נניח שלכל ערך עצמי של T יש ערך מוחלט 1 וגם $\|Tv\| \leq \|v\|$ לכל $v \in V$. הוכיחו כי T אופרטור אוניטרי.

10 נניח ש־ $\mathbf{F} = \mathbf{C}$ ו־ $T \in \mathcal{L}(V)$ אופרטור צמוד לעצמו כך ש־ $\|Tv\| \leq \|v\|$ לכל $v \in V$.

א. הראו כי $\sqrt{I - T^2}$ אופרטור חיובי.

ב. הראו כי $T + i\sqrt{I - T^2}$ אופרטור אוניטרי.

11 נניח ש־ $S \in \mathcal{L}(V)$. הוכיחו כי S אופרטור אוניטרי אם־ורק־אם

$$\{Sv : v \in V \text{ and } \|v\| \leq 1\} = \{v \in V : \|v\| \leq 1\}.$$

12 הוכיחו או הביאו דוגמה נגדית: אם $S \in \mathcal{L}(V)$ הפיך ו־ $\|Sv\| = \|S^{-1}v\|$ לכל $v \in V$, אזי S אוניטרי.

13 Explain why the columns of a square matrix of complex numbers form an orthonormal list in $\mathbf{C}^n$ if and only if the rows of the matrix form an orthonormal list in $\mathbf{C}^n$.

14 Suppose $v \in V$ with $\|v\| = 1$ and $b \in \mathbf{F}$. Also suppose $\dim V \geq 2$. Prove that there exists a unitary operator $S \in \mathcal{L}(V)$ such that $\langle Sv, v\rangle = b$ if and only if $|b| \leq 1$.

15 Suppose T is a unitary operator on V such that $T - I$ is invertible.

(a) Prove that $(T + I)(T - I)^{-1}$ is a skew operator (meaning that it equals the negative of its adjoint).

(b) Prove that if $\mathbf{F} = \mathbf{C}$, then $i(T + I)(T - I)^{-1}$ is a self-adjoint operator.

The function $z \mapsto i(z + 1)(z - 1)^{-1}$ maps the unit circle in $\mathbf{C}$ (except for the point 1) to $\mathbf{R}$. Thus (b) illustrates the analogy between the unitary operators and the unit circle in $\mathbf{C}$, along with the analogy between the self-adjoint operators and $\mathbf{R}$.

16 Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$ is self-adjoint. Prove that $(T + iI)(T - iI)^{-1}$ is a unitary operator and 1 is not an eigenvalue of this operator.

17 Explain why the characterizations of unitary matrices given by 7.57 hold.

18 A square matrix A is called *symmetric* if it equals its transpose. Prove that if A is a symmetric matrix with real entries, then there exists a unitary matrix Q with real entries such that $Q^* A Q$ is a diagonal matrix.

19 Suppose n is a positive integer. For this exercise, we adopt the notation that a typical element z of $\mathbf{C}^n$ is denoted by $z = (z_0, z_1, \ldots, z_{n-1})$. Define linear functionals $\omega_0, \omega_1, \ldots, \omega_{n-1}$ on $\mathbf{C}^n$ by

$$\omega_j(z_0, z_1, \ldots, z_{n-1}) = \frac{1}{\sqrt{n}} \sum_{m=0}^{n-1} z_m e^{-2\pi i j m/n}.$$

The *discrete Fourier transform* is the operator $\mathcal{F} : \mathbf{C}^n \to \mathbf{C}^n$ defined by

$$\mathcal{F}z = (\omega_0(z), \omega_1(z), \ldots, \omega_{n-1}(z)).$$

(a) Show that $\mathcal{F}$ is a unitary operator on $\mathbf{C}^n$.

(b) Show that if $(z_0, \ldots, z_{n-1}) \in \mathbf{C}^n$ and z_n is defined to equal z_0, then

$$\mathcal{F}^{-1}(z_0, z_1, \ldots, z_{n-1}) = \mathcal{F}(z_n, z_{n-1}, \ldots, z_1).$$

(c) Show that $\mathcal{F}^4 = I$.

The discrete Fourier transform has many important applications in data analysis. The usual Fourier transform involves expressions of the form $\int_{-\infty}^{\infty} f(x) e^{-2\pi i t x} dx$ for complex-valued integrable functions f defined on $\mathbf{R}$.

20 Suppose A is a square matrix with linearly independent columns. Prove that there exist unique matrices R and Q such that R is lower triangular with only positive numbers on its diagonal, Q is unitary, and $A = RQ$.

13 הסבירו מדוע העמודות של מטריצה ריבועית של מספרים מרוכבים יוצרות רשימה אורתונורמלית ב־ $\mathbf{C}^n$ אם־ורק־אם השורות של המטריצה יוצרות רשימה אורתונורמלית ב־ $\mathbf{C}^n$.

14 נניח ש־ $v \in V$ עם $\|v\| = 1$ ו־ $b \in \mathbf{F}$. נניח בנוסף ש־ $\dim V \geq 2$. הוכיחו כי קיים אופרטור אוניטרי $S \in \mathcal{L}(V)$ כך ש־ $\langle Sv, v\rangle = b$ אם־ורק־אם $|b| \leq 1$.

15 נניח ש־ T אופרטור אוניטרי מעל V כך ש־ $T - I$ הפיך.

א. הוכיחו כי $(T + I)(T - I)^{-1}$ אופרטור אנטי־צמוד (כלומר שווה למינוס של האופרטור הצמוד שלו).

ב. הוכיחו כי אם $\mathbf{F} = \mathbf{C}$, אזי $i(T + I)(T - I)^{-1}$ אופרטור צמוד לעצמו.

הפונקציה $z \mapsto i(z+1)(z-1)^{-1}$ מעתיקה את מעגל היחידה ב־ $\mathbf{C}$ (פלבד הנקודה 1) אל $\mathbf{R}$. לפיכך (ב) מפחיש את האנלוגיה בין אופרטורים אוניטריים ומעגל היחידה ב־ $\mathbf{C}$, יחד עם האנלוגיה בין אופרטורים הצמודים לעצמם ו־ $\mathbf{R}$.

16 נניח ש־ $\mathbf{F} = \mathbf{C}$ ו־ $T \in \mathcal{L}(V)$ צמוד לעצמו. הוכיחו כי $(T + iI)(T - iI)^{-1}$ אופרטור אוניטרי וכי 1 אינו ערך עצמי של אופרטור זה.

17 הסבירו מדוע האפיונים של מטריצות אוניטריות הנתון ב־ 7.57 מתקיים.

18 מטריצה ריבועית A נקראת סימטרית אם היא שווה למשוחלפת שלה. הוכיחו כי אם A מטריצה סימטרית עם איברים ממשיים, אזי קיימת מטריצה אוניטרית Q עם איברים ממשיים כך ש־ Q^*AQ מטריצה אלכסונית.

19 נניח ש־ n מספר שלם חיובי. עבור תרגיל זה, אנו מאמצים את הסימון שלפיו איבר אופייני z של $\mathbf{C}^n$ מסומן על ידי $z = (z_0, z_1, \dots, z_{n-1})$. נגדיר פונקציונלים ליניאריים $\omega_0, \omega_1, \dots, \omega_{n-1}$ מעל $\mathbf{C}^n$ על ידי

$$\omega_j(z_0, z_1, \dots, z_{n-1}) = \frac{1}{\sqrt{n}} \sum_{m=0}^{n-1} z_m \, e^{-2\pi i j m/n}.$$

התמרת פורייה בדידה היא האופרטור $\mathcal{F} : \mathbf{C}^n \to \mathbf{C}^n$ המוגדר על ידי

$$\mathcal{F}z = (\omega_0(z), \omega_1(z), \dots, \omega_{n-1}(z)).$$

א. הראו כי $\mathcal{F}$ אופרטור אוניטרי מעל $\mathbf{C}^n$.

ב. הראו כי אם $(z_0, \dots, z_{n-1}) \in \mathbf{C}^n$ ו־ z_n מוגדר כשווה ל־ z_0, אזי

$$\mathcal{F}^{-1}(z_0, z_1, \dots, z_{n-1}) = \mathcal{F}(z_n, z_{n-1}, \dots, z_1).$$

ג. הראו כי $\mathcal{F}^4 = I$.

להתמרת פורייה בדידה יש יישומים חשובים רבים בתחום ניתוח נתונים. התמרת פורייה רגילה כוללת ביטויים מהצורה $\int_{-\infty}^{\infty} f(x)e^{-2\pi itx}dx$ עבור פונקציות מרוכבות אינטגרביליות f המוגדרות מעל $\mathbf{R}$.

20 נניח ש־ A מטריצה ריבועית עם עמודות בלתי תלויות ליניארית. הוכיחו כי קיימות מטריצות יחידות R ו־ Q כך ש־ R משולשית תחתונה עם מספרים חיוביים בלבד על האלכסון, Q אוניטרית, ו־ $A = RQ$.

אלגברה ליניארית בדרך הישר, מהדורה רביעית, מחבר: שלדון אקסלר
Linear Algebra Done Right, fourth edition, by Sheldon Axler

7E Singular Value Decomposition

Singular Values

We will need the following result in this section.

7.64 *properties of T^*T*

Suppose $T \in \mathcal{L}(V, W)$. Then

(a) T^*T is a positive operator on V;

(b) $\operatorname{null} T^*T = \operatorname{null} T$;

(c) $\operatorname{range} T^*T = \operatorname{range} T^*$;

(d) $\dim \operatorname{range} T = \dim \operatorname{range} T^* = \dim \operatorname{range} T^*T$.

Proof

(a) We have

$$(T^*T)^* = T^*(T^*)^* = T^*T.$$

Thus T^*T is self-adjoint.

If $v \in V$, then

$$\big\langle (T^*T)v, v \big\rangle = \langle T^*(Tv), v\rangle = \langle Tv, Tv\rangle = \|Tv\|^2 \geq 0.$$

Thus T^*T is a positive operator.

(b) First suppose $v \in \operatorname{null} T^*T$. Then

$$\|Tv\|^2 = \langle Tv, Tv\rangle = \big\langle T^*Tv, v\big\rangle = \langle 0, v\rangle = 0.$$

Thus $Tv = 0$, proving that $\operatorname{null} T^*T \subseteq \operatorname{null} T$.

The inclusion in the other direction is clear, because if $v \in V$ and $Tv = 0$, then $T^*Tv = 0$.

Thus $\operatorname{null} T^*T = \operatorname{null} T$, completing the proof of (b).

(c) We already know from (a) that T^*T is self-adjoint. Thus

$$\operatorname{range} T^*T = (\operatorname{null} T^*T)^\perp = (\operatorname{null} T)^\perp = \operatorname{range} T^*,$$

where the first and last equalities come from 7.6 and the second equality comes from (b).

(d) To verify the first equation in (d), note that

$$\dim \operatorname{range} T = \dim(\operatorname{null} T^*)^\perp = \dim W - \dim \operatorname{null} T^* = \dim \operatorname{range} T^*,$$

where the first equality comes from 7.6(d), the second equality comes from 6.51, and the last equality comes from the fundamental theorem of linear maps (3.21).

The equality $\dim \operatorname{range} T^* = \dim \operatorname{range} T^*T$ follows from (c). ∎

7ל פירוק לערכים סינגולריים

ערכים סינגולריים

אנו נזדקק למסקנה הבאה ביחידה זו.

7.64 תכונות של T^*T

נניח ש־ $T \in \mathcal{L}(V,W)$. אזי

א. T^*T אופרטור חיובי מעל V

ב. $\operatorname{null} T^*T = \operatorname{null} T$

ג. $\operatorname{range} T^*T = \operatorname{range} T^*$

ד. $\dim \operatorname{range} T = \dim \operatorname{range} T^* = \dim \operatorname{range} T^*T$.

הוכחה

א. מתקיים

$$(T^*T)^* = T^*(T^*)^* = T^*T.$$

לפיכך T^*T צמוד לעצמו.

אם $v \in V$, אזי

$$\big\langle (T^*T)v, v \big\rangle = \langle T^*(Tv), v\rangle = \langle Tv, Tv\rangle = \|Tv\|^2 \geq 0.$$

לפיכך T^*T אופרטור חיובי.

ב. תחילה נניח כי $v \in \operatorname{null} T^*T$. אזי

$$\|Tv\|^2 = \langle Tv, Tv\rangle = \langle T^*Tv, v\rangle = \langle 0, v\rangle = 0.$$

לפיכך $Tv = 0$, ובכך הוכחנו כי $\operatorname{null} T^*T \subseteq \operatorname{null} T$.

ההכלה בכיוון השני היא ברורה, מאחר שאם $v \in V$ ו־ $Tv = 0$, אזי $T^*Tv = 0$.

לפיכך $\operatorname{null} T^*T = \operatorname{null} T$, ובכך השלמנו את ההוכחה של (ב).

ג. אנו כבר יודעים מ־ (א) כי T^*T צמוד לעצמו. לפיכך

$$\operatorname{range} T^*T = (\operatorname{null} T^*T)^\perp = (\operatorname{null} T)^\perp = \operatorname{range} T^*,$$

כאשר השוויון הראשון והאחרון נובעים מ־ 7.6 והשוויון השני נובע מ־ (ב).

ד. כדי לוודא את המשוואה הראשונה ב־ (ד), נשים לב כי

$$\dim \operatorname{range} T = \dim (\operatorname{null} T^*)^\perp = \dim W - \dim \operatorname{null} T^* = \dim \operatorname{range} T^*,$$

כאשר השוויון הראשון נובע מ־ 7.6(ד), השוויון השני נובע מ־ 6.51, והשוויון האחרון נובע מהמשפט היסודי של ההעתקות הליניאריות (3.21).

השוויון $\dim \operatorname{range} T^* = \dim \operatorname{range} T^*T$ נובע מ־ (ג).

The eigenvalues of an operator tell us something about the behavior of the operator. Another collection of numbers, called the singular values, is also useful. Eigenspaces and the notation E (used in the examples) were defined in 5.52.

7.65 definition: *singular values*

Suppose $T \in \mathcal{L}(V, W)$. The *singular values* of T are the nonnegative square roots of the eigenvalues of T^*T, listed in decreasing order, each included as many times as the dimension of the corresponding eigenspace of T^*T.

7.66 example: *singular values of an operator on* $\mathbf{F}^4$

Define $T \in \mathcal{L}(\mathbf{F}^4)$ by $T(z_1, z_2, z_3, z_4) = (0, 3z_1, 2z_2, -3z_4)$. A calculation shows that

$$T^*T(z_1, z_2, z_3, z_4) = (9z_1, 4z_2, 0, 9z_4),$$

as you should verify. Thus the standard basis of $\mathbf{F}^4$ diagonalizes T^*T, and we see that the eigenvalues of T^*T are 9, 4, and 0. Also, the dimensions of the eigenspaces corresponding to the eigenvalues are

$$\dim E(9, T^*T) = 2 \quad \text{and} \quad \dim E(4, T^*T) = 1 \quad \text{and} \quad \dim E(0, T^*T) = 1.$$

Taking nonnegative square roots of these eigenvalues of T^*T and using dimension information from above, we conclude that the singular values of T are $3, 3, 2, 0$.

The only eigenvalues of T are -3 and 0. Thus in this case, the collection of eigenvalues did not pick up the number 2 that appears in the definition (and hence the behavior) of T, but the list of singular values does include 2.

7.67 example: *singular values of a linear map from* $\mathbf{F}^4$ *to* $\mathbf{F}^3$

Suppose $T \in \mathcal{L}(\mathbf{F}^4, \mathbf{F}^3)$ has matrix (with respect to the standard bases)

$$\begin{pmatrix} 0 & 0 & 0 & -5 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix}.$$

You can verify that the matrix of T^*T is

$$\begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 25 \end{pmatrix}$$

and that the eigenvalues of the operator T^*T are 25, 2, 0, with $\dim E(25, T^*T) = 1$, $\dim E(2, T^*T) = 1$, and $\dim E(0, T^*T) = 2$. Thus the singular values of T are $5, \sqrt{2}, 0, 0$.

See Exercise 2 for a characterization of the positive singular values.

הערכים העצמיים של אופרטור מלמדים אותנו על התנהגות האופרטור. אוסף אחר של מספרים, הנקרא הערכים הסינגולריים, גם הוא שימושי. מרחבים עצמיים והסימון E (שהם בשימוש בדוגמאות להלן) הוגדרו ב־5.52.

7.65 הגדרה: ערכים סינגולריים

נניח ש־ $T \in \mathcal{L}(V,W)$. הערכים הסינגולריים של T הם השורשים הריבועיים האי־שליליים של הערכים העצמיים של T^*T, מסודרים בסדר יורד, כשכל אחד נכלל מספר פעמים השווה לממד המרחב העצמי המתאים של T^*T.

7.66 דוגמה: ערכים סינגולריים של אופרטור מעל $\mathbf{F}^4$

נגדיר $T \in \mathcal{L}(\mathbf{F}^4)$ על ידי $T(z_1,z_2,z_3,z_4) = (0,3z_1,2z_2,-3z_4)$. חישוב מראה ש־

$$T^*T(z_1,z_2,z_3,z_4) = (9z_1,4z_2,0,9z_4),$$

כפי שעליכם לוודא. לפיכך הבסיס הסטנדרטי של $\mathbf{F}^4$ מלכסן את T^*T, ואנו רואים כי הערכים העצמיים של T^*T הם 9, 4, ו־ 0. בנוסף, הממדים של המרחבים העצמיים השייכים לערכים העצמיים הם

$$\dim E(9,T^*T) = 2, \dim E(4,T^*T) = 1, \dim E(0,T^*T) = 1.$$

בהוצאת שורשים ריבועיים אי־שליליים מערכים עצמיים אלו של T^*T ובצירוף המידע על הממדים לעיל, אנו מסיקים שהערכים הסינגולריים של T הם $3, 3, 2, 0$.

הערכים העצמיים של T הם -3 ו־ 0 בלבד. לפיכך במקרה זה, אוסף הערכים העצמיים לא כלל את המספר 2 המופיע בהגדרה (ולכן ההתנהגות) של T, אך רשימת הערכים הסינגולריים כן כוללת את 2.

7.67 דוגמה: ערכים סינגולריים של העתקה ליניארית מ־ $\mathbf{F}^4$ ל־ $\mathbf{F}^3$

נניח שיש ל־ $T \in \mathcal{L}(\mathbf{F}^4,\mathbf{F}^3)$ מטריצה (ביחס לבסיסים הסטנדרטיים)

$$\begin{pmatrix} 0 & 0 & 0 & -5 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix}.$$

באפשרותכם לוודא כי המטריצה של T^*T היא

$$\begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 25 \end{pmatrix}$$

ובנוסף כי הערכים העצמיים של האופרטור T^*T הם $25, 2, 0$, עם $\dim E(25,T^*T) = 1$, $\dim E(2,T^*T) = 1$, ו־ $\dim E(0,T^*T) = 2$. לפיכך הערכים הסינגולריים של T הם $5, \sqrt{2}, 0, 0$.

ראו תרגיל 2 עבור אפיון של הערכים הסינגולריים החיוביים.

7.68 *role of positive singular values*

Suppose that $T \in \mathcal{L}(V, W)$. Then

(a) T is injective $\iff$ 0 is not a singular value of T;

(b) the number of positive singular values of T equals $\dim \operatorname{range} T$;

(c) T is surjective $\iff$ number of positive singular values of T equals $\dim W$.

Proof The linear map T is injective if and only if $\operatorname{null} T = \{0\}$, which happens if and only if $\operatorname{null} T^*T = \{0\}$ [by 7.64(b)], which happens if and only if 0 is not an eigenvalue of T^*T, which happens if and only if 0 is not a singular value of T, completing the proof of (a).

The spectral theorem applied to T^*T shows that $\dim \operatorname{range} T^*T$ equals the number of positive eigenvalues of T^*T (counting repetitions). Thus 7.64(d) implies that $\dim \operatorname{range} T$ equals the number of positive singular values of T, proving (b).

Use (b) and 2.39 to show that (c) holds. ∎

The table below compares eigenvalues with singular values.

list of eigenvalues	**list of singular values**
context: vector spaces	context: inner product spaces
defined only for linear maps from a vector space to itself	defined for linear maps from an inner product space to a possibly different inner product space
can be arbitrary real numbers (if $\mathbf{F} = \mathbf{R}$) or complex numbers (if $\mathbf{F} = \mathbf{C}$)	are nonnegative numbers
can be the empty list if $\mathbf{F} = \mathbf{R}$	length of list equals dimension of domain
includes 0 $\iff$ operator is not invertible	includes 0 $\iff$ linear map is not injective
no standard order, especially if $\mathbf{F} = \mathbf{C}$	always listed in decreasing order

The next result nicely characterizes isometries in terms of singular values.

7.69 *isometries characterized by having all singular values equal* 1

Suppose that $S \in \mathcal{L}(V, W)$. Then

$$S \text{ is an isometry} \iff \text{all singular values of } S \text{ equal } 1.$$

Proof We have

$$\begin{aligned} S \text{ is an isometry} &\iff S^*S = I \\ &\iff \text{all eigenvalues of } S^*S \text{ equal } 1 \\ &\iff \text{all singular values of } S \text{ equal } 1, \end{aligned}$$

where the first equivalence comes from 7.49 and the second equivalence comes from the spectral theorem (7.29 or 7.31) applied to the self-adjoint operator S^*S. ∎

7.68 תפקיד ערכים סינגולריים חיוביים

נניח ש־ $T \in \mathcal{L}(V,W)$. אזי

א. T חד־חד־ערכית $\iff$ 0 אינו ערך סינגולרי של T

ב. מספר הערכים הסינגולריים החיוביים של T שווה ל־ $\dim \operatorname{range} T$

ג. T עַל $\iff$ מספר הערכים הסינגולריים החיוביים של T שווה ל־ $\dim W$.

הוכחה ההעתקה הליניארית T היא חד־חד־ערכית אם־ורק־אם $\operatorname{null} T = \{0\}$, שהדבר קורה אם־ורק־אם $\operatorname{null} T^*T = \{0\}$ [לפי 7.64(ב)], שהדבר קורה אם־ורק־אם 0 אינו ערך עצמי של T^*T, שהדבר קורה אם־ורק־אם 0 אינו ערך סינגולרי של T, ובכך הושלמה הוכחת (א).

משפט הפירוק הספקטרלי המיושם על T^*T מראה לנו ש־ $\dim \operatorname{range} T^*T$ שווה למספר הערכים העצמיים החיוביים של T^*T (כשאנו סופרים חזרות). לפיכך 7.64(ד) גורר ש־ $\dim \operatorname{range} T$ שווה למספר הערכים הסינגולריים החיוביים של T, מה שמוכיח את (ב).

השתמשו ב־ (ב) וב־ 2.39 כדי להראות כי (ג) מתקיים. ■

הטבלה להלן משווה ערכים עצמיים עם ערכים סינגולריים.

רשימת ערכים עצמיים	רשימת ערכים סינגולריים
הקשר: מרחבים וקטוריים	הקשר: מרחבי מכפלה פנימית
מוגדרת רק עבור העתקות ליניאריות ממרחב וקטורי לעצמו	מוגדרת עבור העתקות ליניאריות ממרחב מכפלה פנימית למרחב מכפלה פנימית זהה או שונה
מספרים שרירותיים ממשיים (כאשר $\mathbf{F} = \mathbf{R}$) או מרוכבים (כאשר $\mathbf{F} = \mathbf{C}$)	מספרים אי־שליליים
יכולה להיות רשימה ריקה אם $\mathbf{F} = \mathbf{R}$	אורך הרשימה שווה לממד התחום
כוללת 0 אם־ורק־אם האופרטור אינו הפיך	כוללת 0 אם־ורק־אם ההעתקה אינה חד־חד־ערכית
ללא סדר מסוים, בפרט אם $\mathbf{F} = \mathbf{C}$	תמיד מופיעים בסדר יורד

המסקנה הבאה מאפיינת באופן נחמד איזומטריות במונחים של ערכים סינגולריים.

7.69 איזומטריות מאופיינות בכך שכל הערכים הסינגולריים שווים 1

נניח ש־ $S \in \mathcal{L}(V,W)$. אזי

S איזומטריה $\iff$ כל הערכים הסינגולריים של S שווים 1.

הוכחה מתקיים

$$\begin{aligned} S \text{ איזומטריה} &\iff S^*S = I \\ &\iff \text{כל הערכים העצמיים של } S^*S \text{ שווים } 1 \\ &\iff \text{כל הערכים הסינגולריים של } S \text{ שווים } 1, \end{aligned}$$

כאשר השקילות הראשונה נובעת מ־ 7.49 והשקילות השנייה נובעת ממשפט הפירוק הספקטרלי (7.29 או 7.31) כשהוא מיושם על האופרטור הצמוד לעצמו S^*S. ■

SVD for Linear Maps and for Matrices

The next result shows that every linear map from V to W has a remarkably clean description in terms of its singular values and orthonormal lists in V and W. In the next section we will see several important applications of the singular value decomposition (often called the SVD).

The singular value decomposition is useful in computational linear algebra because good techniques exist for approximating eigenvalues and eigenvectors of positive operators such as T^*T, whose eigenvalues and eigenvectors lead to the singular value decomposition.

7.70 *singular value decomposition*

Suppose $T \in \mathcal{L}(V, W)$ and the positive singular values of T are $s_1, \dots, s_m$. Then there exist orthonormal lists $e_1, \dots, e_m$ in V and $f_1, \dots, f_m$ in W such that

7.71 $$Tv = s_1\langle v, e_1\rangle f_1 + \cdots + s_m\langle v, e_m\rangle f_m$$

for every $v \in V$.

Proof Let $s_1, \dots, s_n$ denote the singular values of T (thus $n = \dim V$). Because T^*T is a positive operator [see 7.64(a)], the spectral theorem implies that there exists an orthonormal basis $e_1, \dots, e_n$ of V with

7.72 $$T^*Te_k = s_k^2 e_k$$

for each $k = 1, \dots, n$.

For each $k = 1, \dots, m$, let

7.73 $$f_k = \frac{Te_k}{s_k}.$$

If $j, k \in \{1, \dots, m\}$, then

$$\langle f_j, f_k\rangle = \frac{1}{s_j s_k}\langle Te_j, Te_k\rangle = \frac{1}{s_j s_k}\langle e_j, T^*Te_k\rangle = \frac{s_k}{s_j}\langle e_j, e_k\rangle = \begin{cases} 0 & \text{if } j \neq k, \\ 1 & \text{if } j = k. \end{cases}$$

Thus $f_1, \dots, f_m$ is an orthonormal list in W.

If $k \in \{1, \dots, n\}$ and $k > m$, then $s_k = 0$ and hence $T^*Te_k = 0$ (by 7.72), which implies that $Te_k = 0$ [by 7.64(b)].

Suppose $v \in V$. Then

$$\begin{aligned} Tv &= T(\langle v, e_1\rangle e_1 + \cdots + \langle v, e_n\rangle e_n) \\ &= \langle v, e_1\rangle Te_1 + \cdots + \langle v, e_m\rangle Te_m \\ &= s_1\langle v, e_1\rangle f_1 + \cdots + s_m\langle v, e_m\rangle f_m, \end{aligned}$$

where the last index in the first line switched from n to m in the second line because $Te_k = 0$ if $k > m$ (as noted in the paragraph above) and the third line follows from 7.73. The equation above is our desired result. ∎

פירוק לערכים סינגולריים עבור העתקות ליניאריות ומטריצות

הפירוק לערכים סינגולריים שימושי באלגברה ליניארית חישובית מאחר שקיימים טכניקות יעילות לחישוב קירובים עבור ערכים עצמיים ווקטורים עצמיים של אופרטורים חיוביים דוגמת T^*T, שהערכים העצמיים והווקטורים העצמיים שלו מובילים לפירוק לערכים סינגולריים.

המסקנה הבאה מראה שלכל העתקה ליניארית מ־ V אל W יש תיאור נקי להפליא במונחים של הערכים הסינגולריים שלה ורשימות אורתונורמליות ב־ V ו־ W. ביחידה הבאה אנו נראה מספר יישומים חשובים של הפירוק לערכים סינגולריים (לעיתים קרובות נקרא SVD).

7.70 פירוק לערכים סינגולריים

נניח ש־ $T \in \mathcal{L}(V,W)$ והערכים הסינגולריים החיוביים של T הם $s_1, \dots, s_m$. אזי קיימות רשימות אורתונורמליות $e_1, \dots, e_m$ ב־ V ו־ $f_1, \dots, f_m$ ב־ W כך ש־

$$Tv = s_1\langle v, e_1\rangle f_1 + \cdots + s_m\langle v, e_m\rangle f_m \tag{7.71}$$

לכל $v \in V$.

הוכחה יהיו $s_1, \dots, s_n$ הערכים הסינגולריים של T (לפיכך $n = \dim V$). מאחר ש־ T^*T אופרטור חיובי [ראו 7.64(א)], משפט הפירוק הספקטרלי גורר שקיים בסיס אורתונורמלי $e_1, \dots, e_n$ של V עם

$$T^*Te_k = s_k^2 e_k \tag{7.72}$$

לכל $k = 1, \dots, n$.

לכל $k = 1, \dots, m$, יהי

$$f_k = \frac{Te_k}{s_k}. \tag{7.73}$$

אם $j, k \in \{1, \dots, m\}$, אזי

$$\langle f_j, f_k\rangle = \frac{1}{s_j s_k}\langle Te_j, Te_k\rangle = \frac{1}{s_j s_k}\langle e_j, T^*Te_k\rangle = \frac{s_k}{s_j}\langle e_j, e_k\rangle = \begin{cases} 0 & \text{if } j \neq k, \\ 1 & \text{if } j = k. \end{cases}$$

לפיכך $f_1, \dots, f_m$ רשימה אורתונורמלית ב־ W.

אם $k \in \{1, \dots, n\}$ ו־ $k > m$, אזי $s_k = 0$ ולכן $T^*Te_k = 0$ (לפי 7.72), מה שגורר ש־ $Te_k = 0$ [לפי 7.64(ב)].

נניח ש־ $v \in V$. אזי

$$\begin{aligned} Tv &= T(\langle v, e_1\rangle e_1 + \cdots + \langle v, e_n\rangle e_n) \\ &= \langle v, e_1\rangle Te_1 + \cdots + \langle v, e_m\rangle Te_m \\ &= s_1\langle v, e_1\rangle f_1 + \cdots + s_m\langle v, e_m\rangle f_m, \end{aligned}$$

כאשר האינדקס האחרון בשורה הראשונה מוחלף מ־ n ל־ m בשורה השנייה מאחר ש־ $Te_k = 0$ אם $k > m$ (כפי שצוין בפסקה לעיל) והשורה השלישית נובעת מ־ 7.73. המשוואה לעיל היא התוצאה שביקשנו. ∎

Suppose $T \in \mathcal{L}(V, W)$, the positive singular values of T are $s_1, \dots, s_m$, and $e_1, \dots, e_m$ and $f_1, \dots, f_m$ are as in the singular value decomposition 7.70. The orthonormal list $e_1, \dots, e_m$ can be extended to an orthonormal basis $e_1, \dots, e_{\dim V}$ of V and the orthonormal list $f_1, \dots, f_m$ can be extended to an orthonormal basis $f_1, \dots, f_{\dim W}$ of W. The formula 7.71 shows that

$$Te_k = \begin{cases} s_k f_k & \text{if } 1 \le k \le m, \\ 0 & \text{if } m < k \le \dim V. \end{cases}$$

Thus the matrix of T with respect to the orthonormal bases $(e_1, \dots, e_{\dim V})$ and $(f_1, \dots, f_{\dim W})$ has the simple form

$$\mathcal{M}(T, (e_1, \dots, e_{\dim V}), (f_1, \dots, f_{\dim W}))_{j,k} = \begin{cases} s_k & \text{if } 1 \le j = k \le m, \\ 0 & \text{otherwise.} \end{cases}$$

If $\dim V = \dim W$ (as happens, for example, if $W = V$), then the matrix described in the paragraph above is a diagonal matrix. If we extend the definition of diagonal matrix as follows to apply to matrices that are not necessarily square, then we have proved the wonderful result that every linear map from V to W has a diagonal matrix with respect to appropriate orthonormal bases.

7.74 definition: *diagonal matrix*

An M-by-N matrix A is called a *diagonal matrix* if all entries of the matrix are 0 except possibly $A_{k,k}$ for $k = 1, \dots, \min\{M, N\}$.

The table below compares the spectral theorem (7.29 and 7.31) with the singular value decomposition (7.70).

spectral theorem	singular value decomposition
describes only self-adjoint operators (when $\mathbf{F} = \mathbf{R}$) or normal operators (when $\mathbf{F} = \mathbf{C}$)	describes arbitrary linear maps from an inner product space to a possibly different inner product space
produces a single orthonormal basis	produces two orthonormal lists, one for domain space and one for range space, that are not necessarily the same even when range space equals domain space
different proofs depending on whether $\mathbf{F} = \mathbf{R}$ or $\mathbf{F} = \mathbf{C}$	same proof works regardless of whether $\mathbf{F} = \mathbf{R}$ or $\mathbf{F} = \mathbf{C}$

The singular value decomposition gives us a new way to understand the adjoint and the inverse of a linear map. Specifically, the next result shows that given a singular value decomposition of a linear map $T \in \mathcal{L}(V, W)$, we can obtain the adjoint of T simply by interchanging the roles of the e's and the f's (see 7.77). Similarly, we can obtain the pseudoinverse $T^\dagger$ (see 6.68) of T by interchanging the roles of the e's and the f's and replacing each positive singular value s_k of T with $1/s_k$ (see 7.78).

נניח ש־ $T \in \mathcal{L}(V,W)$, הערכים הסינגולריים החיוביים של T הם $s_1, \dots, s_m$, ו־ $e_1, \dots, e_m$ ו־ $f_1, \dots, f_m$ הם כמו בפירוק לערכים סינגולריים 7.70. את הרשימה האורתונורמלית $e_1, \dots, e_m$ ניתן להשלים לכדי בסיס אורתונורמלי של V ואת הרשימה האורתונורמלית $f_1, \dots, f_m$ ניתן להשלים לכדי בסיס אורתונורמלי של W. הנוסחה 7.71 מראה ש־

$$Te_k = \begin{cases} s_k f_k & \text{if } 1 \le k \le m, \\ 0 & \text{if } m < k \le \dim V. \end{cases}$$

לפיכך מטריצת T ביחס לבסיסים האורתונורמליים $(e_1, \dots, e_{\dim V})$ ו־ $(f_1, \dots, f_{\dim W})$ היא בעלת התבנית הפשוטה

$$\mathcal{M}(T, (e_1, \dots, e_{\dim V}), (f_1, \dots, f_{\dim W}))_{j,k} = \begin{cases} s_k & \text{if } 1 \le j = k \le m, \\ 0 & \text{אחרת.} \end{cases}$$

אם $\dim V = \dim W$ (כפי שקורה, לדוגמה, אם $W = V$), אזי המטריצה שתוארה בפסקה לעיל היא מטריצה אלכסונית. אם אנו מרחיבים את ההגדרה של מטריצה אלכסונית כמובא להלן כך שתחול גם עבור מטריצות שאינן בהכרח ריבועיות, אזי הוכחנו את המסקנה הנפלאה כי לכל העתקה ליניארית מ־ V אל W יש מטריצה אלכסונית ביחס לבסיסים אורתונורמליים מתאימים.

7.74 הגדרה: מטריצה אלכסונית

מטריצה A $M \times N$ נקראת מטריצה אלכסונית אם כל איברי המטריצה הם 0 מלבד $A_{k,k}$ עבור $k = 1, \dots, \min\{M, N\}$ שבהם ייתכנו איברים השונים מ־ 0.

הטבלה להלן משווה בין משפט הפירוק הספקטרלי (7.29 ו־ 7.31) לבין הפירוק לערכים סינגולריים (7.70).

משפט הפירוק הספקטרלי	פירוק לערכים סינגולריים
מתאר רק אופרטורים הצמודים לעצמם (כאשר $\mathbf{F} = \mathbf{R}$) או אופרטורים נורמליים (כאשר $\mathbf{F} = \mathbf{C}$)	מתאר העתקות ליניאריות שרירותיות ממרחב מכפלה פנימית אל מרחב מכפלה פנימית זהה או אחר
מפיק בסיס אורתונורמלי יחיד	מפיק שתי רשימות אורתונורמליות, אחת למרחב התחום ואחת למרחב הטווח, שאינם בהכרח אותו המרחב גם כאשר מרחב הטווח שווה למרחב התחום
הוכחות שונות בהתאם לשאלה האם $\mathbf{F} = \mathbf{R}$ או $\mathbf{F} = \mathbf{C}$	אותה הוכחה עובדת בין אם $\mathbf{F} = \mathbf{R}$ ובין אם $\mathbf{F} = \mathbf{C}$

הפירוק לערכים סינגולריים מעניק לנו דרך חדשה להבנת הצמודה וההופכית של העתקה ליניארית. ליתר דיוק, המסקנה הבאה מראה כי בהינתן פירוק לערכים סינגולריים של העתקה ליניארית $T \in \mathcal{L}(V,W)$, נוכל לקבל את הצמודה של T באופן פשוט על ידי כך שנחליף בין התפקידים של ה־ e־ים וה־ f־ים (ראו 7.77). באופן דומה, נוכל לקבל את הפסידו־הופכי $T^\dagger$ (ראו 6.68) של T על ידי כך שנחליף בין התפקידים של ה־ e־ים וה־ f־ים ונחליף כל ערך סינגולרי חיובי s_k של T ב־ $1/s_k$ (ראו 7.78).

Recall that the pseudoinverse $T^\dagger$ in 7.78 below equals the inverse T^{-1} if T is invertible [see 6.69(a)].

7.75 *singular value decomposition of adjoint and pseudoinverse*

Suppose $T \in \mathcal{L}(V, W)$ and the positive singular values of T are $s_1, \dots, s_m$. Suppose $e_1, \dots, e_m$ and $f_1, \dots, f_m$ are orthonormal lists in V and W such that

$$Tv = s_1\langle v, e_1\rangle f_1 + \cdots + s_m\langle v, e_m\rangle f_m \tag{7.76}$$

for every $v \in V$. Then

$$T^*w = s_1\langle w, f_1\rangle e_1 + \cdots + s_m\langle w, f_m\rangle e_m \tag{7.77}$$

and

$$T^\dagger w = \frac{\langle w, f_1\rangle}{s_1} e_1 + \cdots + \frac{\langle w, f_m\rangle}{s_m} e_m \tag{7.78}$$

for every $w \in W$.

Proof If $v \in V$ and $w \in W$ then

$$\begin{aligned}\langle Tv, w\rangle &= \langle s_1\langle v, e_1\rangle f_1 + \cdots + s_m\langle v, e_m\rangle f_m, w\rangle \\ &= s_1\langle v, e_1\rangle\langle f_1, w\rangle + \cdots + s_m\langle v, e_m\rangle\langle f_m, w\rangle \\ &= \langle v, s_1\langle w, f_1\rangle e_1 + \cdots + s_m\langle w, f_m\rangle e_m\rangle.\end{aligned}$$

This implies that

$$T^*w = s_1\langle w, f_1\rangle e_1 + \cdots + s_m\langle w, f_m\rangle e_m,$$

proving 7.77.

To prove 7.78, suppose $w \in W$. Let

$$v = \frac{\langle w, f_1\rangle}{s_1} e_1 + \cdots + \frac{\langle w, f_m\rangle}{s_m} e_m.$$

Apply T to both sides of the equation above, getting

$$\begin{aligned}Tv &= \frac{\langle w, f_1\rangle}{s_1} Te_1 + \cdots + \frac{\langle w, f_m\rangle}{s_m} Te_m \\ &= \langle w, f_1\rangle f_1 + \cdots + \langle w, f_m\rangle f_m \\ &= P_{\text{range } T}\, w,\end{aligned}$$

where the second line holds because 7.76 implies that $Te_k = s_k f_k$ if $k = 1, \dots, m$, and the last line above holds because 7.76 implies that $f_1, \dots, f_m$ spans range T and thus is an orthonormal basis of range T [and hence 6.57(i) applies]. The equation above, the observation that $v \in (\text{null } T)^\perp$ [see Exercise 8(b)], and the definition of $T^\dagger w$ (see 6.68) show that $v = T^\dagger w$, proving 7.78. ∎

נזכיר כי הפסידו־הופכי $T^\dagger$ ב־ 7.78 להלן שווה להופכי T^{-1} אם T הפיכה [ראו 6.69(א)].

7.75 פירוק לערכים סינגולריים של צמודה ופסידו־הופכי

נניח ש־ $T \in \mathcal{L}(V,W)$ והערכים הסינגולריים החיוביים של T הם $s_1, \dots, s_m$. נניח ש־ $e_1, \dots, e_m$ ו־ $f_1, \dots, f_m$ רשימות אורתונורמליות ב־ V ו־ W כך ש־

$$Tv = s_1\langle v, e_1\rangle f_1 + \cdots + s_m\langle v, e_m\rangle f_m \tag{7.76}$$

לכל $v \in V$. אזי

$$T^*w = s_1\langle w, f_1\rangle e_1 + \cdots + s_m\langle w, f_m\rangle e_m \tag{7.77}$$

וגם

$$T^\dagger w = \frac{\langle w, f_1\rangle}{s_1} e_1 + \cdots + \frac{\langle w, f_m\rangle}{s_m} e_m \tag{7.78}$$

לכל $w \in W$.

הוכחה אם $v \in V$ ו־ $w \in W$ אזי

$$\begin{aligned}\langle Tv, w\rangle &= \langle s_1\langle v, e_1\rangle f_1 + \cdots + s_m\langle v, e_m\rangle f_m, w\rangle \\ &= s_1\langle v, e_1\rangle\langle f_1, w\rangle + \cdots + s_m\langle v, e_m\rangle\langle f_m, w\rangle \\ &= \langle v, s_1\langle w, f_1\rangle e_1 + \cdots + s_m\langle w, f_m\rangle e_m\rangle.\end{aligned}$$

מכאן נובע ש־

$$T^*w = s_1\langle w, f_1\rangle e_1 + \cdots + s_m\langle w, f_m\rangle e_m,$$

ובכך הוכחנו את 7.77.

להוכחת 7.78, נניח ש־ $w \in W$. יהי

$$v = \frac{\langle w, f_1\rangle}{s_1} e_1 + \cdots + \frac{\langle w, f_m\rangle}{s_m} e_m.$$

נפעיל את T על שני אגפי המשוואה לעיל ונקבל

$$\begin{aligned}Tv &= \frac{\langle w, f_1\rangle}{s_1} Te_1 + \cdots + \frac{\langle w, f_m\rangle}{s_m} Te_m \\ &= \langle w, f_1\rangle f_1 + \cdots + \langle w, f_m\rangle f_m \\ &= P_{\operatorname{range} T}\, w,\end{aligned}$$

כאשר השורה השנייה מתקיימת מאחר ש־ 7.76 גוררת ש־ $Te_k = s_k f_k$ אם $k = 1, \dots, m$, והשורה האחרונה לעיל מתקיימת מאחר ש־ 7.76 גוררת ש־ $f_1, \dots, f_m$ פורסת את $\operatorname{range} T$ ולפיכך בסיס אורתונורמלי של $\operatorname{range} T$ [ולכן 6.57(ט) מתקיים]. המשוואה לעיל, ההבחנה כי $v \in (\operatorname{null} T)^\perp$ [ראו תרגיל 8(ב)], וההגדרה של $T^\dagger w$ (ראו 6.68) מראות כי $v = T^\dagger w$, ובכך הוכחנו את 7.78. ■

7.79 example: *finding a singular value decomposition*

Define $T \in \mathcal{L}(\mathbf{F}^4, \mathbf{F}^3)$ by $T(x_1, x_2, x_3, x_4) = (-5x_4, 0, x_1 + x_2)$. We want to find a singular value decomposition of T. The matrix of T (with respect to the standard bases) is

$$\begin{pmatrix} 0 & 0 & 0 & -5 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix}.$$

Thus, as discussed in Example 7.67, the matrix of T^*T is

$$\begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 25 \end{pmatrix},$$

and the positive eigenvalues of T^*T are 25, 2, with $\dim E(25, T^*T) = 1$ and $\dim E(2, T^*T) = 1$. Hence the positive singular values of T are $5, \sqrt{2}$.

Thus to find a singular value decomposition of T, we must find an orthonormal list e_1, e_2 in $\mathbf{F}^4$ and an orthonormal list f_1, f_2 in $\mathbf{F}^3$ such that

$$Tv = 5\langle v, e_1\rangle f_1 + \sqrt{2}\langle v, e_2\rangle f_2$$

for all $v \in \mathbf{F}^4$.

An orthonormal basis of $E(25, T^*T)$ is the vector $(0, 0, 0, 1)$; an orthonormal basis of $E(2, T^*T)$ is the vector $\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0, 0\right)$. Thus, following the proof of 7.70, we take

$$e_1 = (0, 0, 0, 1) \quad \text{and} \quad e_2 = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0, 0\right)$$

and

$$f_1 = \frac{Te_1}{5} = (-1, 0, 0) \quad \text{and} \quad f_2 = \frac{Te_2}{\sqrt{2}} = (0, 0, 1).$$

Then, as expected, we see that e_1, e_2 is an orthonormal list in $\mathbf{F}^4$ and f_1, f_2 is an orthonormal list in $\mathbf{F}^3$ and

$$Tv = 5\langle v, e_1\rangle f_1 + \sqrt{2}\langle v, e_2\rangle f_2$$

for all $v \in \mathbf{F}^4$. Thus we have found a singular value decomposition of T.

The next result translates the singular value decomposition from the context of linear maps to the context of matrices. Specifically, the following result gives a factorization of an arbitrary matrix as the product of three nice matrices. The proof gives an explicit construction of these three matrices in terms of the singular value decomposition.

In the next result, the phrase "orthonormal columns" should be interpreted to mean that the columns are orthonormal with respect to the standard Euclidean inner product.

7.79 דוגמה: מציאת הפירוק לערכים סינגולריים

נגדיר $T \in \mathcal{L}(\mathbf{F}^4, \mathbf{F}^3)$ על ידי $T(x_1, x_2, x_3, x_4) = (-5x_4, 0, x_1 + x_2)$. אנו מעוניינים למצוא פירוק לערכים סינגולריים של T. המטריצה של T (ביחס אל הבסיסים הסטנדרטיים) היא

$$\begin{pmatrix} 0 & 0 & 0 & -5 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix}.$$

לפיכך, כפי שציינו בדוגמה 7.67, המטריצה של T^*T היא

$$\begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 25 \end{pmatrix},$$

והערכים העצמיים החיוביים של T^*T הם 25, 2, עם $\dim E(25, T^*T) = 1$ ו־ $\dim E(2, T^*T) = 1$. לכן הערכים הסינגולריים החיוביים של T הם $5, \sqrt{2}$.

לכן כדי למצוא פירוק לערכים סינגולריים של T, עלינו למצוא רשימה אורתונורמלית e_1, e_2 ב־ $\mathbf{F}^4$ ורשימה אורתונורמלית f_1, f_2 ב־ $\mathbf{F}^3$ כך ש־

$$Tv = 5\langle v, e_1\rangle f_1 + \sqrt{2}\langle v, e_2\rangle f_2$$

לכל $v \in \mathbf{F}^4$.

בסיס אורתונורמלי של $E(25, T^*T)$ הוא הווקטור $(0, 0, 0, 1)$; בסיס אורתונורמלי של $E(2, T^*T)$ הוא הווקטור $\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0, 0\right)$. לפיכך, בעקבות ההוכחה של 7.70, ניקח

$$e_1 = (0, 0, 0, 1) \quad \text{ו־} \quad e_2 = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0, 0\right)$$

וגם

$$f_1 = \frac{Te_1}{5} = (-1, 0, 0) \quad \text{ו־} \quad f_2 = \frac{Te_2}{\sqrt{2}} = (0, 0, 1).$$

אזי, כצפוי, אנו רואים כי e_1, e_2 רשימה אורתונורמלית ב־ $\mathbf{F}^4$ ו־ f_1, f_2 רשימה אורתונורמלית ב־ $\mathbf{F}^3$ ו־

$$Tv = 5\langle v, e_1\rangle f_1 + \sqrt{2}\langle v, e_2\rangle f_2$$

לכל $v \in \mathbf{F}^4$. לפיכך מצאנו פירוק לערכים סינגולריים של T.

המסקנה הבאה מתרגמת את הפירוק לערכים סינגולריים מההקשר של העתקות ליניאריות להקשר של מטריצות. ליתר דיוק, המסקנה הבאה מציגה פירוק לגורמים של מטריצה שרירותית כמכפלה של שלוש מטריצות נחמדות. ההוכחה מספקת בנייה מפורשת של שלוש המטריצות הללו במונחים של פירוק לערכים סינגולריים.

במסקנה הבאה, את הביטוי "עמודות אורתונורמליות" יש להבין כך: העמודות אורתונורמליות ביחס אל המכפלה הפנימית האוקלידית הסטנדרטית.

7.80 *matrix version of SVD*

Suppose A is a p-by-n matrix of rank $m \geq 1$. Then there exist a p-by-m matrix B with orthonormal columns, an m-by-m diagonal matrix D with positive numbers on the diagonal, and an n-by-m matrix C with orthonormal columns such that

$$A = BDC^*.$$

Proof Let $T : \mathbf{F}^n \to \mathbf{F}^p$ be the linear map whose matrix with respect to the standard bases equals A. Then $\dim \operatorname{range} T = m$ (by 3.78). Let

7.81 $$Tv = s_1\langle v, e_1\rangle f_1 + \cdots + s_m\langle v, e_m\rangle f_m$$

be a singular value decomposition of T. Let

B = the p-by-m matrix whose columns are $f_1, \ldots, f_m$,
D = the m-by-m diagonal matrix whose diagonal entries are $s_1, \ldots, s_m$,
C = the n-by-m matrix whose columns are $e_1, \ldots, e_m$.

Let $u_1, \ldots, u_m$ denote the standard basis of $\mathbf{F}^m$. If $k \in \{1, \ldots, m\}$ then

$$(AC - BD)u_k = Ae_k - B(s_k u_k) = s_k f_k - s_k f_k = 0.$$

Thus $AC = BD$.

Multiply both sides of this last equation by C^* (the conjugate transpose of C) on the right to get

$$ACC^* = BDC^*.$$

Note that the rows of C^* are the complex conjugates of $e_1, \ldots, e_m$. Thus if $k \in \{1, \ldots, m\}$, then the definition of matrix multiplication shows that $C^* e_k = u_k$; hence $CC^* e_k = e_k$. Thus $ACC^* v = Av$ for all $v \in \operatorname{span}(e_1, \ldots, e_m)$.

If $v \in (\operatorname{span}(e_1, \ldots, e_m))^\perp$, then $Av = 0$ (as follows from 7.81) and $C^* v = 0$ (as follows from the definition of matrix multiplication). Hence $ACC^* v = Av$ for all $v \in (\operatorname{span}(e_1, \ldots, e_m))^\perp$.

Because ACC^* and A agree on $\operatorname{span}(e_1, \ldots, e_m)$ and on $(\operatorname{span}(e_1, \ldots, e_m))^\perp$, we conclude that $ACC^* = A$. Thus the displayed equation above becomes

$$A = BDC^*,$$

as desired. ∎

Note that the matrix A in the result above has pn entries. In comparison, the matrices B, D, and C above have a total of

$$m(p + m + n)$$

entries. Thus if p and n are large numbers and the rank m is considerably less than p and n, then the number of entries that must be stored on a computer to represent A is considerably less than pn.

7.80 גרסת המטריצה של הפירוק לערכים סינגולריים

נניח ש־ A מטריצה $p \times n$ שדרגתה $m \geq 1$. אזי קיימות מטריצה B $p \times m$ עם עמודות אורתונורמליות, מטריצה אלכסונית D $m \times m$ עם מספרים חיוביים על האלכסון, ו־ C מטריצה $n \times m$ עם עמודות אורתונורמליות כך ש־

$$A = BDC^*.$$

הוכחה תהי $T: \mathbf{F}^n \to \mathbf{F}^p$ ההעתקה הליניארית שהמטריצה שלה ביחס לבסיסים הסטנדרטיים שווה ל־ A. אזי $\dim \operatorname{range} T = m$ (לפי 3.78). יהי

$$Tv = s_1\langle v, e_1\rangle f_1 + \cdots + s_m\langle v, e_m\rangle f_m \tag{7.81}$$

הפירוק לערכים סינגולריים של T. תהיינה

B המטריצה $p \times m$ שעמודותיה הן $f_1, \ldots, f_m$,
D המטריצה האלכסונית $m \times m$ שאיברי האלכסון שלה הם $s_1, \ldots, s_m$,
C המטריצה $n \times m$ שעמודותיה הן $e_1, \ldots, e_m$.

יהיו $u_1, \ldots, u_m$ הבסיס הסטנדרטי של $\mathbf{F}^m$. אם $k \in \{1, \ldots, m\}$ אזי

$$(AC - BD)u_k = Ae_k - B(s_k u_k) = s_k f_k - s_k f_k = 0.$$

לפיכך $AC = BD$.

נכפיל את שני אגפי משוואה זו ב־ C^* (הצמודה־משוחלפת של C) מימין כדי לקבל

$$ACC^* = BDC^*.$$

נשים לב שהשורות של C^* הן הצמודים המרוכבים של $e_1, \ldots, e_m$. לפיכך אם $k \in \{1, \ldots, m\}$, אזי מההגדרה של כפל מטריצות נקבל ש־ $C^*e_k = u_k$; לכן $CC^*e_k = e_k$. לפיכך $ACC^*v = Av$ לכל $v \in \operatorname{span}(e_1, \ldots, e_m)$.

אם $v \in (\operatorname{span}(e_1, \ldots, e_m))^\perp$, אזי $Av = 0$ (כפי שנובע מ־ 7.81) וגם $C^*v = 0$ (כפי שנובע מהגדרת כפל מטריצות). לכן $ACC^*v = Av$ לכל $v \in (\operatorname{span}(e_1, \ldots, e_m))^\perp$.

מאחר ש־ ACC^* ו־ A מתלכדות על $\operatorname{span}(e_1, \ldots, e_m)$ ועל $(\operatorname{span}(e_1, \ldots, e_m))^\perp$, אנו מסיקים ש־ $ACC^* = A$. לפיכך המשוואה שהוצגה לעיל הופכת ל־

$$A = BDC^*,$$

כפי שרצינו. ■

שימו לב שלמטריצה A במסקנה לעיל יש pn איברים. בהשוואה, למטריצות B, D ו־ C לעיל יש יחדיו

$$m(p + m + n)$$

איברים. לפיכך אם p ו־ n הם מספרים גדולים והדרגה m קטנה באופן ניכר מ־ p ו־ n, אזי מספר האיברים שעלינו לאחסן במחשב להצגת A קטן באופן ניכר מ־ pn.

Exercises 7E

1 Suppose $T \in \mathcal{L}(V, W)$. Show that $T = 0$ if and only if all singular values of T are 0.

2 Suppose $T \in \mathcal{L}(V, W)$ and $s > 0$. Prove that s is a singular value of T if and only if there exist nonzero vectors $v \in V$ and $w \in W$ such that

$$Tv = sw \quad \text{and} \quad T^*w = sv.$$

The vectors v, w satisfying both equations above are called a ***Schmidt pair****. Erhard Schmidt introduced the concept of singular values in 1907.*

3 Give an example of $T \in \mathcal{L}(\mathbf{C}^2)$ such that 0 is the only eigenvalue of T and the singular values of T are $5, 0$.

4 Suppose that $T \in \mathcal{L}(V, W)$, s_1 is the largest singular value of T, and s_n is the smallest singular value of T. Prove that

$$\{\|Tv\| : v \in V \text{ and } \|v\| = 1\} = [s_n, s_1].$$

5 Suppose $T \in \mathcal{L}(\mathbf{C}^2)$ is defined by $T(x, y) = (-4y, x)$. Find the singular values of T.

6 Find the singular values of the differentiation operator $D \in \mathcal{L}(\mathcal{P}_2(\mathbf{R}))$ defined by $Dp = p'$, where the inner product on $\mathcal{P}_2(\mathbf{R})$ is as in Example 6.34.

7 Suppose that $T \in \mathcal{L}(V)$ is self-adjoint or that $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$ is normal. Let $\lambda_1, \dots, \lambda_n$ be the eigenvalues of T, each included in this list as many times as the dimension of the corresponding eigenspace. Show that the singular values of T are $|\lambda_1|, \dots, |\lambda_n|$, after these numbers have been sorted into decreasing order.

8 Suppose $T \in \mathcal{L}(V, W)$. Suppose $s_1 \ge s_2 \ge \cdots \ge s_m > 0$ and $e_1, \dots, e_m$ is an orthonormal list in V and $f_1, \dots, f_m$ is an orthonormal list in W such that

$$Tv = s_1\langle v, e_1\rangle f_1 + \cdots + s_m\langle v, e_m\rangle f_m$$

for every $v \in V$.

(a) Prove that $f_1, \dots, f_m$ is an orthonormal basis of range T.
(b) Prove that $e_1, \dots, e_m$ is an orthonormal basis of $(\operatorname{null} T)^\perp$.
(c) Prove that $s_1, \dots, s_m$ are the positive singular values of T.
(d) Prove that if $k \in \{1, \dots, m\}$, then e_k is an eigenvector of T^*T with corresponding eigenvalue s_k^2.
(e) Prove that

$$TT^*w = s_1^2\langle w, f_1\rangle f_1 + \cdots + s_m^2\langle w, f_m\rangle f_m$$

for all $w \in W$.

תרגילים 7ל

1 נניח ש־ $T \in \mathcal{L}(V, W)$. הראו כי $T = 0$ אם־ורק־אם כל הערכים הסינגולריים של T הם 0.

2 נניח ש־ $T \in \mathcal{L}(V, W)$ ו־ $s > 0$. הוכיחו כי s ערך סינגולרי של T אם־ורק־אם קיימים וקטורים השונים מאפס $v \in V$ ו־ $w \in W$ כך ש־

$$Tv = sw \quad \text{וגם} \quad T^*w = sv.$$

הווקטורים v, w המקיימים את המשוואות לעיל נקראים **צמד שמידט**. ארהרד שמידט הציג את רעיון הערכים הסינגולריים בשנת 1907.

3 הביאו דוגמה של $T \in \mathcal{L}(\mathbf{C}^2)$ כך ש־ 0 הוא הערך העצמי היחיד של T והערכים הסינגולריים של T הם $5, 0$.

4 נניח ש־ $T \in \mathcal{L}(V, W)$, s_1 הוא הערך הסינגולרי הגבוה ביותר של T, ו־ s_n הערך הסינגולרי הנמוך ביותר של T. הוכיחו כי

$$\{\|Tv\| : v \in V \text{ and } \|v\| = 1\} = [s_n, s_1].$$

5 נניח ש־ $T \in \mathcal{L}(\mathbf{C}^2)$ מוגדר על ידי $T(x, y) = (-4y, x)$. מצאו את הערכים הסינגולריים של T.

6 מצאו את הערכים הסינגולריים של אופרטור הגזירה $D \in \mathcal{L}(\mathcal{P}_2(\mathbf{R}))$ המוגדר על ידי $Dp = p'$, כאשר המכפלה הפנימית מעל $\mathcal{P}_2(\mathbf{R})$ היא כמו בדוגמה 6.34.

7 נניח ש־ $T \in \mathcal{L}(V)$ צמוד לעצמו או ש־ $\mathbf{F} = \mathbf{C}$ ו־ $T \in \mathcal{L}(V)$ נורמלי. יהיו $\lambda_1, \dots, \lambda_n$ הערכים העצמיים של T, כאשר כל אחד מופיע ברשימה מספר פעמים הזהה לממד המרחב העצמי השייך. הראו כי הערכים הסינגולריים של T הם $|\lambda_1|, \dots, |\lambda_n|$, אחרי שמספרים אלו מוינו בסדר יורד.

8 נניח ש־ $T \in \mathcal{L}(V, W)$. נניח ש־ $s_1 \geq s_2 \geq \cdots \geq s_m > 0$ רשימה אורתונורמלית ב־ V ו־ $f_1, \dots, f_m$ רשימה אורתונורמלית ב־ W כך ש־

$$Tv = s_1\langle v, e_1\rangle f_1 + \cdots + s_m\langle v, e_m\rangle f_m$$

לכל $v \in V$.

א. הוכיחו כי $f_1, \dots, f_m$ בסיס אורתונורמלי של $\operatorname{range} T$.

ב. הוכיחו כי $e_1, \dots, e_m$ בסיס אורתונורמלי של $(\operatorname{null} T)^\perp$.

ג. הוכיחו כי $s_1, \dots, s_m$ הם הערכים הסינגולריים החיוביים של T.

ד. הוכיחו כי אם $k \in \{1, \dots, m\}$, אזי e_k וקטור עצמי של T^*T שהערך העצמי השייך לו הוא s_k^2.

ה. הוכיחו כי

$$TT^*w = s_1^2\langle w, f_1\rangle f_1 + \cdots + s_m^2\langle w, f_m\rangle f_m$$

לכל $w \in W$.

9 Suppose $T \in \mathcal{L}(V, W)$. Show that T and T^* have the same positive singular values.

10 Suppose $T \in \mathcal{L}(V, W)$ has singular values $s_1, \dots, s_n$. Prove that if T is an invertible linear map, then T^{-1} has singular values

$$\frac{1}{s_n}, \dots, \frac{1}{s_1}.$$

11 Suppose that $T \in \mathcal{L}(V, W)$ and $v_1, \dots, v_n$ is an orthonormal basis of V. Let $s_1, \dots, s_n$ denote the singular values of T.

(a) Prove that $\|Tv_1\|^2 + \cdots + \|Tv_n\|^2 = s_1^2 + \cdots + s_n^2$.

(b) Prove that if $W = V$ and T is a positive operator, then

$$\langle Tv_1, v_1 \rangle + \cdots + \langle Tv_n, v_n \rangle = s_1 + \cdots + s_n.$$

See the comment after Exercise 5 in Section 7A.

12 (a) Give an example of a finite-dimensional vector space and an operator T on it such that the singular values of T^2 do not equal the squares of the singular values of T.

(b) Suppose $T \in \mathcal{L}(V)$ is normal. Prove that the singular values of T^2 equal the squares of the singular values of T.

13 Suppose $T_1, T_2 \in \mathcal{L}(V)$. Prove that T_1 and T_2 have the same singular values if and only if there exist unitary operators $S_1, S_2 \in \mathcal{L}(V)$ such that $T_1 = S_1 T_2 S_2$.

14 Suppose $T \in \mathcal{L}(V, W)$. Let s_n denote the smallest singular value of T. Prove that $s_n \|v\| \le \|Tv\|$ for every $v \in V$.

15 Suppose $T \in \mathcal{L}(V)$ and $s_1 \ge \cdots \ge s_n$ are the singular values of T. Prove that if λ is an eigenvalue of T, then $s_1 \ge |\lambda| \ge s_n$.

16 Suppose $T \in \mathcal{L}(V, W)$. Prove that $(T^*)^\dagger = (T^\dagger)^*$.

Compare the result in this exercise to the analogous result for invertible linear maps [see 7.5(f)].

17 Suppose $T \in \mathcal{L}(V)$. Prove that T is self-adjoint if and only if $T^\dagger$ is self-adjoint.

Matrices unfold
Singular values gleam like stars
Order in chaos shines

—written by ChatGPT with input *haiku about SVD*

9 נניח ש־ $T \in \mathcal{L}(V,W)$. הראו כי ל־ T ול־ T^* יש אותם ערכים סינגולריים חיוביים.

10 נניח שיש ל־ $T \in \mathcal{L}(V,W)$ ערכים סינגולריים $s_1, \dots, s_n$. הוכיחו כי אם T העתקה ליניארית הפיכה, אזי יש ל־ T^{-1} ערכים סינגולריים

$$\frac{1}{s_n}, \dots, \frac{1}{s_1}.$$

11 נניח ש־ $T \in \mathcal{L}(V,W)$ ו־ $v_1, \dots, v_n$ בסיס אורתונורמלי של V. יהיו $s_1, \dots, s_n$ הערכים הסינגולריים של T.

א. הוכיחו כי $\|Tv_1\|^2 + \cdots + \|Tv_n\|^2 = s_1^2 + \cdots + s_n^2$.

ב. הוכיחו כי אם $W = V$ ו־ T אופרטור חיובי, אזי

$$\langle Tv_1, v_1\rangle + \cdots + \langle Tv_n, v_n\rangle = s_1 + \cdots + s_n.$$

ראו ההערה שאחרי תרגיל 5 ביחידה 7ז.

12 א. הביאו דוגמה של מרחב וקטורי מממד סופי ואופרטור T מעליו כך שהערכים הסינגולריים של T^2 אינם שווים לריבועים של הערכים הסינגולריים של T.

ב. נניח ש־ $T \in \mathcal{L}(V)$ נורמלי. הוכיחו כי הערכים הסינגולריים של T^2 שווים לריבועים של הערכים הסינגולריים של T.

13 נניח ש־ $T_1, T_2 \in \mathcal{L}(V)$. הוכיחו כי ל־ T_1 ול־ T_2 יש אותם ערכים סינגולריים אם־ורק־אם קיימים אופרטורים אוניטריים $S_1, S_2 \in \mathcal{L}(V)$ כך ש־ $T_1 = S_1 T_2 S_2$.

14 נניח ש־ $T \in \mathcal{L}(V,W)$. יהי s_n הערך הסינגולרי הנמוך ביותר של T. הוכיחו כי $s_n\|v\| \le \|Tv\|$ לכל $v \in V$.

15 נניח ש־ $T \in \mathcal{L}(V)$ ו־ $s_1 \ge \cdots \ge s_n$ הערכים הסינגולריים של T. הוכיחו כי אם λ ערך עצמי של T, אזי $s_1 \ge |\lambda| \ge s_n$.

16 נניח ש־ $T \in \mathcal{L}(V,W)$. הוכיחו כי $(T^*)^\dagger = (T^\dagger)^*$.

השוו את הפסקנה בתרגיל זה לפסקנה האנלוגית עבור העתקות ליניאריות הפיכות [ראו 7.5(ו)].

17 נניח ש־ $T \in \mathcal{L}(V)$. הוכיחו כי T צמוד לעצמו אם־ורק־אם $T^\dagger$ צמוד לעצמו.

מַטְרִיצָה–תְּהוּ
שָׁלֹשׁ יָפוֹת מִתְלַכְּדוֹת
מְתֹהוּ–סֵדֶר

—נכתב על ידי קלוד אופוס 4.8
קלט: haiku about SVD in Hebrew

7F Consequences of Singular Value Decomposition

Norms of Linear Maps

The singular value decomposition leads to the following upper bound for $\|Tv\|$.

7.82 *upper bound for $\|Tv\|$*

Suppose $T \in \mathcal{L}(V, W)$. Let s_1 be the largest singular value of T. Then

$$\|Tv\| \le s_1\|v\|$$

for all $v \in V$.

Proof Let $s_1, \dots, s_m$ denote the positive singular values of T, and let $e_1, \dots, e_m$ be an orthonormal list in V and $f_1, \dots, f_m$ be an orthonormal list in W that provide a singular value decomposition of T. Thus

For a lower bound on $\|Tv\|$, look at Exercise 14 in Section 7E.

7.83 $$Tv = s_1\langle v, e_1\rangle f_1 + \cdots + s_m\langle v, e_m\rangle f_m$$

for all $v \in V$. Hence if $v \in V$ then

$$\begin{aligned}\|Tv\|^2 &= s_1^2\,|\langle v, e_1\rangle|^2 + \cdots + s_m^2\,|\langle v, e_m\rangle|^2\\ &\le s_1^2\big(|\langle v, e_1\rangle|^2 + \cdots + |\langle v, e_m\rangle|^2\big)\\ &\le s_1^2\,\|v\|^2,\end{aligned}$$

where the last inequality follows from Bessel's inequality (6.26). Taking square roots of both sides of the inequality above shows that $\|Tv\| \le s_1\|v\|$, as desired. ∎

Suppose $T \in \mathcal{L}(V, W)$ and s_1 is the largest singular value of T. The result above shows that

7.84 $$\|Tv\| \le s_1 \text{ for all } v \in V \text{ with } \|v\| \le 1.$$

Taking $v = e_1$ in 7.83 shows that $Te_1 = s_1 f_1$. Because $\|f_1\| = 1$, this implies that $\|Te_1\| = s_1$. Thus because $\|e_1\| = 1$, the inequality in 7.84 leads to the equation

7.85 $$\max\{\|Tv\| : v \in V \text{ and } \|v\| \le 1\} = s_1.$$

The equation above is the motivation for the following definition, which defines the norm of T to be the left side of the equation above without needing to refer to singular values or the singular value decomposition.

7.86 definition: *norm of a linear map, $\|\cdot\|$*

Suppose $T \in \mathcal{L}(V, W)$. Then the *norm* of T, denoted by $\|T\|$, is defined by

$$\|T\| = \max\{\|Tv\| : v \in V \text{ and } \|v\| \le 1\}.$$

ז׳ מסקנות מהפירוק לערכים סינגולריים

נורמות של העתקות ליניאריות

הפירוק לערכים סינגולריים מוביל לחסם העליון הבא עבור $\|Tv\|$.

7.82 חסם עליון של $\|Tv\|$

נניח ש־ $T \in \mathcal{L}(V,W)$. יהי s_1 הערך הסינגולרי הגבוה ביותר של T. אזי

$$\|Tv\| \le s_1\|v\|$$

לכל $v \in V$.

עבור החסם התחתון של $\|Tv\|$, ראו תרגיל 14 בחידה ז׳.

הוכחה נסמן ב־ $s_1, \dots, s_m$ את הערכים הסינגולריים החיוביים של T, ותהי $e_1, \dots, e_m$ רשימה אורתונורמלית ב־ V ו־ $f_1, \dots, f_m$ רשימה אורתונורמלית ב־ W המספקות פירוק לערכים סינגולריים של T. לפיכך

$$Tv = s_1\langle v, e_1\rangle f_1 + \cdots + s_m\langle v, e_m\rangle f_m \tag{7.83}$$

לכל $v \in V$. לכן אם $v \in V$ אזי

$$\begin{aligned}\|Tv\|^2 &= s_1^2\,|\langle v, e_1\rangle|^2 + \cdots + s_m^2\,|\langle v, e_m\rangle|^2\\ &\le s_1^2\left(|\langle v, e_1\rangle|^2 + \cdots + |\langle v, e_m\rangle|^2\right)\\ &\le s_1^2\,\|v\|^2,\end{aligned}$$

כאשר האי־שוויון האחרון נובע מאי־שוויון בֶּסֶל (6.26). הוצאת שורשים משני אגפי האי־שוויון לעיל מראה ש־ $\|Tv\| \le s_1\|v\|$, כפי שרצינו. ■

נניח ש־ $T \in \mathcal{L}(V,W)$ ו־ s_1 הערך הסינגולרי הגבוה ביותר של T. המסקנה לעיל מראה ש־

$$\|Tv\| \le s_1 \text{ for all } v \in V \text{ with } \|v\| \le 1. \tag{7.84}$$

כשניקח $v = e_1$ ב־ 7.83 נקבל ש־ $Te_1 = s_1 f_1$. מאחר ש־ $\|f_1\| = 1$, מכאן נובע ש־ $\|Te_1\| = s_1$. לפיכך מאחר ש־ $\|e_1\| = 1$, האי־שוויון ב־ 7.84 מוביל למשוואה

$$\max\{\|Tv\| : v \in V \text{ and } \|v\| \le 1\} = s_1. \tag{7.85}$$

המשוואה לעיל היא המוטיבציה להגדרה הבאה, המגדירה את הנורמה של T להיות האגף השמאלי של המשוואה לעיל ללא הצורך בהפניה לערכים סינגולריים או לפירוק לערכים סינגולריים.

7.86 הגדרה: נורמה של העתקה ליניארית, $\|\cdot\|$

נניח ש־ $T \in \mathcal{L}(V,W)$. אזי הנורמה של T, מסומנת על ידי $\|T\|$, מוגדרת על ידי

$$\|T\| = \max\{\|Tv\| : v \in V \text{ וגם } \|v\| \le 1\}.$$

In general, the maximum of an infinite set of nonnegative numbers need not exist. However, the discussion before 7.86 shows that the maximum in the definition of the norm of a linear map T from V to W does indeed exist (and equals the largest singular value of T).

We now have two different uses of the word *norm* and the notation $\|\cdot\|$. Our first use of this notation was in connection with an inner product on V, when we defined $\|v\| = \sqrt{\langle v, v\rangle}$ for each $v \in V$. Our second use of the norm notation and terminology is with the definition we just made of $\|T\|$ for $T \in \mathcal{L}(V, W)$. The norm $\|T\|$ for $T \in \mathcal{L}(V, W)$ does not usually come from taking an inner product of T with itself (see Exercise 21). You should be able to tell from the context and from the symbols used which meaning of the norm is intended.

The properties of the norm on $\mathcal{L}(V, W)$ listed below look identical to properties of the norm on an inner product space (see 6.9 and 6.17). The inequality in (d) is called the *triangle inequality*, thus using the same terminology that we used for the norm on V. For the reverse triangle inequality, see Exercise 1.

7.87 *basic properties of norms of linear maps*

Suppose $T \in \mathcal{L}(V, W)$. Then

(a) $\|T\| \geq 0$;

(b) $\|T\| = 0 \iff T = 0$;

(c) $\|\lambda T\| = |\lambda| \, \|T\|$ for all $\lambda \in \mathbf{F}$;

(d) $\|S + T\| \leq \|S\| + \|T\|$ for all $S \in \mathcal{L}(V, W)$.

Proof

(a) Because $\|Tv\| \geq 0$ for every $v \in V$, the definition of $\|T\|$ implies that $\|T\| \geq 0$.

(b) Suppose $\|T\| = 0$. Thus $Tv = 0$ for all $v \in V$ with $\|v\| \leq 1$. If $u \in V$ with $u \neq 0$, then

$$Tu = \|u\| \, T\left(\frac{u}{\|u\|}\right) = 0,$$

where the last equality holds because $u/\|u\|$ has norm 1. Because $Tu = 0$ for all $u \in V$, we have $T = 0$.

Conversely, if $T = 0$ then $Tv = 0$ for all $v \in V$ and hence $\|T\| = 0$.

(c) Suppose $\lambda \in \mathbf{F}$. Then

$$\begin{aligned}\|\lambda T\| &= \max\{\|\lambda Tv\| : v \in V \text{ and } \|v\| \leq 1\} \\ &= |\lambda| \max\{\|Tv\| : v \in V \text{ and } \|v\| \leq 1\} \\ &= |\lambda| \, \|T\|.\end{aligned}$$

(d) Suppose $S \in \mathcal{L}(V, W)$. The definition of $\|S + T\|$ implies that there exists $v \in V$ such that $\|v\| \leq 1$ and $\|S + T\| = \|(S + T)v\|$. Now

$$\|S + T\| = \|(S + T)v\| = \|Sv + Tv\| \leq \|Sv\| + \|Tv\| \leq \|S\| + \|T\|,$$

completing the proof of (d). ∎

באופן כללי, המקסימום של קבוצה אינסופית של מספרים אי־שליליים אינו קיים בהכרח. עם זאת, מה שהצגנו לפני ההגדרה ב־ 7.86 מראה כי המקסימום בהגדרה של הנורמה של העתקה ליניארית T מ־ V אל W קיים בהחלט (ושווה לערך הסינגולרי הגבוה ביותר של T).

כעת יש לנו שני שימושים שונים של המונח נורמה והסימון $\|\cdot\|$. השימוש הראשון שלנו בסימון קשור למכפלה פנימית מעל V, כאשר הגדרנו $\|v\| = \sqrt{\langle v, v\rangle}$ לכל $v \in V$. השימוש השני שלנו במינוח ובסימון הנורמה הוא ההגדרה שזה עתה נתנו ל־ $\|T\|$ עבור $T \in \mathcal{L}(V, W)$. הנורמה $\|T\|$ של $T \in \mathcal{L}(V, W)$ אינה מתקבלת בדרך כלל מחישוב המכפלה הפנימית של T עם עצמה (ראו תרגיל 21). ביכולתכם להבחין מתוך ההקשר והסימונים לאיזו מהמשמעויות של נורמה מתכוונים.

התכונות של הנורמה מעל $\mathcal{L}(V, W)$ מפורטות להלן ונראות זהות לתכונות הנורמה מעל מרחב מכפלה פנימית (ראו 6.9 ו־ 6.17). האי־שוויון ב־ (ד) נקרא האי־שוויון המשולש, ולפיכך נעשה שימוש באותה טרמינולוגיה שבה השתמשנו עבור הנורמה מעל V. לאי־שוויון המשולש ההפוך, ראו תרגיל 1.

7.87 תכונות בסיסיות של נורמות של העתקות ליניאריות

נניח ש־ $T \in \mathcal{L}(V, W)$. אזי

א. $\|T\| \geq 0$

ב. $\|T\| = 0 \iff T = 0$

ג. $\|\lambda T\| = |\lambda| \|T\|$ לכל $\lambda \in \mathbf{F}$

ד. $\|S + T\| \leq \|S\| + \|T\|$ לכל $S \in \mathcal{L}(V, W)$.

הוכחה

א. מאחר ש־ $\|Tv\| \geq 0$ לכל $v \in V$, ההגדרה של $\|T\|$ גוררת ש־ $\|T\| \geq 0$.

ב. נניח ש־ $\|T\| = 0$. לפיכך $Tv = 0$ לכל $v \in V$ עם $\|v\| \leq 1$. אם $u \in V$ עם $u \neq 0$, אזי

$$Tu = \|u\| T\left(\frac{u}{\|u\|}\right) = 0,$$

כאשר השוויון האחרון מתקיים מאחר שהנורמה של $u/\|u\|$ היא 1. מאחר ש־ $Tu = 0$ לכל $u \in V$, מתקיים $T = 0$.

בכיוון ההפוך, אם $T = 0$ אזי $Tv = 0$ לכל $v \in V$ ולכן $\|T\| = 0$.

ג. נניח ש־ $\lambda \in \mathbf{F}$. אזי

$$\begin{aligned}\|\lambda T\| &= \max\{\|\lambda Tv\| : v \in V \text{ וגם } \|v\| \leq 1\} \\ &= |\lambda| \max\{\|Tv\| : v \in V \text{ וגם } \|v\| \leq 1\} \\ &= |\lambda| \|T\|.\end{aligned}$$

ד. נניח ש־ $S \in \mathcal{L}(V, W)$. ההגדרה של $\|S+T\|$ גוררת שקיים $v \in V$ כך ש־ $\|v\| \leq 1$ וגם $\|S + T\| = \|(S + T)v\|$. כעת

$$\|S + T\| = \|(S + T)v\| = \|Sv + Tv\| \leq \|Sv\| + \|Tv\| \leq \|S\| + \|T\|,$$

ובכך השלמנו את ההוכחה של (ד). ■

For $S, T \in \mathcal{L}(V, W)$, the quantity $\|S - T\|$ is often called the distance between S and T. Informally, think of the condition that $\|S - T\|$ is a small number as meaning that S and T are close together. For example, Exercise 9 asserts that for every $T \in \mathcal{L}(V)$, there is an invertible operator as close to T as we wish.

7.88 *alternative formulas for $\|T\|$*

Suppose $T \in \mathcal{L}(V, W)$. Then

(a) $\|T\|$ = the largest singular value of T;

(b) $\|T\| = \max\{\|Tv\| : v \in V \text{ and } \|v\| = 1\}$;

(c) $\|T\|$ = the smallest number c such that $\|Tv\| \le c\|v\|$ for all $v \in V$.

Proof

(a) See 7.85.

(b) Let $v \in V$ be such that $0 < \|v\| \le 1$. Let $u = v/\|v\|$. Then

$$\|u\| = \left\|\frac{v}{\|v\|}\right\| = 1 \quad \text{and} \quad \|Tu\| = \left\|T\Big(\frac{v}{\|v\|}\Big)\right\| = \frac{\|Tv\|}{\|v\|} \ge \|Tv\|.$$

Thus when finding the maximum of $\|Tv\|$ with $\|v\| \le 1$, we can restrict attention to vectors in V with norm 1, proving (b).

(c) Suppose $v \in V$ and $v \ne 0$. Then the definition of $\|T\|$ implies that

$$\left\|T\Big(\frac{v}{\|v\|}\Big)\right\| \le \|T\|,$$

which implies that

7.89 $$\|Tv\| \le \|T\|\,\|v\|.$$

Now suppose $c \ge 0$ and $\|Tv\| \le c\|v\|$ for all $v \in V$. This implies that

$$\|Tv\| \le c$$

for all $v \in V$ with $\|v\| \le 1$. Taking the maximum of the left side of the inequality above over all $v \in V$ with $\|v\| \le 1$ shows that $\|T\| \le c$. Thus $\|T\|$ is the smallest number c such that $\|Tv\| \le c\|v\|$ for all $v \in V$. ∎

When working with norms of linear maps, you will probably frequently use the inequality 7.89.

For computing an approximation of the norm of a linear map T given the matrix of T with respect to some orthonormal bases, 7.88(a) is likely to be most useful. The matrix of T^*T is quickly computable from matrix multiplication. Then a computer can be asked to find an approximation for the largest eigenvalue of T^*T (excellent numerical algorithms exist for this purpose). Then taking the square root and using 7.88(a) gives an approximation for the norm of T (which usually cannot be computed exactly).

עבור $S, T \in \mathcal{L}(V, W)$, הגודל $\|S - T\|$ לעיתים קרובות נקרא המרחק בין S ל־T. באופן בלתי רשמי, חשבו על מצב שבו $\|S - T\|$ הוא מספר קטן במובן ש־S ו־T קרובים זה לזה. לדוגמה, תרגיל 9 טוען שלכל $T \in \mathcal{L}(V)$, ישנו אופרטור הפיך קרוב ל־T ככל שנרצה.

7.88 נוסחאות חלופיות עבור $\|T\|$

נניח ש־$T \in \mathcal{L}(V, W)$. אזי

א. $\|T\| =$ הערך הסינגולרי הגבוה ביותר של T

ב. $\|T\| = \max\{\|Tv\| : v \in V \text{ וגם } \|v\| = 1\}$

ג. $\|T\| =$ המספר הנמוך ביותר c כך ש־$\|Tv\| \le c\|v\|$ לכל $v \in V$.

הוכחה

א. ראו 7.85.

ב. יהי $v \in V$ כך ש־$0 < \|v\| \le 1$. יהי $u = v/\|v\|$. אזי

$$\|u\| = \left\|\frac{v}{\|v\|}\right\| = 1 \quad \text{ו־} \quad \|Tu\| = \left\|T\Big(\frac{v}{\|v\|}\Big)\right\| = \frac{\|Tv\|}{\|v\|} \ge \|Tv\|.$$

לפיכך כאשר אנו מוצאים את המקסימום של $\|Tv\|$ עם $\|v\| \le 1$, נוכל לצמצם את תשומת הלב אל וקטורים ב־V עם נורמה 1, ובכך הוכחנו את (ב).

ג. נניח ש־$v \in V$ ו־$v \ne 0$. אזי ההגדרה של $\|T\|$ גוררת ש־

$$\left\|T\Big(\frac{v}{\|v\|}\Big)\right\| \le \|T\|,$$

מה שגורר ש־

$$\|Tv\| \le \|T\|\,\|v\|. \tag{7.89}$$

כעת נניח ש־$c \ge 0$ ו־$\|Tv\| \le c\|v\|$ לכל $v \in V$. מכאן נובע ש־

$$\|Tv\| \le c$$

לכל $v \in V$ עם $\|v\| \le 1$. כשניקח את המקסימום של האגף השמאלי של האי־שוויון לעיל מבין כל $v \in V$ עם $\|v\| \le 1$ נקבל ש־$\|T\| \le c$. לפיכך $\|T\|$ הוא המספר הנמוך ביותר c כך ש־$\|Tv\| \le c\|v\|$ לכל $v \in V$. ■

כשתעבדו עם נורמות של העתקות ליניאריות, ככל הנראה לעיתים קרובות תשתמשו באי־שוויון 7.89.

לחישוב קירובים של הנורמה של העתקה ליניארית T בהינתן המטריצה של T ביחס לבסיסים אורתונורמליים כלשהם, 7.88(א) יהיה ככל הנראה השימושי ביותר. המטריצה של T^*T ניתנת לחישוב באופן מהיר באמצעות כפל מטריצות. בשלב זה נוכל להשתמש במחשב למציאת קירוב לערך העצמי הגבוה ביותר של T^*T (קיימים אלגוריתמים מספריים מצוינים למטרה זו). הוצאת שורשים ריבועיים ושימוש ב־7.88(א) נותנים בשלב זה קירוב לנורמה של T (אשר בדרך כלל לא ניתנת לחישוב מדויק).

You should verify all assertions in the example below.

7.90 example: *norms*

- If I denotes the usual identity operator on V, then $\|I\| = 1$.
- If $T \in \mathcal{L}(\mathbf{F}^n)$ and the matrix of T with respect to the standard basis of $\mathbf{F}^n$ consists of all 1's, then $\|T\| = n$.
- If $T \in \mathcal{L}(V)$ and V has an orthonormal basis consisting of eigenvectors of T with corresponding eigenvalues $\lambda_1, \ldots, \lambda_n$, then $\|T\|$ is the maximum of the numbers $|\lambda_1|, \ldots, |\lambda_n|$.
- Suppose $T \in \mathcal{L}(\mathbf{R}^5)$ is the operator whose matrix (with respect to the standard basis) is the 5-by-5 matrix whose entry in row j, column k is $1/(j^2 + k)$. Standard mathematical software shows that the largest singular value of T is approximately 0.8 and the smallest singular value of T is approximately 10^{-6}. Thus $\|T\| \approx 0.8$ and (using Exercise 10 in Section 7E) $\|T^{-1}\| \approx 10^6$. It is not possible to find exact formulas for these norms.

A linear map and its adjoint have the same norm, as shown by the next result.

7.91 *norm of the adjoint*

Suppose $T \in \mathcal{L}(V, W)$. Then $\|T^*\| = \|T\|$.

Proof Suppose $w \in W$. Then

$$\|T^*w\|^2 = \langle T^*w, T^*w\rangle = \langle TT^*w, w\rangle \le \|TT^*w\|\,\|w\| \le \|T\|\,\|T^*w\|\,\|w\|.$$

The inequality above implies that

$$\|T^*w\| \le \|T\|\,\|w\|,$$

which along with 7.88(c) implies that $\|T^*\| \le \|T\|$.

Replacing T with T^* in the inequality $\|T^*\| \le \|T\|$ and then using the equation $(T^*)^* = T$ shows that $\|T\| \le \|T^*\|$. Thus $\|T^*\| = \|T\|$, as desired. ∎

You may want to construct an alternative proof of the result above using Exercise 9 in Section 7E, which asserts that a linear map and its adjoint have the same positive singular values.

Approximation by Linear Maps with Lower-Dimensional Range

The next result is a spectacular application of the singular value decomposition. It says that to best approximate a linear map by a linear map whose range has dimension at most k, chop off the singular value decomposition after the first k terms. Specifically, the linear map T_k in the next result has the property that $\dim \operatorname{range} T_k = k$ and T_k minimizes the distance to T among all linear maps with range of dimension at most k. This result leads to algorithms for compressing huge matrices while preserving their most important information.

עליכם לוודא את כל הטענות בדוגמה להלן.

7.90 דוגמה: נורמות

- אם I מסמן את אופרטור הזהות הרגיל מעל V, אזי $\|I\| = 1$.
- אם $T \in \mathcal{L}(\mathbf{F}^n)$ וכל איברי המטריצה של T ביחס לבסיס הסטנדרטי של $\mathbf{F}^n$ הם 1, אזי $\|T\| = n$.
- אם $T \in \mathcal{L}(V)$ ויש ל־ V בסיס אורתונורמלי המורכב מווקטורים עצמיים של T שהערכים העצמיים השייכים להם הם $\lambda_1, \dots, \lambda_n$, אזי $\|T\|$ היא המקסימום של המספרים $|\lambda_1|, \dots, |\lambda_n|$.
- נניח ש־ $T \in \mathcal{L}(\mathbf{R}^5)$ הוא האופרטור שהמטריצה שלו (ביחס לבסיס הסטנדרטי) היא מטריצה 5×5 שהאיבר בשורה j, עמודה k הוא $1/(j^2 + k)$. תוכנה מתמטית סטנדרטית מראה שהערך הסינגולרי הגבוה ביותר של T הוא 0.8 בקירוב והערך הסינגולרי הנמוך ביותר של T הוא 10^{-6} בקירוב. לפיכך $\|T\| \approx 0.8$ (ולפי תרגיל 10 ביחידה 5ז) $\|T^{-1}\| \approx 10^6$. אין אפשרות למצוא את הנוסחאות המדויקות לנורמות אלו.

להעתקה ליניארית ולהעתקה הצמודה שלה יש אותה נורמה, כפי שמראה לנו המסקנה הבאה.

7.91 נורמה של הצמודה

נניח ש־ $T \in \mathcal{L}(V, W)$. אזי $\|T\| = \|T^*\|$.

הוכחה נניח ש־ $w \in W$. אזי

$$\|T^*w\|^2 = \langle T^*w, T^*w\rangle = \langle TT^*w, w\rangle \le \|TT^*w\|\,\|w\| \le \|T\|\,\|T^*w\|\,\|w\|.$$

האי־שוויון לעיל גורר ש־

$$\|T^*w\| \le \|T\|\,\|w\|,$$

אשר יחד עם 7.88(ג) גורר ש־ $\|T\| \le \|T^*\|$.

החלפת T ב־ T^* באי־שוויון $\|T\| \le \|T^*\|$ ואז שימוש במשוואה $(T^*)^* = T$ מראה לנו ש־ $\|T^*\| \le \|T\|$. לפיכך $\|T\| = \|T^*\|$, כפי שרצינו. ■

ייתכן שתרצו להרכיב הוכחה חלופית עבור המסקנה לעיל בהתבסס על תרגיל 9 ביחידה 5ז, הטוען כי להעתקה ליניארית ולהעתקה הצמודה שלה יש אותם ערכים סינגולריים חיוביים.

קירובים באמצעות העתקות ליניאריות עם ממד טווח נמוך יותר

המסקנה הבאה היא יישום מרהיב של הפירוק לערכים סינגולריים. היא קובעת כי כדי לחשב את הקירוב הטוב ביותר של העתקה ליניארית באמצעות העתקה ליניארית שממד הטווח שלה הוא לכל היותר k, יש לחתוך את הפירוק לערכים סינגולריים אחרי k הביטויים הראשונים. ליתר דיוק, להעתקה הליניארית T_k במסקנה הבאה התכונה ש־ $\dim \operatorname{range} T_k = k$ ו־ T_k ממזערת את המרחק ל־ T מבין כל ההעתקות הליניאריות עם ממד טווח k לכל היותר. מסקנה זו מובילה לאלגוריתמים לדחיסת מטריצות ענק תוך שימור המידע החשוב ביותר האצור בהן.

7.92 *best approximation by linear map whose range has dimension $\le k$*

Suppose $T \in \mathcal{L}(V, W)$ and $s_1 \ge \cdots \ge s_m$ are the positive singular values of T. Suppose $1 \le k < m$. Then

$$\min\{\|T - S\| : S \in \mathcal{L}(V, W) \text{ and } \dim \operatorname{range} S \le k\} = s_{k+1}.$$

Furthermore, if

$$Tv = s_1\langle v, e_1\rangle f_1 + \cdots + s_m\langle v, e_m\rangle f_m$$

is a singular value decomposition of T and $T_k \in \mathcal{L}(V, W)$ is defined by

$$T_k v = s_1\langle v, e_1\rangle f_1 + \cdots + s_k\langle v, e_k\rangle f_k$$

for each $v \in V$, then $\dim \operatorname{range} T_k = k$ and $\|T - T_k\| = s_{k+1}$.

Proof If $v \in V$ then

$$\begin{aligned}
\|(T - T_k)v\|^2 &= \|s_{k+1}\langle v, e_{k+1}\rangle f_{k+1} + \cdots + s_m\langle v, e_m\rangle f_m\|^2 \\
&= s_{k+1}^2 |\langle v, e_{k+1}\rangle|^2 + \cdots + s_m^2 |\langle v, e_m\rangle|^2 \\
&\le s_{k+1}^2\big(|\langle v, e_{k+1}\rangle|^2 + \cdots + |\langle v, e_m\rangle|^2\big) \\
&\le s_{k+1}^2 \|v\|^2.
\end{aligned}$$

Thus $\|T - T_k\| \le s_{k+1}$. The equation $(T - T_k)e_{k+1} = s_{k+1}f_{k+1}$ now shows that $\|T - T_k\| = s_{k+1}$.

Suppose $S \in \mathcal{L}(V, W)$ and $\dim \operatorname{range} S \le k$. Thus $Se_1, \ldots, Se_{k+1}$, which is a list of length $k + 1$, is linearly dependent. Hence there exist $a_1, \ldots, a_{k+1} \in \mathbf{F}$, not all 0, such that

$$a_1 Se_1 + \cdots + a_{k+1}Se_{k+1} = 0.$$

Now $a_1e_1 + \cdots + a_{k+1}e_{k+1} \ne 0$ because $a_1, \ldots, a_{k+1}$ are not all 0. We have

$$\begin{aligned}
\|(T - S)(a_1e_1 + \cdots + a_{k+1}e_{k+1})\|^2 &= \|T(a_1e_1 + \cdots + a_{k+1}e_{k+1})\|^2 \\
&= \|s_1a_1f_1 + \cdots + s_{k+1}a_{k+1}f_{k+1}\|^2 \\
&= s_1^2|a_1|^2 + \cdots + s_{k+1}^2|a_{k+1}|^2 \\
&\ge s_{k+1}^2(|a_1|^2 + \cdots + |a_{k+1}|^2) \\
&= s_{k+1}^2\|a_1e_1 + \cdots + a_{k+1}e_{k+1}\|^2.
\end{aligned}$$

Because $a_1e_1 + \cdots + a_{k+1}e_{k+1} \ne 0$, the inequality above implies that

$$\|T - S\| \ge s_{k+1}.$$

Thus $S = T_k$ minimizes $\|T - S\|$ among $S \in \mathcal{L}(V, W)$ with $\dim \operatorname{range} S \le k$. ∎

For other examples of the use of the singular value decomposition in best approximation, see Exercise 22, which finds a subspace of given dimension on which the restriction of a linear map is as small as possible, and Exercise 27, which finds a unitary operator that is as close as possible to a given operator.

7.92 הקירוב הטוב ביותר באמצעות העתקה ליניארית שמימד הטווח $\le k$

נניח ש־ $T \in \mathcal{L}(V,W)$ ו־ $s_1 \ge \cdots \ge s_m$ הערכים הסינגולריים החיוביים של T.
נניח ש־ $1 \le k < m$. אזי

$$\min\{\|T-S\| : S \in \mathcal{L}(V,W) \text{ וגם } \dim \operatorname{range} S \le k\} = s_{k+1}.$$

יתר על כן, אם

$$Tv = s_1\langle v, e_1\rangle f_1 + \cdots + s_m\langle v, e_m\rangle f_m$$

הוא פירוק לערכים סינגולריים של T ו־ $T_k \in \mathcal{L}(V,W)$ מוגדר על ידי

$$T_k v = s_1\langle v, e_1\rangle f_1 + \cdots + s_k\langle v, e_k\rangle f_k$$

לכל $v \in V$, אזי $\dim \operatorname{range} T_k = k$ ו־ $\|T - T_k\| = s_{k+1}$.

הוכחה אם $v \in V$ אזי

$$\begin{aligned}
\|(T-T_k)v\|^2 &= \|s_{k+1}\langle v, e_{k+1}\rangle f_{k+1} + \cdots + s_m\langle v, e_m\rangle f_m\|^2 \\
&= s_{k+1}^2 |\langle v, e_{k+1}\rangle|^2 + \cdots + s_m^2 |\langle v, e_m\rangle|^2 \\
&\le s_{k+1}^2\left(|\langle v, e_{k+1}\rangle|^2 + \cdots + |\langle v, e_m\rangle|^2\right) \\
&\le s_{k+1}^2 \|v\|^2.
\end{aligned}$$

לפיכך $\|T - T_k\| \le s_{k+1}$. המשוואה $(T-T_k)e_{k+1} = s_{k+1}f_{k+1}$ כעת מראה לנו ש־ $\|T - T_k\| = s_{k+1}$.

נניח ש־ $S \in \mathcal{L}(V,W)$ ו־ $\dim \operatorname{range} S \le k$. לפיכך $Se_1, \ldots, Se_{k+1}$, שהיא רשימה באורך $k+1$, תלויה ליניארית. לכן קיימים $a_1, \ldots, a_{k+1} \in \mathbf{F}$, שלא כולם 0, כך ש־

$$a_1 Se_1 + \cdots + a_{k+1}Se_{k+1} = 0.$$

כעת $a_1e_1 + \cdots + a_{k+1}e_{k+1} \ne 0$ מאחר ש־ $a_1, \ldots, a_{k+1}$ לא כולם 0. מתקיים

$$\begin{aligned}
\|(T-S)(a_1e_1 + \cdots + a_{k+1}e_{k+1})\|^2 &= \|T(a_1e_1 + \cdots + a_{k+1}e_{k+1})\|^2 \\
&= \|s_1a_1f_1 + \cdots + s_{k+1}a_{k+1}f_{k+1}\|^2 \\
&= s_1^2|a_1|^2 + \cdots + s_{k+1}^2|a_{k+1}|^2 \\
&\ge s_{k+1}^2(|a_1|^2 + \cdots + |a_{k+1}|^2) \\
&= s_{k+1}^2\|a_1e_1 + \cdots + a_{k+1}e_{k+1}\|^2.
\end{aligned}$$

מאחר ש־ $a_1e_1 + \cdots + a_{k+1}e_{k+1} \ne 0$, האי־שוויון לעיל גורר ש־

$$\|T - S\| \ge s_{k+1}.$$

לפיכך $S = T_k$ ממזערת את $\|T-S\|$ מבין $S \in \mathcal{L}(V,W)$ עם $\dim \operatorname{range} S \le k$. ■

לדוגמאות נוספות של שימוש בפירוק לערכים סינגולריים בבעיות קירוב מיטבי, ראו תרגיל 22, המוצא תת־מרחב מממד נתון אשר מעליו הצמצום של העתקה ליניארית הוא קטן ככל הניתן, ותרגיל 27, המוצא אופרטור אוניטרי הקרוב ביותר לאופרטור נתון.

Polar Decomposition

Recall our discussion before 7.54 of the analogy between complex numbers z with $|z| = 1$ and unitary operators. Continuing with this analogy, note that every complex number z except 0 can be written in the form

$$z = \left(\frac{z}{|z|}\right)|z|$$

$$= \left(\frac{z}{|z|}\right)\sqrt{\bar{z}z},$$

where the first factor, namely, $z/|z|$, has absolute value 1.

Our analogy leads us to guess that every operator $T \in \mathcal{L}(V)$ can be written as a unitary operator times $\sqrt{T^*T}$. That guess is indeed correct. The corresponding result is called the polar decomposition, which gives a beautiful description of an arbitrary operator on V.

Note that if $T \in \mathcal{L}(V)$, then T^*T is a positive operator [as was shown in 7.64(a)]. Thus the operator $\sqrt{T^*T}$ makes sense and is well defined as a positive operator on V.

The polar decomposition that we are about to state and prove says that every operator on V is the product of a unitary operator and a positive operator. Thus we can write an arbitrary operator on V as the product of two nice operators, each of which comes from a class that we can completely describe and that we understand reasonably well. The unitary operators are described by 7.55 if $\mathbf{F} = \mathbf{C}$; the positive operators are described by the real and complex spectral theorems (7.29 and 7.31).

Specifically, consider the case $\mathbf{F} = \mathbf{C}$, and suppose

$$T = S\sqrt{T^*T}$$

is a polar decomposition of an operator $T \in \mathcal{L}(V)$, where S is a unitary operator. Then there is an orthonormal basis of V with respect to which S has a diagonal matrix, and there is an orthonormal basis of V with respect to which $\sqrt{T^*T}$ has a diagonal matrix. **Warning:** There may not exist an orthonormal basis that simultaneously puts the matrices of both S and $\sqrt{T^*T}$ into these nice diagonal forms—S may require one orthonormal basis and $\sqrt{T^*T}$ may require a different orthonormal basis.

However (still assuming that $\mathbf{F} = \mathbf{C}$), if T is normal, then an orthonormal basis of V can be chosen such that both S and $\sqrt{T^*T}$ have diagonal matrices with respect to this basis—see Exercise 31. The converse is also true: If $T \in \mathcal{L}(V)$ and $T = S\sqrt{T^*T}$ for some unitary operator $S \in \mathcal{L}(V)$ such that S and $\sqrt{T^*T}$ both have diagonal matrices with respect to the same orthonormal basis of V, then T is normal. This holds because T then has a diagonal matrix with respect to this same orthonormal basis, which implies that T is normal [by the equivalence of (c) and (a) in 7.31].

פירוק פולרי

נזכיר את הדיון שהיה לנו לפני 7.54 הנוגע לאנלוגיה בין המספרים המרוכבים z שעבורם $|z| = 1$ לבין האופרטורים האוניטריים. בהמשך לאנלוגיה זו, נשים לב שכל מספר מרוכב z מלבד 0 יכול להיכתב באופן הבא

$$z = \left(\frac{z}{|z|}\right)|z|$$

$$= \left(\frac{z}{|z|}\right)\sqrt{\overline{z}z},$$

כאשר הערך המוחלט של הגורם הראשון, קרי, $z/|z|$, הוא 1. אנלוגיה זו מובילה אותנו אל הניחוש כי כל אופרטור $T \in \mathcal{L}(V)$ ניתן לכתיבה כאופרטור אוניטרי כפול $\sqrt{T^*T}$. ניחוש זה הוא אכן נכון. המסקנה המתאימה נקראת הפירוק הפולרי, המעניק תיאור יפה לאופרטור שרירותי מעל V.

נשים לב כי אם $T \in \mathcal{L}(V)$, אזי T^*T אופרטור חיובי [כפי שהראינו ב־ 7.64(א)]. לפיכך האופרטור $\sqrt{T^*T}$ הגיוני ומוגדר היטב כאופרטור חיובי מעל V.

הפירוק הפולרי שאנו עומדים להציג ולהוכיח טוען כי כל אופרטור מעל V הוא מכפלה של אופרטור אוניטרי ואופרטור חיובי. לפיכך אנו יכולים לכתוב אופרטור שרירותי מעל V כמכפלה של שני אופרטורים נחמדים, כאשר כל אחד משויך לקבוצה שאנו יכולים לתאר בצורה מלאה ולהבין באופן טוב יחסית. האופרטורים האוניטריים מתוארים לפי 7.55 אם $\mathbf{F} = \mathbf{C}$; האופרטורים החיוביים מתוארים על ידי משפטי הפירוק הספקטרלי, הממשי והמרוכב (7.29 ו־ 7.31).

ליתר דיוק, נתבונן במקרה $\mathbf{F} = \mathbf{C}$, ונניח ש־

$$T = S\sqrt{T^*T}$$

הוא פירוק פולרי של אופרטור $T \in \mathcal{L}(V)$, כאשר S הוא אופרטור אוניטרי. אזי קיים בסיס אורתונורמלי של V אשר ביחס אליו יש ל־ S מטריצה אלכסונית, וקיים בסיס אורתונורמלי של V אשר ביחס אליו יש ל־ $\sqrt{T^*T}$ מטריצה אלכסונית. **אזהרה:** ייתכן שלא קיים בסיס אורתונורמלי אחד אשר בו־זמנית מארגן את המטריצות של שני האופרטורים S ו־ $\sqrt{T^*T}$ לתוך אותן צורות אלכסוניות נחמדות—ייתכן כי S דורש בסיס אורתונורמלי אחד בעוד $\sqrt{T^*T}$ דורש בסיס אורתונורמלי אחר.

בכל זאת (עדיין תחת ההנחה ש־ $\mathbf{F} = \mathbf{C}$), אם T נורמלי, אזי ניתן לבחור בסיס אורתונורמלי של V כך שלשני האופרטורים, ל־ S ול־ $\sqrt{T^*T}$ יש מטריצות אלכסוניות ביחס לבסיס זה—ראו תרגיל 31. ההיפך גם הוא נכון: אם $T \in \mathcal{L}(V)$ ו־ $T = S\sqrt{T^*T}$ עבור אופרטור אוניטרי $S \in \mathcal{L}(V)$ כלשהו כך שלשני האופרטורים, ל־ S ול־ $\sqrt{T^*T}$ יש מטריצות אלכסוניות ביחס לאותו בסיס אורתונורמלי של V, אזי T נורמלי. זה מתקיים מאחר שיש ל־ T מטריצה אלכסונית ביחס אל אותו בסיס אורתונורמלי, מה שגורר ש־ T נורמלי [לפי השקילות של (א) ו־ (ג) ב־ 7.31].

The polar decomposition below is valid on both real and complex inner product spaces and for all operators on those spaces.

7.93 *polar decomposition*

Suppose $T \in \mathcal{L}(V)$. Then there exists a unitary operator $S \in \mathcal{L}(V)$ such that

$$T = S\sqrt{T^*T}.$$

Proof Let $s_1, \dots, s_m$ be the positive singular values of T, and let $e_1, \dots, e_m$ and $f_1, \dots, f_m$ be orthonormal lists in V such that

7.94 $$Tv = s_1\langle v, e_1\rangle f_1 + \cdots + s_m\langle v, e_m\rangle f_m$$

for every $v \in V$. Extend $e_1, \dots, e_m$ and $f_1, \dots, f_m$ to orthonormal bases $e_1, \dots, e_n$ and $f_1, \dots, f_n$ of V.

Define $S \in \mathcal{L}(V)$ by

$$Sv = \langle v, e_1\rangle f_1 + \cdots + \langle v, e_n\rangle f_n$$

for each $v \in V$. Then

$$\begin{aligned} \|Sv\|^2 &= \|\langle v, e_1\rangle f_1 + \cdots + \langle v, e_n\rangle f_n\|^2 \\ &= |\langle v, e_1\rangle|^2 + \cdots + |\langle v, e_n\rangle|^2 \\ &= \|v\|^2. \end{aligned}$$

Thus S is a unitary operator.

Applying T^* to both sides of 7.94 and then using the formula for T^* given by 7.77 shows that

$$T^*Tv = s_1^2\langle v, e_1\rangle e_1 + \cdots + s_m^2\langle v, e_m\rangle e_m$$

for every $v \in V$. Thus if $v \in V$, then

$$\sqrt{T^*T}v = s_1\langle v, e_1\rangle e_1 + \cdots + s_m\langle v, e_m\rangle e_m$$

because the operator that sends v to the right side of the equation above is a positive operator whose square equals T^*T. Now

$$\begin{aligned} S\sqrt{T^*T}v &= S(s_1\langle v, e_1\rangle e_1 + \cdots + s_m\langle v, e_m\rangle e_m) \\ &= s_1\langle v, e_1\rangle f_1 + \cdots + s_m\langle v, e_m\rangle f_m \\ &= Tv, \end{aligned}$$

where the last equation follows from 7.94. ∎

Exercise 27 shows that the unitary operator S produced in the proof above is as close as a unitary operator can be to T.

Alternative proofs of the polar decomposition directly use the spectral theorem, avoiding the singular value decomposition. However, the proof above seems cleaner than those alternative proofs.

הפירוק הפולרי להלן תקף במרחבי מכפלה פנימית ממשיים ומרוכבים כאחד, ועבור כל האופרטורים מעל מרחבים אלו.

7.93 פירוק פולרי

נניח ש־ $T \in \mathcal{L}(V)$. אזי קיים אופרטור אוניטרי $S \in \mathcal{L}(V)$ כך ש־
$$T = S\sqrt{T^*T}.$$

הוכחה יהיו $s_1, \dots, s_m$ הערכים הסינגולריים החיוביים של T, ותהיינה $e_1, \dots, e_m$ ו־ $f_1, \dots, f_m$ רשימות אורתונורמליות ב־ V כך ש־

$$Tv = s_1\langle v, e_1\rangle f_1 + \cdots + s_m\langle v, e_m\rangle f_m \tag{7.94}$$

לכל $v \in V$. נשלים את $e_1, \dots, e_m$ ו־ $f_1, \dots, f_m$ לכדי בסיסים אורתונורמליים $e_1, \dots, e_n$ ו־ $f_1, \dots, f_n$ של V.

נגדיר $S \in \mathcal{L}(V)$ על ידי

$$Sv = \langle v, e_1\rangle f_1 + \cdots + \langle v, e_n\rangle f_n$$

לכל $v \in V$. אזי

$$\begin{aligned}\|Sv\|^2 &= \|\langle v, e_1\rangle f_1 + \cdots + \langle v, e_n\rangle f_n\|^2 \\ &= |\langle v, e_1\rangle|^2 + \cdots + |\langle v, e_n\rangle|^2 \\ &= \|v\|^2.\end{aligned}$$

לפיכך S אופרטור אוניטרי.

הפעלת T^* על שני האגפים של 7.94 ואז שימוש בנוסחה עבור T^* הנתונה ב־ 7.77 מראה לנו כי

$$T^*Tv = s_1^2\langle v, e_1\rangle e_1 + \cdots + s_m^2\langle v, e_m\rangle e_m$$

לכל $v \in V$. לפיכך אם $v \in V$, אזי

$$\sqrt{T^*T}v = s_1\langle v, e_1\rangle e_1 + \cdots + s_m\langle v, e_m\rangle e_m$$

מאחר שהאופרטור השולח את v אל האגף הימני של המשוואה לעיל הוא אופרטור חיובי שריבועו שווה ל־ T^*T. כעת

$$\begin{aligned}S\sqrt{T^*T}v &= S(s_1\langle v, e_1\rangle e_1 + \cdots + s_m\langle v, e_m\rangle e_m) \\ &= s_1\langle v, e_1\rangle f_1 + \cdots + s_m\langle v, e_m\rangle f_m \\ &= Tv,\end{aligned}$$

כאשר המשוואה האחרונה נובעת מ־ 7.94. ∎

תרגיל 27 מראה שהאופרטור האוניטרי S שהופק בהוכחה לעיל הוא הקרוב ביותר האפשרי ל־ T מבין האופרטורים האוניטריים.

הוכחות חלופיות לפירוק הפולרי משתמשות במשפט הפירוק הספקטרלי באופן ישיר, תוך הימנעות מהפירוק לערכים סינגולריים. בכל זאת, ההוכחה לעיל נראית נקייה יותר מאותן הוכחות חלופיות.

Operators Applied to Ellipsoids and Parallelepipeds

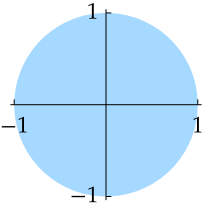
7.95 definition: *ball, B*

The *ball* in V of radius 1 centered at 0, denoted by B, is defined by

$$B = \{v \in V : \|v\| < 1\}.$$

If $\dim V = 2$, the word *disk* is sometimes used instead of *ball*. However, using *ball* in all dimensions is less confusing. Similarly, if $\dim V = 2$, then the word *ellipse* is sometimes used instead of the word *ellipsoid* that we are about to define. Again, using *ellipsoid* in all dimensions is less confusing.

You can think of the ellipsoid defined below as obtained by starting with the ball B and then stretching by a factor of s_k along each f_k-axis.



The ball B in $\mathbf{R}^2$.

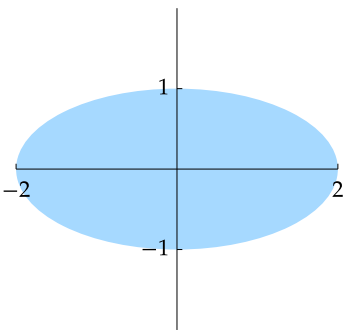
7.96 definition: *ellipsoid, $E(s_1 f_1, \dots, s_n f_n)$, principal axes*

Suppose that $f_1, \dots, f_n$ is an orthonormal basis of V and $s_1, \dots, s_n$ are positive numbers. The *ellipsoid* $E(s_1 f_1, \dots, s_n f_n)$ with *principal axes* $s_1 f_1, \dots, s_n f_n$ is defined by

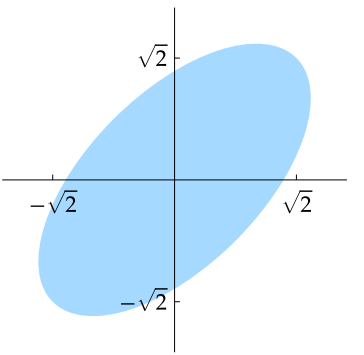
$$E(s_1 f_1, \dots, s_n f_n) = \left\{v \in V : \frac{|\langle v, f_1\rangle|^2}{s_1^2} + \cdots + \frac{|\langle v, f_n\rangle|^2}{s_n^2} < 1\right\}.$$

The ellipsoid notation $E(s_1 f_1, \dots, s_n f_n)$ does not explicitly include the inner product space V, even though the definition above depends on V. However, the inner product space V should be clear from the context and also from the requirement that $f_1, \dots, f_n$ be an orthonormal basis of V.

7.97 example: *ellipsoids*



The ellipsoid $E(2f_1, f_2)$ in $\mathbf{R}^2$, where f_1, f_2 is the standard basis of $\mathbf{R}^2$.



The ellipsoid $E(2f_1, f_2)$ in $\mathbf{R}^2$, where $f_1 = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$ and $f_2 = \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$.

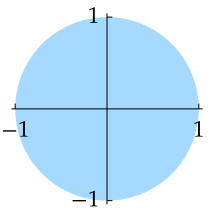
יישום אופרטורים על אליפסואידים ומקבילונים

7.95 הגדרה: כדור, B

הכדור ב־ V שרדיוסו 1 ושמרכזו ב־ 0, מסומן על ידי B, מוגדר על ידי
$$B = \{v \in V : \|v\| < 1\}.$$

אם $\dim V = 2$, לעיתים נעשה שימוש במונח עיגול במקום כדור. בכל זאת, שימוש בכדור בכל הממדים הוא פחות מבלבל. באופן דומה, אם $\dim V = 2$, לעיתים נעשה שימוש במונח אליפסה במקום המונח אליפסואיד שאנו עומדים להגדיר. שוב, השימוש באליפסואיד בכל הממדים פחות מבלבל.

תוכלו לחשוב על האליפסואיד המוגדר להלן כזה שהושג בכך שמתחילים עם הכדור B ואז מותחים פי s_k לאורך כל ציר f_k.



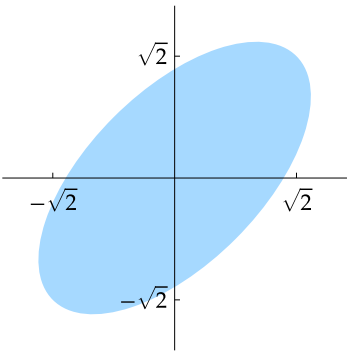
הכדור B ב־ $\mathbf{R}^2$.

7.96 הגדרה: אליפסואיד, $E(s_1 f_1, \dots, s_n f_n)$, צירים ראשיים

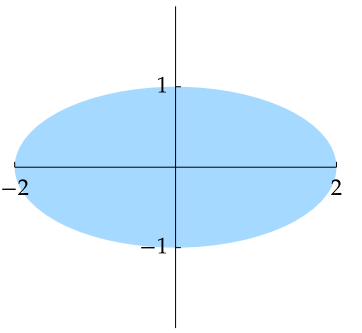
נניח ש־ $f_1, \dots, f_n$ בסיס אורתונורמלי של V ו־ $s_1, \dots, s_n$ מספרים חיוביים. האליפסואיד $E(s_1 f_1, \dots, s_n f_n)$ עם צירים ראשיים $s_1 f_1, \dots, s_n f_n$ מוגדר על ידי
$$E(s_1 f_1, \dots, s_n f_n) = \left\{v \in V : \frac{|\langle v, f_1\rangle|^2}{s_1^2} + \cdots + \frac{|\langle v, f_n\rangle|^2}{s_n^2} < 1\right\}.$$

סימון האליפסואיד $E(s_1 f_1, \dots, s_n f_n)$ אינו כולל באופן מפורש את מרחב המכפלה הפנימית V, אף שההגדרה לעיל תלויה ב־ V. בכל זאת, מרחב המכפלה הפנימית V אמור להיות ברור מההקשר ובנוסף מהדרישה ש־ $f_1, \dots, f_n$ יהיה בסיס אורתונורמלי של V.

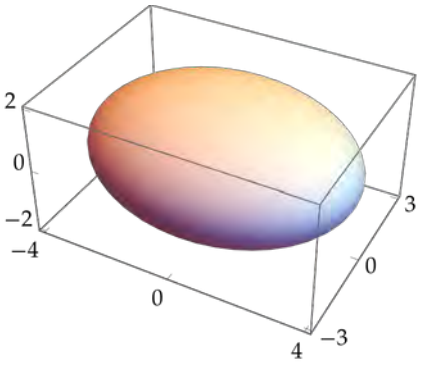
7.97 דוגמה: אליפסואידים



האליפסואיד $E(2f_1, f_2)$ ב־ $\mathbf{R}^2$, כאשר $f_1 = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$ ו־ $f_2 = \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$.



האליפסואיד $E(2f_1, f_2)$ ב־ $\mathbf{R}^2$, כאשר f_1, f_2 הוא הבסיס הסטנדרטי של $\mathbf{R}^2$.



The ellipsoid $E(4f_1, 3f_2, 2f_3)$ in $\mathbf{R}^3$, where f_1, f_2, f_3 is the standard basis of $\mathbf{R}^3$.

The ellipsoid $E(f_1, \dots, f_n)$ equals the ball B in V for every orthonormal basis $f_1, \dots, f_n$ of V [by Parseval's identity 6.30(b)].

7.98 notation: $T(\Omega)$

For T a function defined on V and $\Omega \subseteq V$, define $T(\Omega)$ by

$$T(\Omega) = \{Tv : v \in \Omega\}.$$

Thus if T is a function defined on V, then $T(V) = \text{range } T$.

The next result states that every invertible operator $T \in \mathcal{L}(V)$ maps the ball B in V onto an ellipsoid in V. The proof shows that the principal axes of this ellipsoid come from the singular value decomposition of T.

7.99 *invertible operator takes ball to ellipsoid*

Suppose $T \in \mathcal{L}(V)$ is invertible. Then T maps the ball B in V onto an ellipsoid in V.

Proof Suppose T has singular value decomposition

7.100 $$Tv = s_1\langle v, e_1\rangle f_1 + \cdots + s_n\langle v, e_n\rangle f_n$$

for all $v \in V$; here $s_1, \dots, s_n$ are the singular values of T and $e_1, \dots, e_n$ and $f_1, \dots, f_n$ are both orthonormal bases of V. We will show that $T(B) = E(s_1 f_1, \dots, s_n f_n)$.

First suppose $v \in B$. Because T is invertible, none of the singular values $s_1, \dots, s_n$ equals 0 (see 7.68). Thus 7.100 implies that

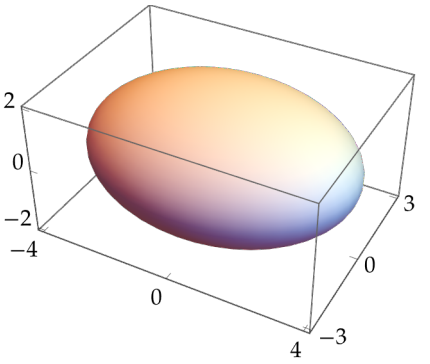
$$\frac{|\langle Tv, f_1\rangle|^2}{s_1^2} + \cdots + \frac{|\langle Tv, f_n\rangle|^2}{s_n^2} = |\langle v, e_1\rangle|^2 + \cdots + |\langle v, e_n\rangle|^2 < 1.$$

Thus $Tv \in E(s_1 f_1, \dots, s_n f_n)$. Hence $T(B) \subseteq E(s_1 f_1, \dots, s_n f_n)$.

To prove inclusion in the other direction, now suppose $w \in E(s_1 f_1, \dots, s_n f_n)$. Let

$$v = \frac{\langle w, f_1\rangle}{s_1} e_1 + \cdots + \frac{\langle w, f_n\rangle}{s_n} e_n.$$

Then $\|v\| < 1$ and 7.100 implies that $Tv = \langle w, f_1\rangle f_1 + \cdots + \langle w, f_n\rangle f_n = w$. Thus $T(B) \supseteq E(s_1 f_1, \dots, s_n f_n)$.



האליפסואיד $E(4f_1, 3f_2, 2f_3)$ ב־ $\mathbf{R}^3$, כאשר f_1, f_2, f_3 הבסיס הסטנדרטי של $\mathbf{R}^3$.

האליפסואיד $E(f_1, \dots, f_n)$ שווה לכדור B ב־ V לכל בסיס אורתונורמלי $f_1, \dots, f_n$ של V [לפי שוויון פרסבל 6.30(ב)].

7.98 סימון: $T(\Omega)$

עבור T פונקציה המוגדרת מעל V ו־ $\Omega \subseteq V$, נגדיר $T(\Omega)$ על ידי

$$T(\Omega) = \{Tv : v \in \Omega\}.$$

לפיכך אם T פונקציה המוגדרת מעל V, אזי $T(V) = \operatorname{range} T$.

המסקנה הבאה קובעת שכל אופרטור הפיך $T \in \mathcal{L}(V)$ מעתיק את הכדור B ב־ V על אליפסואיד ב־ V. ההוכחה מראה כי הצירים הראשיים של אליפסואיד זה נובעים מהפירוק לערכים סינגולריים של T.

7.99 אופרטור הפיך מעתיק כדור לאליפסואיד

נניח ש־ $T \in \mathcal{L}(V)$ הפיך. אזי T מעתיק את הכדור B ב־ V על אליפסואיד ב־ V.

הוכחה נניח שיש ל־ T פירוק לערכים סינגולריים

$$Tv = s_1\langle v, e_1\rangle f_1 + \cdots + s_n\langle v, e_n\rangle f_n \tag{7.100}$$

לכל $v \in V$. כאן $s_1, \dots, s_n$ הם הערכים הסינגולריים של T ו־ $e_1, \dots, e_n$ ו־ $f_1, \dots, f_n$ שניהם בסיסים אורתונורמליים של V. אנו נראה כי $T(B) = E(s_1 f_1, \dots, s_n f_n)$.

תחילה נניח כי $v \in B$. מאחר ש־ T הפיך, אף אחד מהערכים הסינגולריים $s_1, \dots, s_n$ אינו שווה 0 (ראו 7.68). לפיכך 7.100 גורר ש־

$$\frac{|\langle Tv, f_1\rangle|^2}{s_1^2} + \cdots + \frac{|\langle Tv, f_n\rangle|^2}{s_n^2} = |\langle v, e_1\rangle|^2 + \cdots + |\langle v, e_n\rangle|^2 < 1.$$

לפיכך $Tv \in E(s_1 f_1, \dots, s_n f_n)$. לכן $T(B) \subseteq E(s_1 f_1, \dots, s_n f_n)$.

להוכחת ההכלה בכיוון השני, כעת נניח ש־ $w \in E(s_1 f_1, \dots, s_n f_n)$. יהי

$$v = \frac{\langle w, f_1\rangle}{s_1} e_1 + \cdots + \frac{\langle w, f_n\rangle}{s_n} e_n.$$

אזי $\|v\| < 1$ ו־ 7.100 גוררים ש־ $Tv = \langle w, f_1\rangle f_1 + \cdots + \langle w, f_n\rangle f_n = w$. לפיכך $T(B) \supseteq E(s_1 f_1, \dots, s_n f_n)$. ■

אלגברה ליניארית בדרך הישר, מהדורה רביעית, מחבר: שלדון אקסלר
Linear Algebra Done Right, fourth edition, by Sheldon Axler

We now use the previous result to show that invertible operators take all ellipsoids, not just the ball of radius 1, to ellipsoids.

7.101 *invertible operator takes ellipsoids to ellipsoids*

Suppose $T \in \mathcal{L}(V)$ is invertible and E is an ellipsoid in V. Then $T(E)$ is an ellipsoid in V.

Proof There exist an orthonormal basis $f_1, \dots, f_n$ of V and positive numbers $s_1, \dots, s_n$ such that $E = E(s_1 f_1, \dots, s_n f_n)$. Define $S \in \mathcal{L}(V)$ by

$$S(a_1 f_1 + \cdots + a_n f_n) = a_1 s_1 f_1 + \cdots + a_n s_n f_n.$$

Then S maps the ball B of V onto E, as you can verify. Thus

$$T(E) = T(S(B)) = (TS)(B).$$

The equation above and 7.99, applied to TS, show that $T(E)$ is an ellipsoid in V. ∎

Recall (see 3.95) that if $u \in V$ and $\Omega \subseteq V$ then $u + \Omega$ is defined by

$$u + \Omega = \{u + w : w \in \Omega\}.$$

Geometrically, the sets Ω and $u + \Omega$ look the same, but they are in different locations.

In the following definition, if $\dim V = 2$ then the word *parallelogram* is often used instead of *parallelepiped*.

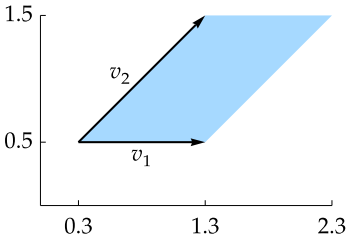
7.102 definition: $P(v_1, \dots, v_n)$, *parallelepiped*

Suppose $v_1, \dots, v_n$ is a basis of V. Let

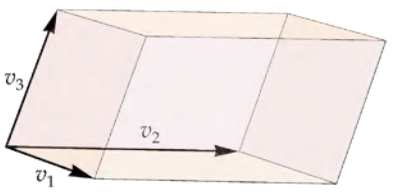
$$P(v_1, \dots, v_n) = \{a_1 v_1 + \cdots + a_n v_n : a_1, \dots, a_n \in (0, 1)\}.$$

A *parallelepiped* is a set of the form $u + P(v_1, \dots, v_n)$ for some $u \in V$. The vectors $v_1, \dots, v_n$ are called the *edges* of this parallelepiped.

7.103 example: *parallelepipeds*



The parallelepiped $(0.3, 0.5) + P((1, 0), (1, 1))$ *in* $\mathbf{R}^2$.



A parallelepiped in $\mathbf{R}^3$.

כעת אנו משתמשים במסקנה הקודמת כדי להראות כי אופרטורים הפיכים מעתיקים את כל האליפסואידים, לא רק את הכדור שרדיוסו 1, אל אליפסואידים.

7.101 אופרטור הפיך מעתיק אליפסואידים אל אליפסואידים

נניח ש־ $T \in \mathcal{L}(V)$ הפיך ו־ E אליפסואיד ב־ V. אזי $T(E)$ אליפסואיד ב־ V.

הוכחה קיימים בסיס אורתונורמלי $f_1, \dots, f_n$ של V ומספרים חיוביים $s_1, \dots, s_n$ כך ש־ $E = E(s_1 f_1, \dots, s_n f_n)$. נגדיר $S \in \mathcal{L}(V)$ על ידי

$$S(a_1 f_1 + \cdots + a_n f_n) = a_1 s_1 f_1 + \cdots + a_n s_n f_n.$$

אזי S מעתיק את הכדור B של V על E, כפי שבאפשרותכם לוודא. לפיכך

$$T(E) = T(S(B)) = (TS)(B).$$

המשוואה לעיל ו־ 7.99, מיושמות על TS, מראות כי $T(E)$ אליפסואיד ב־ V. ■

נזכיר (ראו 3.95) כי אם $u \in V$ ו־ $\Omega \subseteq V$ אזי $u + \Omega$ מוגדר על ידי

$$u + \Omega = \{u + w : w \in \Omega\}.$$

מבחינה גאומטרית, הקבוצות Ω ו־ $u + \Omega$ נראות זהות, אך הן במיקומים שונים.

בהגדרה הבאה, אם $\dim V = 2$ אזי לעיתים קרובות נעשה שימוש במילה מקבילית במקום במקבילון.

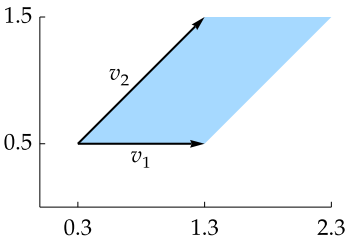
7.102 הגדרה: מקבילון, $P(v_1, \dots, v_n)$

נניח ש־ $v_1, \dots, v_n$ בסיס של V. יהי

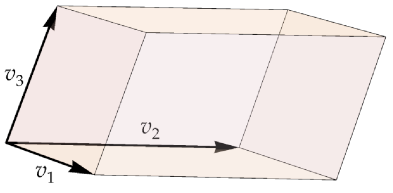
$$P(v_1, \dots, v_n) = \{a_1 v_1 + \cdots + a_n v_n : a_1, \dots, a_n \in (0, 1)\}.$$

מקבילון הוא קבוצה מהצורה $u + P(v_1, \dots, v_n)$ עבור $u \in V$ כלשהו. הווקטורים $v_1, \dots, v_n$ נקראים הצלעות של מקבילון זה.

7.103 דוגמה: מקבילונים



המקבילון $(0.3, 0.5) + P((1, 0), (1, 1))$ ב־ $\mathbf{R}^2$.



מקבילון ב־ $\mathbf{R}^3$.

7.104 *invertible operator takes parallelepipeds to parallelepipeds*

Suppose $u \in V$, $v_1, \dots, v_n$ is a basis of V, and $T \in \mathcal{L}(V)$ is invertible. Then

$$T(u + P(v_1, \dots, v_n)) = Tu + P(Tv_1, \dots, Tv_n).$$

Proof Because T is invertible, the list $Tv_1, \dots, Tv_n$ is a basis of V. The linearity of T implies that

$$T(u + a_1v_1 + \cdots + a_nv_n) = Tu + a_1Tv_1 + \cdots + a_nTv_n$$

for all $a_1, \dots, a_n \in (0, 1)$. Thus $T(u + P(v_1, \dots, v_n)) = Tu + P(Tv_1, \dots, Tv_n)$. ∎

Just as the rectangles are distinguished among the parallelograms in $\mathbf{R}^2$, we give a special name to the parallelepipeds in V whose defining edges are orthogonal to each other.

7.105 definition: *box*

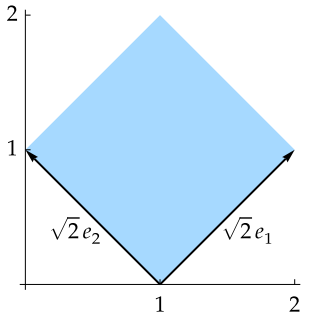
A *box* in V is a set of the form

$$u + P(r_1e_1, \dots, r_ne_n),$$

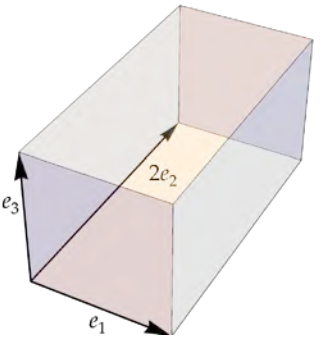
where $u \in V$ and $r_1, \dots, r_n$ are positive numbers and $e_1, \dots, e_n$ is an orthonormal basis of V.

Note that in the special case of $\mathbf{R}^2$ each box is a rectangle, but the terminology *box* can be used in all dimensions.

7.106 example: *boxes*



The box $(1, 0) + P(\sqrt{2}e_1, \sqrt{2}e_2)$, *where* $e_1 = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$ *and* $e_2 = \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$.



The box $P(e_1, 2e_2, e_3)$, *where* e_1, e_2, e_3 *is the standard basis of* $\mathbf{R}^3$.

Suppose $T \in \mathcal{L}(V)$ is invertible. Then T maps every parallelepiped in V to a parallelepiped in V (by 7.104). In particular, T maps every box in V to a parallelepiped in V. This raises the question of whether T maps some boxes in V to boxes in V. The following result answers this question, with the help of the singular value decomposition.

7.104 אופרטור הפיך מעתיק מקבילונים אל מקבילונים

נניח ש־ $u \in V$ ו־ $v_1, \dots, v_n$ בסיס של V, ו־ $T \in \mathcal{L}(V)$ הפיך. אזי
$$T(u + P(v_1, \dots, v_n)) = Tu + P(Tv_1, \dots, Tv_n).$$

הוכחה מאחר ש־ T הפיך, הרשימה $Tv_1, \dots, Tv_n$ בסיס של V. הליניאריות של T גוררת ש־
$$T(u + a_1 v_1 + \cdots + a_n v_n) = Tu + a_1 Tv_1 + \cdots + a_n Tv_n$$
לכל $a_1, \dots, a_n \in (0,1)$. לפיכך $T(u + P(v_1, \dots, v_n)) = Tu + P(Tv_1, \dots, Tv_n)$. ■

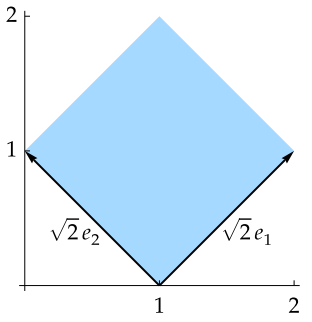
בדומה למלבנים, המובחנים מבין המקביליות ב־ $\mathbf{R}^2$, אנו מעניקים שם מיוחד למקבילון ב־ V שהצלעות המגדירות שלו אורתוגונליות זו לזו.

7.105 הגדרה: תיבה

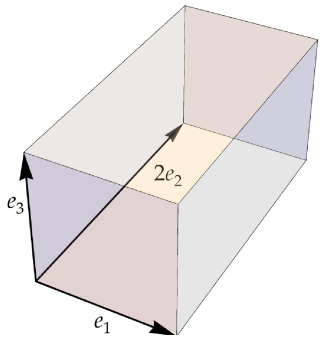
תיבה ב־ V היא קבוצה מהצורה
$$u + P(r_1 e_1, \dots, r_n e_n),$$
כאשר $u \in V$ ו־ $r_1, \dots, r_n$ מספרים חיוביים ו־ $e_1, \dots, e_n$ בסיס אורתונורמלי של V.

שימו לב שבמקרה הפרטי של $\mathbf{R}^2$ כל תיבה היא מלבן, אך ניתן להשתמש במונח תיבה בכל הממדים.

7.106 דוגמה: תיבות



התיבה $P(\sqrt{2}e_1, \sqrt{2}e_2) + (1,0)$, כאשר $e_1 = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$ ו־ $e_2 = \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$.



התיבה $P(e_1, 2e_2, e_3)$, כאשר e_1, e_2, e_3 הוא הבסיס הסטנדרטי של $\mathbf{R}^3$.

נניח ש־ $T \in \mathcal{L}(V)$ הפיך. אזי T מעתיק כל מקבילון ב־ V אל מקבילון ב־ V (לפי 7.104). בפרט, T מעתיק כל תיבה ב־ V אל מקבילון ב־ V. הדבר מעלה את השאלה האם T מעתיק תיבות כלשהן ב־ V אל תיבות ב־ V. המסקנה הבאה עונה על שאלה זו, בסיוע הפירוק לערכים סינגולריים.

7.107 *every invertible operator takes some boxes to boxes*

Suppose $T \in \mathcal{L}(V)$ is invertible. Suppose T has singular value decomposition

$$Tv = s_1\langle v, e_1\rangle f_1 + \cdots + s_n\langle v, e_n\rangle f_n,$$

where $s_1, \ldots, s_n$ are the singular values of T and $e_1, \ldots, e_n$ and $f_1, \ldots, f_n$ are orthonormal bases of V and the equation above holds for all $v \in V$. Then T maps the box $u + P(r_1e_1, \ldots, r_ne_n)$ onto the box $Tu + P(r_1s_1f_1, \ldots, r_ns_nf_n)$ for all positive numbers $r_1, \ldots, r_n$ and all $u \in V$.

Proof If $a_1, \ldots, a_n \in (0, 1)$ and $r_1, \ldots, r_n$ are positive numbers and $u \in V$, then

$$T(u + a_1r_1e_1 + \cdots + a_nr_ne_n) = Tu + a_1r_1s_1f_1 + \cdots + a_nr_ns_nf_n.$$

Thus $T(u + P(r_1e_1, \ldots, r_ne_n)) = Tu + P(r_1s_1f_1, \ldots, r_ns_nf_n)$. ∎

Volume via Singular Values

Our goal in this subsection is to understand how an operator changes the volume of subsets of its domain. Because notions of volume belong to analysis rather than to linear algebra, we will work only with an intuitive notion of volume. Our intuitive approach to volume can be converted into appropriate correct definitions, correct statements, and correct proofs using the machinery of analysis.

Our intuition about volume works best in real inner product spaces. Thus the assumption that $\mathbf{F} = \mathbf{R}$ will appear frequently in the rest of this subsection.

If $\dim V = n$, then by *volume* we will mean n-dimensional volume. You should be familiar with this concept in $\mathbf{R}^3$. When $n = 2$, this is usually called area instead of volume, but for consistency we use the word volume in all dimensions. The most fundamental intuition about volume is that the volume of a box (whose defining edges are by definition orthogonal to each other) is the product of the lengths of the defining edges. Thus we make the following definition.

7.108 definition: *volume of a box*

Suppose $\mathbf{F} = \mathbf{R}$. If $u \in V$ and $r_1, \ldots, r_n$ are positive numbers and $e_1, \ldots, e_n$ is an orthonormal basis of V, then

$$\text{volume}(u + P(r_1e_1, \ldots, r_ne_n)) = r_1 \times \cdots \times r_n.$$

The definition above agrees with the familiar formulas for the area (which we are calling the volume) of a rectangle in $\mathbf{R}^2$ and for the volume of a box in $\mathbf{R}^3$. For example, the first box in Example 7.106 has two-dimensional volume (or area) 2 because the defining edges of that box have length $\sqrt{2}$ and $\sqrt{2}$. The second box in Example 7.106 has three-dimensional volume 2 because the defining edges of that box have length 1, 2, and 1.

7.107 כל אופרטור הפיך מעתיק תיבות כלשהן אל תיבות

נניח ש־ $T \in \mathcal{L}(V)$ הפיך. נניח שיש ל־ T פירוק לערכים סינגולריים

$$Tv = s_1\langle v, e_1\rangle f_1 + \cdots + s_n\langle v, e_n\rangle f_n,$$

כאשר $s_1, \ldots, s_n$ הם הערכים הסינגולריים של T ו־ $e_1, \ldots, e_n$ ו־ $f_1, \ldots, f_n$ הם בסיסים אורתונורמליים של V והמשוואה לעיל מתקיימת לכל $v \in V$. אזי T מעתיק את התיבה $u + P(r_1e_1, \ldots, r_ne_n)$ על התיבה $Tu + P(r_1s_1f_1, \ldots, r_ns_nf_n)$ לכל המספרים החיוביים $r_1, \ldots, r_n$ ולכל $u \in V$.

הוכחה אם $a_1, \ldots, a_n \in (0, 1)$ מספרים חיוביים ו־ $r_1, \ldots, r_n$ ו־ $u \in V$, אזי

$$T(u + a_1r_1e_1 + \cdots + a_nr_ne_n) = Tu + a_1r_1s_1f_1 + \cdots + a_nr_ns_nf_n.$$

לפיכך $T(u + P(r_1e_1, \ldots, r_ne_n)) = Tu + P(r_1s_1f_1, \ldots, r_ns_nf_n)$. ■

נפח באמצעות ערכים סינגולריים

מטרתנו בתת־יחידה זו של הפרק היא להבין כיצד אופרטור משנה את הנפח של תתי־קבוצות בתחומו. מאחר שמושגים של נפח שייכים לאנליזה ולא לאלגברה ליניארית, אנו נעבוד במונחים אינטואיטיביים של נפח. ניתן להמיר את גישתנו האינטואיטיבית לנפח להגדרות נכונות ומתאימות, משפטים נכונים והוכחות נכונות, תוך שימוש בכלי האנליזה.

האינטואיציה שלנו לגבי נפח עובדת באופן מיטבי במרחבי מכפלה פנימית ממשיים. לפיכך ההנחה ש־ $\mathbf{F} = \mathbf{R}$ תופיע לעיתים קרובות בתת־יחידה זו.

אם $\dim V = n$, אזי כשאנו מציינים נפח אנו מתכוונים לנפח n־ממדי. אתם אמורים להכיר את רעיון זה כבר ב־ $\mathbf{R}^3$. כאשר $n = 2$, בדרך כלל הדבר נקרא שטח במקום נפח, אך לשם העקביות אנו נשתמש במילה נפח בכל הממדים. האינטואיציה הבסיסית ביותר לגבי נפח היא שנפח של תיבה (זו אשר הצלעות המגדירות שלה לפי הגדרה אורתוגונליות זו לזו) הוא המכפלה של אורכי הצלעות המגדירות. לפיכך נציג את ההגדרה הבאה.

7.108 הגדרה: נפח תיבה, volume

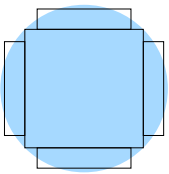
נניח ש־ $\mathbf{F} = \mathbf{R}$. אם $u \in V$ ו־ $r_1, \ldots, r_n$ מספרים חיוביים ו־ $e_1, \ldots, e_n$ בסיס אורתונורמלי של V, אזי

$$\text{volume}(u + P(r_1e_1, \ldots, r_ne_n)) = r_1 \times \cdots \times r_n.$$

ההגדרה לעיל תואמת את הנוסחאות המוכרות עבור חישוב השטח (שלו אנו קוראים כעת נפח) של מלבן ב־ $\mathbf{R}^2$ ושל נפח תיבה ב־ $\mathbf{R}^3$. לדוגמה, לתיבה הראשונה בדוגמה 7.106 נפח דו־ממדי (או שטח) שערכו 2 מאחר שאורכי הצלעות המגדירות של תיבה זו הם $\sqrt{2}$ ו־ $\sqrt{2}$. נפח התיבה השנייה בדוגמה 7.106 הוא תלת־ממדי וערכו 2 מאחר שאורכי הצלעות המגדירות של תיבה זו הם 2, 1 ו־ 1.

To define the volume of a subset of V, approximate the subset by a finite collection of disjoint boxes, and then add up the volumes of the approximating collection of boxes. As we approximate a subset of V more accurately by disjoint unions of more boxes, we get a better approximation to the volume.

These ideas should remind you of how the Riemann integral is defined by approximating the area under a curve by a disjoint collection of rectangles. This discussion leads to the following nonrigorous but intuitive definition.



Volume of this ball ≈ sum of the volumes of the five boxes.

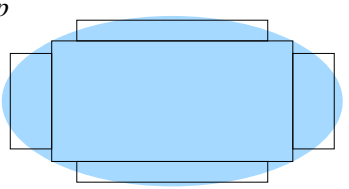
7.109 definition: *volume*

Suppose $\mathbf{F} = \mathbf{R}$ and $\Omega \subseteq V$. Then the *volume* of Ω, denoted by volume Ω, is approximately the sum of the volumes of a collection of disjoint boxes that approximate Ω.

We are ignoring many reasonable questions by taking an intuitive approach to volume. For example, if we approximate Ω by boxes with respect to one basis, do we get the same volume if we approximate Ω by boxes with respect to a different basis? If Ω_1 and Ω_2 are disjoint subsets of V, is volume$(\Omega_1 \cup \Omega_2) =$ volume Ω_1 + volume Ω_2? Provided that we consider only reasonably nice subsets of V, techniques of analysis show that both these questions have affirmative answers that agree with our intuition about volume.

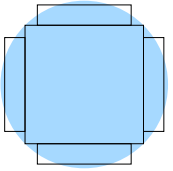
7.110 example: *volume change by a linear map*

Suppose that $T \in \mathcal{L}(\mathbf{R}^2)$ is defined by $Tv = 2\langle v, e_1\rangle e_1 + \langle v, e_2\rangle e_2$, where e_1, e_2 is the standard basis of $\mathbf{R}^2$. This linear map stretches vectors along the e_1-axis by a factor of 2 and leaves vectors along the e_2-axis unchanged. The ball approximated by five boxes above gets mapped by T to the ellipsoid shown here. Each of the five boxes in the original figure gets mapped to a box of twice the width and the same height as in the original figure. Hence each box gets mapped to a box of twice the volume (area) as in the original figure. The sum of the volumes of the five new boxes approximates the volume of the ellipsoid. Thus T changes the volume of the ball by a factor of 2.



Each box here has twice the width and the same height as the boxes in the previous figure.

In the example above, T maps boxes with respect to the basis e_1, e_2 to boxes with respect to the same basis; thus we can see how T changes volume. In general, an operator maps boxes to parallelepipeds that are not boxes. However, if we choose the right basis (coming from the singular value decomposition!), then boxes with respect to that basis get mapped to boxes with respect to a possibly different basis, as shown in 7.107. This observation leads to a natural proof of the following result.



הנפח של כדור זה ≈ סכום הנפחים של חמש התיבות.

להגדרת הנפח של תת־קבוצה של V, בצעו קירוב של התת־קבוצה באמצעות אוסף בגודל סופי של תיבות זרות, ולאחר מכן סכמו את הנפחים של תיבות הקירוב. ככל שהקירוב של התת־קבוצה של V יהיה מדויק יותר בכך שנבצע איחוד זר של עוד תיבות, נקבל קירוב טוב יותר לנפח.

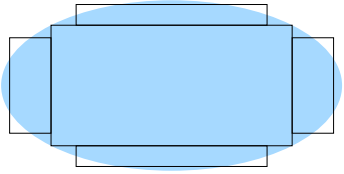
רעיונות אלו אמורים להזכיר לכם את אינטגרל רימן המוגדר על ידי קירוב השטח תחת העקום על ידי אוסף זר של מלבנים. דיון זה מוביל אל ההגדרה הלא ריגורוזית אך אינטואיטיבית הבאה.

7.109 הגדרה: נפח

נניח ש־ $\mathbf{F} = \mathbf{R}$ ו־ $\Omega \subseteq V$. אזי הנפח של Ω, מסומן על ידי $\operatorname{volume}\Omega$, הוא בקירוב סכום הנפחים של אוסף של תיבות זרות המקרבות את Ω.

אנו מתעלמים משאלות הגיוניות רבות בכך שאנו נוקטים בגישה האינטואיטיבית לנפח. לדוגמה, אם נקרב את Ω באמצעות תיבות ביחס לבסיס אחד, האם נקבל את אותו הנפח אם נקרב את Ω באמצעות תיבות ביחס לבסיס אחר? אם Ω_1 ו־ Ω_2 שתי תתי־קבוצות זרות של V, האם $\operatorname{volume}(\Omega_1 \cup \Omega_2) = \operatorname{volume}\Omega_1 + \operatorname{volume}\Omega_2$? בהינתן שאנו לוקחים בחשבון רק תתי־קבוצות נחמדות של V, טכניקות של אנליזה מראות כי לשתי שאלות אלו תשובות חיוביות העולות בקנה אחד עם האינטואיציה שלנו בנוגע לנפח.

7.110 דוגמה: שינוי נפח באמצעות העתקה ליניארית



ביחס לתיבות באיור הקודם, רוחב התיבות כאן הוא כפול בעוד גובה התיבות שווה.

נניח ש־ $T \in \mathcal{L}(\mathbf{R}^2)$ מוגדר על ידי $Tv = 2\langle v, e_1\rangle e_1 + \langle v, e_2\rangle e_2$, כאשר e_1, e_2 הוא הבסיס הסטנדרטי של $\mathbf{R}^2$. העתקה ליניארית זו מותחת את הווקטורים לאורך ציר ה־ e_1 בפקטור של 2 ומותירה את הווקטורים לאורך ציר ה־ e_2 ללא שינוי. הכדור המקורב על ידי חמש התיבות לעיל מועתק על ידי T אל האליפסואיד המוצג כאן. כל אחת מחמש התיבות באיור המקורי מועתקת אל תיבה ברוחב כפול ובגובה זהה לזה שבאיור המקורי. לכן כל תיבה מועתקת לתיבה שנפחה (שטחה) כפול מזה שבאיור המקורי. סכום הנפחים של חמש התיבות החדשות הוא הקירוב לנפח האליפסואיד. לפיכך T משנה את נפח הכדור בפקטור של 2.

בדוגמה לעיל, T מעתיקה תיבות ביחס אל הבסיס e_1, e_2 אל תיבות ביחס לאותו הבסיס; לפיכך אנו יכולים לראות כיצד T משנה נפח. באופן כללי, אופרטור מעתיק תיבות אל מקבילונים שאינם תיבות. בכל זאת, אם נבחר את הבסיס הנכון (זה המגיע מהפירוק לערכים סינגולריים!), אזי תיבות ביחס לבסיס זה מועתקות לתיבות ביחס לבסיס זהה או שונה, כפי שמוצג ב־ 7.107. התבוננות זו מובילה להוכחה טבעית של המסקנה הבאה.

7.111 *volume changes by a factor of the product of the singular values*

Suppose $\mathbf{F} = \mathbf{R}$, $T \in \mathcal{L}(V)$ is invertible, and $\Omega \subseteq V$. Then

$$\text{volume } T(\Omega) = (\text{product of singular values of } T)(\text{volume } \Omega).$$

Proof Suppose T has singular value decomposition

$$Tv = s_1\langle v, e_1\rangle f_1 + \cdots + s_n\langle v, e_n\rangle f_n$$

for all $v \in V$, where $e_1, \ldots, e_n$ and $f_1, \ldots, f_n$ are orthonormal bases of V.

Approximate Ω by boxes of the form $u + P(r_1 e_1, \ldots, r_n e_n)$, which have volume $r_1 \times \cdots \times r_n$. The operator T maps each box $u + P(r_1 e_1, \ldots, r_n e_n)$ onto the box $Tu + P(r_1 s_1 f_1, \ldots, r_n s_n f_n)$, which has volume $(s_1 \times \cdots \times s_n)(r_1 \times \cdots \times r_n)$.

The operator T maps a collection of boxes that approximate Ω onto a collection of boxes that approximate $T(\Omega)$. Because T changes the volume of each box in a collection that approximates Ω by a factor of $s_1 \times \cdots \times s_n$, the linear map T changes the volume of Ω by the same factor. ∎

Suppose $T \in \mathcal{L}(V)$. As we will see when we get to determinants, the product of the singular values of T equals $|\det T|$; see 9.60 and 9.61.

Properties of an Operator as Determined by Its Eigenvalues

We conclude this chapter by presenting the table below. The context of this table is a finite-dimensional complex inner product space. The first column of the table shows a property that a normal operator on such a space might have. The second column of the table shows a subset of $\mathbf{C}$ such that the operator has the corresponding property if and only if all eigenvalues of the operator lie in the specified subset. For example, the first row of the table states that a normal operator is invertible if and only if all its eigenvalues are nonzero (this first row is the only one in the table that does not need the hypothesis that the operator is normal).

Make sure you can explain why all results in the table hold. For example, the last row of the table holds because the norm of an operator equals its largest singular value (by 7.85) and the singular values of a normal operator, assuming $\mathbf{F} = \mathbf{C}$, equal the absolute values of the eigenvalues (by Exercise 7 in Section 7E).

properties of a normal operator	**eigenvalues are contained in**
invertible	$\mathbf{C}\backslash\{0\}$
self-adjoint	$\mathbf{R}$
skew	$\{\lambda \in \mathbf{C} : \operatorname{Re}\lambda = 0\}$
orthogonal projection	$\{0, 1\}$
positive	$[0, \infty)$
unitary	$\{\lambda \in \mathbf{C} : \lvert\lambda\rvert = 1\}$
norm is less than 1	$\{\lambda \in \mathbf{C} : \lvert\lambda\rvert < 1\}$

7.111 פקטור שינוי הנפח הוא מכפלת הערכים הסינגולריים

נניח ש־ $T \in \mathcal{L}(V)$, $\mathbf{F} = \mathbf{R}$ הפיך, ו־ $\Omega \subseteq V$. אזי
$\text{volume } T(\Omega) = (\text{מכפלת הערכים הסינגולריים של } T)(\text{volume } \Omega)$.

הוכחה נניח שיש ל־ T פירוק לערכים סינגולריים

$$Tv = s_1\langle v, e_1\rangle f_1 + \cdots + s_n\langle v, e_n\rangle f_n$$

לכל $v \in V$, כאשר $e_1, \ldots, e_n$ ו־ $f_1, \ldots, f_n$ בסיסים אורתונורמליים של V.

נקרב את Ω באמצעות תיבות מהצורה $u + P(r_1 e_1, \ldots, r_n e_n)$, שהנפח שלהן הוא $r_1 \times \cdots \times r_n$. האופרטור T מעתיק כל תיבה $u + P(r_1 e_1, \ldots, r_n e_n)$ על התיבה $Tu + P(r_1 s_1 f_1, \ldots, r_n s_n f_n)$, שהנפח שלה הוא $(s_1 \times \cdots \times s_n)(r_1 \times \cdots \times r_n)$.

האופרטור T מעתיק אוסף תיבות קירוב ל־ Ω על אוסף תיבות קירוב ל־ $T(\Omega)$. מאחר ש־ T משנה את הנפח של כל תיבה באוסף המקרב את Ω בפקטור של $s_1 \times \cdots \times s_n$, ההעתקה הליניארית T משנה את הנפח של Ω באותו הפקטור. ■

נניח ש־ $T \in \mathcal{L}(V)$. כפי שנראה כאשר נגיע לנושא הדטרמיננטות, המכפלה של הערכים הסינגולריים של T שווה ל־ $|\det T|$; ראו 9.60 ו־ 9.61.

תכונות של אופרטור כפי שנקבעות על ידי הערכים העצמיים שלו

אנו חותמים את פרק זה בהצגת הטבלה להלן. ההקשר של טבלה זו הוא מרחב מכפלה פנימית מרוכב מממד סופי. העמודה הראשונה של הטבלה מציגה תכונה שייתכן שתהיה לאופרטור נורמלי מעל מרחב זה. העמודה השנייה של הטבלה מציגה תת־קבוצה של $\mathbf{C}$ כך שלאופרטור תהיה התכונה המתאימה אם־ורק־אם כל הערכים העצמיים של האופרטור נמצאים בתת־הקבוצה המצוינת. לדוגמה, השורה הראשונה בטבלה קובעת כי אופרטור נורמלי הוא הפיך אם־ורק־אם כל הערכים העצמיים שלו שונים מאפס (שורה ראשונה זו היא היחידה בטבלה שאינה זקוקה להנחה כי האופרטור נורמלי).

ודאו שביכולתכם להסביר מדוע כל התוצאות בטבלה מתקיימות. לדוגמה, השורה האחרונה של הטבלה מתקיימת מאחר שהנורמה של אופרטור שווה לערך הסינגולרי הגבוה ביותר (לפי 7.85) והערכים הסינגולריים של אופרטור נורמלי, בהנחה ש־ $\mathbf{F} = \mathbf{C}$, שווים לערכים המוחלטים של הערכים העצמיים (לפי תרגיל 7 ביחידה 5ז).

תכונות של אופרטור נורמלי	ערכים עצמיים מוכלים ב־
הפיך	$\mathbf{C} \setminus \{0\}$
צמוד לעצמו	$\mathbf{R}$
אנטי־צמוד	$\{\lambda \in \mathbf{C} : \operatorname{Re} \lambda = 0\}$
היטל אורתוגונלי	$\{0, 1\}$
חיובי	$[0, \infty)$
אוניטרי	$\{\lambda \in \mathbf{C} : \lvert\lambda\rvert = 1\}$
נורמה קטנה מ־ 1	$\{\lambda \in \mathbf{C} : \lvert\lambda\rvert < 1\}$

Exercises 7F

1 Prove that if $S, T \in \mathcal{L}(V, W)$, then $\big|\, \|S\| - \|T\| \,\big| \le \|S - T\|$.

The inequality above is called the ***reverse triangle inequality****.*

2 Suppose that $T \in \mathcal{L}(V)$ is self-adjoint or that $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$ is normal. Prove that

$$\|T\| = \max\{|\lambda| : \lambda \text{ is an eigenvalue of } T\}.$$

3 Suppose $T \in \mathcal{L}(V, W)$ and $v \in V$. Prove that

$$\|Tv\| = \|T\| \, \|v\| \iff T^*Tv = \|T\|^2 v.$$

4 Suppose $T \in \mathcal{L}(V, W)$, $v \in V$, and $\|Tv\| = \|T\| \, \|v\|$. Prove that if $u \in V$ and $\langle u, v\rangle = 0$, then $\langle Tu, Tv\rangle = 0$.

5 Suppose U is a finite-dimensional inner product space, $T \in \mathcal{L}(V, U)$, and $S \in \mathcal{L}(U, W)$. Prove that

$$\|ST\| \le \|S\| \, \|T\|.$$

6 Prove or give a counterexample: If $S, T \in \mathcal{L}(V)$, then $\|ST\| = \|TS\|$.

7 Show that defining $d(S, T) = \|S - T\|$ for $S, T \in \mathcal{L}(V, W)$ makes d a metric on $\mathcal{L}(V, W)$.

This exercise is intended for readers who are familiar with metric spaces.

8 (a) Prove that if $T \in \mathcal{L}(V)$ and $\|I - T\| < 1$, then T is invertible.

(b) Suppose that $S \in \mathcal{L}(V)$ is invertible. Prove that if $T \in \mathcal{L}(V)$ and $\|S - T\| < 1/\|S^{-1}\|$, then T is invertible.

This exercise shows that the set of invertible operators in $\mathcal{L}(V)$ is an open subset of $\mathcal{L}(V)$, using the metric defined in Exercise 7.

9 Suppose $T \in \mathcal{L}(V)$. Prove that for every $\epsilon > 0$, there exists an invertible operator $S \in \mathcal{L}(V)$ such that $0 < \|T - S\| < \epsilon$.

10 Suppose $\dim V > 1$ and $T \in \mathcal{L}(V)$ is not invertible. Prove that for every $\epsilon > 0$, there exists $S \in \mathcal{L}(V)$ such that $0 < \|T - S\| < \epsilon$ and S is not invertible.

11 Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$. Prove that for every $\epsilon > 0$ there exists a diagonalizable operator $S \in \mathcal{L}(V)$ such that $0 < \|T - S\| < \epsilon$.

12 Suppose $T \in \mathcal{L}(V)$ is a positive operator. Show that $\big\|\sqrt{T}\,\big\| = \sqrt{\|T\|}$.

13 Suppose $S, T \in \mathcal{L}(V)$ are positive operators. Show that

$$\|S - T\| \le \max\{\|S\|, \|T\|\} \le \|S + T\|.$$

14 Suppose U and W are subspaces of V such that $\|P_U - P_W\| < 1$. Prove that $\dim U = \dim W$.

תרגילים 7ט

1 הוכיחו כי אם $S, T \in \mathcal{L}(V, W)$, אזי $\big|\, \|S\| - \|T\| \,\big| \le \|S - T\|$.

האי־שוויון לעיל נקרא **אי־שוויון המשולש ההפוך**.

2 נניח ש־ $T \in \mathcal{L}(V)$ צמוד לעצמו או ש־ $\mathbf{F} = \mathbf{C}$ ו־ T נורמלי. הוכיחו כי הנורמה של T שווה לערך המוחלט הגדול ביותר מבין הערכים העצמיים של T.

3 נניח ש־ $T \in \mathcal{L}(V, W)$ ו־ $v \in V$. הוכיחו כי

$$\|Tv\| = \|T\| \, \|v\| \iff T^*Tv = \|T\|^2 v.$$

4 נניח ש־ $v \in V$, $T \in \mathcal{L}(V, W)$, ו־ $\|Tv\| = \|T\| \, \|v\|$. הוכיחו כי אם $u \in V$ ו־ $\langle u, v\rangle = 0$, אזי $\langle Tu, Tv\rangle = 0$.

5 נניח ש־ U מרחב מכפלה פנימית מממד סופי, $T \in \mathcal{L}(V, U)$, ו־ $S \in \mathcal{L}(U, W)$. הוכיחו כי

$$\|ST\| \le \|S\| \, \|T\|.$$

6 הוכיחו או הביאו דוגמה נגדית: אם $S, T \in \mathcal{L}(V)$, אזי $\|ST\| = \|TS\|$.

7 הראו כי הגדרת $d(S, T) = \|S - T\|$ עבור $S, T \in \mathcal{L}(V, W)$ הופכת את d למטריקה מעל $\mathcal{L}(V, W)$.

תרגיל זה מיועד לקוראים בעלי היכרות עם מרחבים מטריים.

8 א. הוכיחו כי אם $T \in \mathcal{L}(V)$ ו־ $\|I - T\| < 1$, אזי T הפיך.

ב. נניח ש־ $S \in \mathcal{L}(V)$ הפיך. הוכיחו כי אם $T \in \mathcal{L}(V)$ ו־ $\|S - T\| < 1/\|S^{-1}\|$, אזי T הפיך.

תרגיל זה מראה כי הקבוצה של אופרטורים הפיכים ב־ $\mathcal{L}(V)$ היא תת־קבוצה פתוחה של $\mathcal{L}(V)$, בשימוש במטריקה המוגדרת בתרגיל 7.

9 נניח ש־ $T \in \mathcal{L}(V)$. הוכיחו כי לכל $\epsilon > 0$, קיים אופרטור הפיך $S \in \mathcal{L}(V)$ כך ש־ $0 < \|T - S\| < \epsilon$.

10 נניח ש־ $\dim V > 1$ ו־ $T \in \mathcal{L}(V)$ אינו הפיך. הוכיחו כי לכל $\epsilon > 0$, קיים $S \in \mathcal{L}(V)$ כך ש־ $0 < \|T - S\| < \epsilon$ ו־ S אינו הפיך.

11 נניח ש־ $\mathbf{F} = \mathbf{C}$ ו־ $T \in \mathcal{L}(V)$. הוכיחו כי לכל $\epsilon > 0$ קיים אופרטור לכסין $S \in \mathcal{L}(V)$ כך ש־ $0 < \|T - S\| < \epsilon$.

12 נניח ש־ $T \in \mathcal{L}(V)$ אופרטור חיובי. הראו כי $\big\|\sqrt{T}\,\big\| = \sqrt{\|T\|}$.

13 נניח ש־ $S, T \in \mathcal{L}(V)$ אופרטורים חיוביים. הראו כי

$$\|S - T\| \le \max\{\|S\|, \|T\|\} \le \|S + T\|.$$

14 נניח ש־ U ו־ W תתי־מרחב של V כך ש־ $\|P_U - P_W\| < 1$. הוכיחו כי $\dim U = \dim W$.

15 Define $T \in \mathcal{L}(\mathbf{F}^3)$ by

$$T(z_1, z_2, z_3) = (z_3, 2z_1, 3z_2).$$

Find (explicitly) a unitary operator $S \in \mathcal{L}(\mathbf{F}^3)$ such that $T = S\sqrt{T^*T}$.

16 Suppose $S \in \mathcal{L}(V)$ is a positive invertible operator. Prove that there exists $\delta > 0$ such that T is a positive operator for every self-adjoint operator $T \in \mathcal{L}(V)$ with $\|S - T\| < \delta$.

17 Prove that if $u \in V$ and φ_u is the linear functional on V defined by the equation $\varphi_u(v) = \langle v, u\rangle$, then $\|\varphi_u\| = \|u\|$.

Here we are thinking of the scalar field $\mathbf{F}$ *as an inner product space with* $\langle \alpha, \beta\rangle = \alpha\overline{\beta}$ *for all* $\alpha, \beta \in \mathbf{F}$. *Thus* $\|\varphi_u\|$ *means the norm of* φ_u *as a linear map from* V *to* $\mathbf{F}$.

18 Suppose $e_1, \dots, e_n$ is an orthonormal basis of V and $T \in \mathcal{L}(V, W)$.

(a) Prove that $\max\{\|Te_1\|, \dots, \|Te_n\|\} \le \|T\| \le (\|Te_1\|^2 + \cdots + \|Te_n\|^2)^{1/2}$.

(b) Prove that $\|T\| = (\|Te_1\|^2 + \cdots + \|Te_n\|^2)^{1/2}$ if and only if $\dim \operatorname{range} T \le 1$.

Here $e_1, \dots, e_n$ *is an arbitrary orthonormal basis of* V, *not necessarily connected with a singular value decomposition of* T. *If* $s_1, \dots, s_n$ *is the list of singular values of* T, *then the right side of the inequality above equals* $(s_1^2 + \cdots + s_n^2)^{1/2}$, *as was shown in Exercise 11(a) in Section 7E.*

19 Prove that if $T \in \mathcal{L}(V, W)$, then $\|T^*T\| = \|T\|^2$.

This formula for $\|T^*T\|$ *leads to the important subject of* C^**-algebras.*

20 Suppose $T \in \mathcal{L}(V)$ is normal. Prove that $\|T^k\| = \|T\|^k$ for every positive integer k.

21 Suppose $\dim V > 1$ and $\dim W > 1$. Prove that the norm on $\mathcal{L}(V, W)$ does not come from an inner product. In other words, prove that there does not exist an inner product on $\mathcal{L}(V, W)$ such that

$$\max\{\|Tv\| : v \in V \text{ and } \|v\| \le 1\} = \sqrt{\langle T, T\rangle}$$

for all $T \in \mathcal{L}(V, W)$.

22 Suppose $T \in \mathcal{L}(V, W)$. Let $n = \dim V$ and let $s_1 \ge \cdots \ge s_n$ denote the singular values of T. Prove that if $1 \le k \le n$, then

$$\min\{\|T|_U\| : U \text{ is a subspace of } V \text{ with } \dim U = k\} = s_{n-k+1}.$$

23 Suppose $T \in \mathcal{L}(V, W)$. Show that T is uniformly continuous with respect to the metrics on V and W that arise from the norms on those spaces (see Exercise 23 in Section 6B).

24 Suppose $T \in \mathcal{L}(V)$ is invertible. Prove that

$$\|T^{-1}\| = \|T\|^{-1} \iff \frac{T}{\|T\|} \text{ is a unitary operator.}$$

15 נגדיר $T \in \mathcal{L}(\mathbf{F}^3)$ על ידי

$$T(z_1, z_2, z_3) = (z_3, 2z_1, 3z_2).$$

מצאו (במפורש) אופרטור אוניטרי $S \in \mathcal{L}(\mathbf{F}^3)$ כך ש־ $T = S\sqrt{T^*T}$.

16 נניח ש־ $S \in \mathcal{L}(V)$ אופרטור חיובי הפיך. הוכיחו כי קיים $\delta > 0$ כך ש־ T אופרטור חיובי לכל אופרטור צמוד לעצמו $T \in \mathcal{L}(V)$ עם $\|S - T\| < \delta$.

17 הוכיחו כי אם $u \in V$ ו־ φ_u הפונקציונל הליניארי מעל V המוגדר על ידי המשוואה $\varphi_u(v) = \langle v, u\rangle$, אזי $\|\varphi_u\| = \|u\|$.

כאן אנו חושבים על שדה הסקלרים $\mathbf{F}$ כמרחב מכפלה פנימית עם $\langle \alpha, \beta\rangle = \alpha\overline{\beta}$ לכל $\alpha, \beta \in \mathbf{F}$. לפיכך $\|\varphi_u\|$ משמעותו היא הנורמה של φ_u כהעתקה ליניארית מ־ V אל $\mathbf{F}$.

18 נניח ש־ $e_1, \dots, e_n$ בסיס אורתונורמלי של V ו־ $T \in \mathcal{L}(V, W)$.

א. הוכיחו כי $\max\{\|Te_1\|, \dots, \|Te_n\|\} \le \|T\| \le (\|Te_1\|^2 + \cdots + \|Te_n\|^2)^{1/2}$.

ב. הוכיחו כי $\|T\| = (\|Te_1\|^2 + \cdots + \|Te_n\|^2)^{1/2}$ אם־ורק־אם $\dim \operatorname{range} T \le 1$.

כאן $e_1, \dots, e_n$ בסיס אורתונורמלי שרירותי של V, שאינו בהכרח קשור לפירוק לערכים סינגולריים של T. אם $s_1, \dots, s_n$ היא הרשימה של ערכים סינגולריים של T, אזי האגף הימני של האי־שוויון לעיל שווה ל־ $(s_1^2 + \cdots + s_n^2)^{1/2}$, כפי שהוצג בתרגיל 11(א) ביחידה 7ז.

19 הוכיחו כי אם $T \in \mathcal{L}(V, W)$, אזי $\|T^*T\| = \|T\|^2$.

נוסחה זו עבור $\|T^*T\|$ מוכיחה לנושא החשוב של אלגברת סי כוכב.

20 נניח ש־ $T \in \mathcal{L}(V)$ נורמלי. הוכיחו כי $\|T^k\| = \|T\|^k$ לכל מספר שלם חיובי k.

21 נניח ש־ $\dim V > 1$ ו־ $\dim W > 1$. הוכיחו כי הנורמה מעל $\mathcal{L}(V, W)$ אינה נובעת ממכפלה פנימית. במילים אחרות, הוכיחו כי לא קיימת מכפלה פנימית מעל $\mathcal{L}(V, W)$ כך ש־

$$\max\{\|Tv\| : v \in V \text{ and } \|v\| \le 1\} = \sqrt{\langle T, T\rangle}$$

לכל $T \in \mathcal{L}(V, W)$.

22 נניח ש־ $T \in \mathcal{L}(V, W)$. יהי $n = \dim V$ ונסמן ב־ $s_1 \ge \cdots \ge s_n$ את הערכים הסינגולריים של T. הוכיחו כי אם $1 \le k \le n$, אזי

$$\min\{\|T|_U\| : U \text{ תת־מרחב של } V \text{ עם } \dim U = k\} = s_{n-k+1}.$$

23 נניח ש־ $T \in \mathcal{L}(V, W)$. הראו כי T רציפה במידה שווה ביחס אל המטריקות מעל V ו־ W הנובעות מהנורמות של אותם מרחבים (ראו תרגיל 23 ביחידה 21).

24 נניח ש־ $T \in \mathcal{L}(V)$ הפיך. הוכיחו כי

$$\|T^{-1}\| = \|T\|^{-1}$$

אם־ורק־אם $\frac{T}{\|T\|}$ אופרטור אוניטרי.

25 Fix $u, x \in V$ with $u \neq 0$. Define $T \in \mathcal{L}(V)$ by $Tv = \langle v, u\rangle x$ for every $v \in V$. Prove that

$$\sqrt{T^*T}v = \frac{\|x\|}{\|u\|}\langle v, u\rangle u$$

for every $v \in V$.

26 Suppose $T \in \mathcal{L}(V)$. Prove that T is invertible if and only if there exists a unique unitary operator $S \in \mathcal{L}(V)$ such that $T = S\sqrt{T^*T}$.

27 Suppose $T \in \mathcal{L}(V)$ and $s_1, \dots, s_n$ are the singular values of T. Let $e_1, \dots, e_n$ and $f_1, \dots, f_n$ be orthonormal bases of V such that

$$Tv = s_1\langle v, e_1\rangle f_1 + \cdots + s_n\langle v, e_n\rangle f_n$$

for all $v \in V$. Define $S \in \mathcal{L}(V)$ by

$$Sv = \langle v, e_1\rangle f_1 + \cdots + \langle v, e_n\rangle f_n.$$

(a) Show that S is unitary and $\|T - S\| = \max\{|s_1 - 1|, \dots, |s_n - 1|\}$.
(b) Show that if $E \in \mathcal{L}(V)$ is unitary, then $\|T - E\| \geq \|T - S\|$.

This exercise finds a unitary operator S that is as close as possible (among the unitary operators) to a given operator T.

28 Suppose $T \in \mathcal{L}(V)$. Prove that there exists a unitary operator $S \in \mathcal{L}(V)$ such that $T = \sqrt{TT^*}\, S$.

29 Suppose $T \in \mathcal{L}(V)$.

(a) Use the polar decomposition to show that there exists a unitary operator $S \in \mathcal{L}(V)$ such that $TT^* = ST^*TS^*$.
(b) Show how (a) implies that T and T^* have the same singular values.

30 Suppose $T \in \mathcal{L}(V)$, $S \in \mathcal{L}(V)$ is a unitary operator, and $R \in \mathcal{L}(V)$ is a positive operator such that $T = SR$. Prove that $R = \sqrt{T^*T}$.

This exercise shows that if we write T as the product of a unitary operator and a positive operator (as in the polar decomposition 7.93), then the positive operator equals $\sqrt{T^*T}$.

31 Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$ is normal. Prove that there exists a unitary operator $S \in \mathcal{L}(V)$ such that $T = S\sqrt{T^*T}$ and such that S and $\sqrt{T^*T}$ both have diagonal matrices with respect to the same orthonormal basis of V.

32 Suppose that $T \in \mathcal{L}(V, W)$ and $T \neq 0$. Let $s_1, \dots, s_m$ denote the positive singular values of T. Show that there exists an orthonormal basis $e_1, \dots, e_m$ of $(\operatorname{null} T)^\perp$ such that

$$T\Big(E\Big(\frac{e_1}{s_1}, \dots, \frac{e_m}{s_m}\Big)\Big)$$

equals the ball in $\operatorname{range} T$ of radius 1 centered at 0.

25 נְקַבֵּע $u, x \in V$ עם $u \neq 0$. נגדיר $T \in \mathcal{L}(V)$ על ידי $Tv = \langle v, u\rangle x$ לכל $v \in V$. הוכיחו כי

$$\sqrt{T^*T}v = \frac{\|x\|}{\|u\|}\langle v, u\rangle u$$

לכל $v \in V$.

26 נניח ש־ $T \in \mathcal{L}(V)$. הוכיחו כי T הפיך אם־ורק־אם קיים אופרטור אוניטרי יחיד $S \in \mathcal{L}(V)$ כך ש־ $T = S\sqrt{T^*T}$.

27 נניח ש־ $T \in \mathcal{L}(V)$ ו־ $s_1, \dots, s_n$ הם הערכים הסינגולריים של T. יהיו $e_1, \dots, e_n$ ו־ $f_1, \dots, f_n$ בסיסים אורתונורמליים של V כך ש־

$$Tv = s_1\langle v, e_1\rangle f_1 + \cdots + s_n\langle v, e_n\rangle f_n$$

לכל $v \in V$. נגדיר $S \in \mathcal{L}(V)$ על ידי

$$Sv = \langle v, e_1\rangle f_1 + \cdots + \langle v, e_n\rangle f_n.$$

א. הראו כי S אוניטרי וגם $\|T - S\| = \max\{|s_1 - 1|, \dots, |s_n - 1|\}$.

ב. הראו כי אם $E \in \mathcal{L}(V)$ אוניטרי, אזי $\|T - E\| \geq \|T - S\|$.

תרגיל זה מוצא אופרטור אוניטרי S הקרוב ככל הניתן (מבין האופרטורים האוניטריים) לאופרטור נתון T.

28 נניח ש־ $T \in \mathcal{L}(V)$. הוכיחו כי קיים אופרטור אוניטרי $S \in \mathcal{L}(V)$ כך ש־ $T = \sqrt{TT^*}\, S$.

29 נניח ש־ $T \in \mathcal{L}(V)$.

א. השתמשו בפירוק הפולרי כדי להראות כי קיים אופרטור אוניטרי $S \in \mathcal{L}(V)$ כך ש־ $TT^* = ST^*TS^*$.

ב. הראו כיצד (א) גורר שיש ל־ T ו־ T^* אותם ערכים סינגולריים.

30 נניח ש־ $T \in \mathcal{L}(V)$, $S \in \mathcal{L}(V)$ אופרטור אוניטרי, ו־ $R \in \mathcal{L}(V)$ אופרטור חיובי כך ש־ $T = SR$. הוכיחו כי $R = \sqrt{T^*T}$.

תרגיל זה מראה כי אם נכתוב את T כמכפלה של אופרטור אוניטרי ואופרטור חיובי (כמו בפירוק הפולרי 7.93), אזי האופרטור החיובי שווה ל־ $\sqrt{T^*T}$.

31 נניח ש־ $\mathbf{F} = \mathbf{C}$ ו־ $T \in \mathcal{L}(V)$ נורמלי. הוכיחו כי קיים אופרטור אוניטרי $S \in \mathcal{L}(V)$ כך ש־ $T = S\sqrt{T^*T}$ וכן לשני האופרטורים S ו־ $\sqrt{T^*T}$ יש מטריצות אלכסוניות ביחס אל אותו הבסיס האורתונורמלי של V.

32 נניח ש־ $T \in \mathcal{L}(V, W)$ ו־ $T \neq 0$. יהיו $s_1, \dots, s_m$ הערכים הסינגולריים החיוביים של T. הראו כי קיים בסיס אורתונורמלי $e_1, \dots, e_m$ של $(\operatorname{null} T)^\perp$ כך ש־

$$T\Big(E\big(\frac{e_1}{s_1}, \dots, \frac{e_m}{s_m}\big)\Big)$$

שווה לכדור ב־ $\operatorname{range} T$ שרדיוסו 1 ושמרכזו ב־ 0.

Chapter 8

Operators on Complex Vector Spaces

In this chapter we delve deeper into the structure of operators, with most of the attention on complex vector spaces. Some of the results in this chapter apply to both real and complex vector spaces; thus we do not make a standing assumption that $\mathbf{F} = \mathbf{C}$. Also, an inner product does not help with this material, so we return to the general setting of a finite-dimensional vector space.

Even on a finite-dimensional complex vector space, an operator may not have enough eigenvectors to form a basis of the vector space. Thus we will consider the closely related objects called generalized eigenvectors. We will see that for each operator on a finite-dimensional complex vector space, there is a basis of the vector space consisting of generalized eigenvectors of the operator. The generalized eigenspace decomposition then provides a good description of arbitrary operators on a finite-dimensional complex vector space.

Nilpotent operators, which are operators that when raised to some power equal 0, have an important role in these investigations. Nilpotent operators provide a key tool in our proof that every invertible operator on a finite-dimensional complex vector space has a square root and in our approach to Jordan form.

This chapter concludes by defining the trace and proving its key properties.

standing assumptions for this chapter

- $\mathbf{F}$ denotes $\mathbf{R}$ or $\mathbf{C}$.
- V denotes a finite-dimensional nonzero vector space over $\mathbf{F}$.

David Iliff CC BY-SA

The Long Room of the Old Library at the University of Dublin, where William Hamilton (1805–1865) was a student and then a faculty member. Hamilton proved a special case of what we now call the Cayley–Hamilton theorem in 1853.

פרק ח

אופרטורים מעל מרחבים וקטוריים מרוכבים

בפרק זה נעמיק חקר בנוגע למבנה של אופרטורים, כאשר מירב תשומת הלב תוקדש למרחבים וקטוריים מרוכבים. חלק מהמסקנות בפרק זה חלות על מרחבים וקטוריים ממשיים ומרוכבים כאחד; לפיכך לא נצא מנקודת ההנחה כי $\mathbf{F} = \mathbf{C}$. בנוסף, מכפלה פנימית אינה מסייעת בחומר זה, כך שאנו שבים אל המסגרת הכללית של מרחב וקטורי מממד סופי.

אפילו מעל מרחב וקטורי מרוכב מממד סופי, ייתכן שלאופרטור לא יהיו די וקטורים עצמיים להגדרת בסיס של המרחב הווקטורי. לפיכך נתבונן באובייקטים קשורים וקרובים הנקראים וקטורים עצמיים מוכללים. נראה כי לכל אופרטור מעל מרחב וקטורי מרוכב מממד סופי, יש בסיס של המרחב הווקטורי המורכב מווקטורים עצמיים מוכללים של האופרטור. הפירוק למרחבים עצמיים מוכללים יספק לנו תיאור טוב של אופרטורים שרירותיים מעל מרחב וקטורי מרוכב מממד סופי.

לאופרטורים נילפוטנטיים, שהם אופרטורים השווים 0 כאשר הם מֻעֲלִים בחזקה כלשהי, יש תפקיד חשוב בחקירות אלו. אופרטורים נילפוטנטיים מספקים כלי מרכזי בהוכחה כי לכל אופרטור הפיך מעל מרחב וקטורי מרוכב מממד סופי יש שורש ריבועי וכן בגישתנו לצורת ז'ורדן.

פרק זה נחתם בהגדרת העקבה ובהוכחת תכונותיה המרכזיות.

הנחות המוצא לפרק זה

- $\mathbf{F}$ מסמן $\mathbf{R}$ או $\mathbf{C}$.
- V מסמן מרחב וקטורי מממד סופי השונה מאפס מעל $\mathbf{F}$.

David Iliff CC BY-SA

החדר הארוך של הספרייה הישנה באוניברסיטת דבלין, שבה ויליאם המילטון (1805–1865) היה בתחילה סטודנט ולאחר מכן חבר סגל. המילטון הוכיח מקרה פרטי של מה שקרוי כעת משפט קיילי־המילטון בשנת 1853.

אלגברה ליניארית בדרך הישר, מהדורה רביעית, מחבר: שלדון אקסלר
Linear Algebra Done Right, fourth edition, by Sheldon Axler

8A Generalized Eigenvectors and Nilpotent Operators

Null Spaces of Powers of an Operator

We begin this chapter with a study of null spaces of powers of an operator.

8.1 *sequence of increasing null spaces*

Suppose $T \in \mathcal{L}(V)$. Then

$$\{0\} = \operatorname{null} T^0 \subseteq \operatorname{null} T^1 \subseteq \cdots \subseteq \operatorname{null} T^k \subseteq \operatorname{null} T^{k+1} \subseteq \cdots .$$

Proof Suppose k is a nonnegative integer and $v \in \operatorname{null} T^k$. Then $T^k v = 0$, which implies that $T^{k+1}v = T(T^k v) = T(0) = 0$. Thus $v \in \operatorname{null} T^{k+1}$. Hence $\operatorname{null} T^k \subseteq \operatorname{null} T^{k+1}$, as desired. ∎

The following result states that if two consecutive terms in the sequence of subspaces above are equal, then all later terms in the sequence are equal.

For similar results about decreasing sequences of ranges, see Exercises 6, 7, and 8.

8.2 *equality in the sequence of null spaces*

Suppose $T \in \mathcal{L}(V)$ and m is a nonnegative integer such that

$$\operatorname{null} T^m = \operatorname{null} T^{m+1}.$$

Then

$$\operatorname{null} T^m = \operatorname{null} T^{m+1} = \operatorname{null} T^{m+2} = \operatorname{null} T^{m+3} = \cdots .$$

Proof Let k be a positive integer. We want to prove that

$$\operatorname{null} T^{m+k} = \operatorname{null} T^{m+k+1}.$$

We already know from 8.1 that $\operatorname{null} T^{m+k} \subseteq \operatorname{null} T^{m+k+1}$.

To prove the inclusion in the other direction, suppose $v \in \operatorname{null} T^{m+k+1}$. Then

$$T^{m+1}(T^k v) = T^{m+k+1} v = 0.$$

Hence

$$T^k v \in \operatorname{null} T^{m+1} = \operatorname{null} T^m.$$

Thus $T^{m+k}v = T^m(T^k v) = 0$, which means that $v \in \operatorname{null} T^{m+k}$. This implies that $\operatorname{null} T^{m+k+1} \subseteq \operatorname{null} T^{m+k}$, completing the proof. ∎

The result above raises the question of whether there exists a nonnegative integer m such that $\operatorname{null} T^m = \operatorname{null} T^{m+1}$. The next result shows that this equality holds at least when m equals the dimension of the vector space on which T operates.

חו"ל וקטורים עצמיים מוכללים ואופרטורים נילפוטנטיים

מרחבי האפס של חזקות אופרטור

אנו פותחים פרק זה בחקר מרחבי האפס של חזקות אופרטור.

8.1 סדרת מרחבי אפס ההולכים וגדלים

נניח ש־ $T \in \mathcal{L}(V)$. אזי

$$\{0\} = \operatorname{null} T^0 \subseteq \operatorname{null} T^1 \subseteq \cdots \subseteq \operatorname{null} T^k \subseteq \operatorname{null} T^{k+1} \subseteq \cdots .$$

הוכחה נניח ש־ k מספר שלם אי־שלילי ו־ $v \in \operatorname{null} T^k$. אזי $T^k v = 0$, מה שגורר ש־ $T^{k+1} v = T(T^k v) = T(0) = 0$. לפיכך $v \in \operatorname{null} T^{k+1}$. לכן $\operatorname{null} T^k \subseteq \operatorname{null} T^{k+1}$, כפי שרצינו. ■

המסקנה הבאה קובעת כי אם שני איברים רצופים בסדרה של תתי־המרחב לעיל שווים, אזי כל האיברים המופיעים בסדרה לאחר מכן שווים.

לתוצאות דומות עבור סדרת טווחים ההולכים וקטנים, ראו תרגילים 6, 7 ו־8.

8.2 שוויון בסדרה של מרחבי אפס

נניח ש־ $T \in \mathcal{L}(V)$ ו־ m מספר שלם אי־שלילי כך ש־

$$\operatorname{null} T^m = \operatorname{null} T^{m+1}.$$

אזי

$$\operatorname{null} T^m = \operatorname{null} T^{m+1} = \operatorname{null} T^{m+2} = \operatorname{null} T^{m+3} = \cdots .$$

הוכחה יהי k מספר שלם חיובי. ברצוננו להוכיח כי

$$\operatorname{null} T^{m+k} = \operatorname{null} T^{m+k+1}.$$

אנו כבר יודעים מ־ 8.1 כי $\operatorname{null} T^{m+k} \subseteq \operatorname{null} T^{m+k+1}$.

להוכחת ההכלה בכיוון השני, נניח ש־ $v \in \operatorname{null} T^{m+k+1}$. אזי

$$T^{m+1}(T^k v) = T^{m+k+1} v = 0.$$

לכן

$$T^k v \in \operatorname{null} T^{m+1} = \operatorname{null} T^m.$$

לפיכך $T^{m+k} v = T^m(T^k v) = 0$, ומשמעות הדבר היא ש־ $v \in \operatorname{null} T^{m+k}$. מכאן נובע ש־ $\operatorname{null} T^{m+k+1} \subseteq \operatorname{null} T^{m+k}$, ובכך הושלמה ההוכחה. ■

המסקנה לעיל מעלה את השאלה האם קיים מספר שלם אי־שלילי m כך ש־ $\operatorname{null} T^m = \operatorname{null} T^{m+1}$. המסקנה הבאה מראה כי שוויון זה מתקיים לכל הפחות כאשר m שווה לממד המרחב הווקטורי שמעליו T פועל.

8.3 *null spaces stop growing*

Suppose $T \in \mathcal{L}(V)$. Then

$$\operatorname{null} T^{\dim V} = \operatorname{null} T^{\dim V+1} = \operatorname{null} T^{\dim V+2} = \cdots.$$

Proof We only need to prove that $\operatorname{null} T^{\dim V} = \operatorname{null} T^{\dim V+1}$ (by 8.2). Suppose this is not true. Then, by 8.1 and 8.2, we have

$$\{0\} = \operatorname{null} T^0 \subsetneq \operatorname{null} T^1 \subsetneq \cdots \subsetneq \operatorname{null} T^{\dim V} \subsetneq \operatorname{null} T^{\dim V+1},$$

where the symbol $\subsetneq$ means "contained in but not equal to". At each of the strict inclusions in the chain above, the dimension increases by at least 1. Thus $\dim \operatorname{null} T^{\dim V+1} \geq \dim V + 1$, a contradiction because a subspace of V cannot have a larger dimension than $\dim V$. ∎

It is not true that $V = \operatorname{null} T \oplus \operatorname{range} T$ for every $T \in \mathcal{L}(V)$. However, the next result can be a useful substitute.

8.4 *V is the direct sum of* $\operatorname{null} T^{\dim V}$ *and* $\operatorname{range} T^{\dim V}$

Suppose $T \in \mathcal{L}(V)$. Then

$$V = \operatorname{null} T^{\dim V} \oplus \operatorname{range} T^{\dim V}.$$

Proof Let $n = \dim V$. First we show that

8.5 $$(\operatorname{null} T^n) \cap (\operatorname{range} T^n) = \{0\}.$$

Suppose $v \in (\operatorname{null} T^n) \cap (\operatorname{range} T^n)$. Then $T^n v = 0$, and there exists $u \in V$ such that $v = T^n u$. Applying T^n to both sides of the last equation shows that $T^n v = T^{2n} u$. Hence $T^{2n} u = 0$, which implies that $T^n u = 0$ (by 8.3). Thus $v = T^n u = 0$, completing the proof of 8.5.

Now 8.5 implies that $\operatorname{null} T^n + \operatorname{range} T^n$ is a direct sum (by 1.46). Also,

$$\dim(\operatorname{null} T^n \oplus \operatorname{range} T^n) = \dim \operatorname{null} T^n + \dim \operatorname{range} T^n = \dim V,$$

where the first equality above comes from 3.94 and the second equality comes from the fundamental theorem of linear maps (3.21). The equation above implies that $\operatorname{null} T^n \oplus \operatorname{range} T^n = V$ (see 2.39), as desired. ∎

For an improvement of the result above, see Exercise 19.

8.6 example: $\mathbf{F}^3 = \operatorname{null} T^3 \oplus \operatorname{range} T^3$ *for* $T \in \mathcal{L}(\mathbf{F}^3)$

Suppose $T \in \mathcal{L}(\mathbf{F}^3)$ is defined by

$$T(z_1, z_2, z_3) = (4z_2, 0, 5z_3).$$

8.3 מרחבי אפס חזקים מלגדול

נניח ש־ $T \in \mathcal{L}(V)$. אזי

$$\operatorname{null} T^{\dim V} = \operatorname{null} T^{\dim V+1} = \operatorname{null} T^{\dim V+2} = \cdots.$$

הוכחה כל שעלינו להוכיח הוא ש־ $\operatorname{null} T^{\dim V} = \operatorname{null} T^{\dim V+1}$ (לפי 8.2). נניח בשלילה שהשוויון אינו מתקיים. אזי, לפי 8.1 ו־ 8.2, נקבל

$$\{0\} = \operatorname{null} T^0 \subsetneq \operatorname{null} T^1 \subsetneq \cdots \subsetneq \operatorname{null} T^{\dim V} \subsetneq \operatorname{null} T^{\dim V+1},$$

כאשר הסימן $\subsetneq$ משמעותו היא "מוכל אך אינו שווה ל...". בכל אחת מההכלות ממש בשרשרת לעיל, הממד גדל לכל הפחות ב־ 1. לפיכך $\dim \operatorname{null} T^{\dim V+1} \geq \dim V + 1$, תוצאה בסתירה לכך שממד תת־מרחב של V אינו יכול להיות גדול יותר מ־ $\dim V$. ∎

אין זה נכון לקבוע כי $V = \operatorname{null} T \oplus \operatorname{range} T$ לכל $T \in \mathcal{L}(V)$. בכל זאת, המסקנה הבאה יכולה להוות תחליף שימושי.

8.4 הסכום הישר של $\operatorname{null} T^{\dim V}$ ו־ $\operatorname{range} T^{\dim V}$

נניח ש־ $T \in \mathcal{L}(V)$. אזי

$$V = \operatorname{null} T^{\dim V} \oplus \operatorname{range} T^{\dim V}.$$

הוכחה יהי $n = \dim V$. ראשית נראה כי

$$(\operatorname{null} T^n) \cap (\operatorname{range} T^n) = \{0\}. \tag{8.5}$$

נניח ש־ $v \in (\operatorname{null} T^n) \cap (\operatorname{range} T^n)$. אזי $T^n v = 0$, ובנוסף, קיים $u \in V$ כך ש־ $v = T^n u$. הפעלת T^n על שני אגפי המשוואה האחרונה מראה ש־ $T^{2n} u = T^n v$. לכן $T^{2n} u = 0$, מה שגורר ש־ $T^n u = 0$ (לפי 8.3). לפיכך $v = T^n u = 0$, ובכך הושלמה הוכחת 8.5.

כעת, 8.5 גורר ש־ $\operatorname{null} T^n + \operatorname{range} T^n$ סכום ישר (לפי 1.46). בנוסף,

$$\dim(\operatorname{null} T^n \oplus \operatorname{range} T^n) = \dim \operatorname{null} T^n + \dim \operatorname{range} T^n = \dim V,$$

כאשר השוויון הראשון לעיל נובע מ־ 3.94 והשוויון השני נובע מהמשפט היסודי של ההעתקות הליניאריות (3.21). המשוואה לעיל גוררת ש־ $\operatorname{null} T^n \oplus \operatorname{range} T^n = V$ (ראו 2.39), כפי שרצינו. ∎

לשיפור המסקנה לעיל, ראו תרגיל 19.

8.6 דוגמה: $\mathbf{F}^3 = \operatorname{null} T^3 \oplus \operatorname{range} T^3$ עבור $T \in \mathcal{L}(\mathbf{F}^3)$

נניח ש־ $T \in \mathcal{L}(\mathbf{F}^3)$ מוגדר על ידי

$$T(z_1, z_2, z_3) = (4z_2, 0, 5z_3).$$

אזי $\operatorname{null} T = \{(z_1, 0, 0) : z_1 \in \mathbf{F}\}$ ו־ $\operatorname{range} T = \{(z_1, 0, z_3) : z_1, z_3 \in \mathbf{F}\}$. לפיכך

Then $\operatorname{null} T = \{(z_1, 0, 0) : z_1 \in \mathbf{F}\}$ and $\operatorname{range} T = \{(z_1, 0, z_3) : z_1, z_3 \in \mathbf{F}\}$. Thus $\operatorname{null} T \cap \operatorname{range} T \neq \{0\}$. Hence $\operatorname{null} T + \operatorname{range} T$ is not a direct sum. Also note that $\operatorname{null} T + \operatorname{range} T \neq \mathbf{F}^3$. However, we have $T^3(z_1, z_2, z_3) = (0, 0, 125z_3)$. Thus we see that

$$\operatorname{null} T^3 = \{(z_1, z_2, 0) : z_1, z_2 \in \mathbf{F}\} \quad \text{and} \quad \operatorname{range} T^3 = \{(0, 0, z_3) : z_3 \in \mathbf{F}\}.$$

Hence $\mathbf{F}^3 = \operatorname{null} T^3 \oplus \operatorname{range} T^3$, as expected by 8.4.

Generalized Eigenvectors

Some operators do not have enough eigenvectors to lead to good descriptions of their behavior. Thus in this subsection we introduce the concept of generalized eigenvectors, which will play a major role in our description of the structure of an operator.

To understand why we need more than eigenvectors, let's examine the question of describing an operator by decomposing its domain into invariant subspaces. Fix $T \in \mathcal{L}(V)$. We seek to describe T by finding a "nice" direct sum decomposition

$$V = V_1 \oplus \cdots \oplus V_n,$$

where each V_k is a subspace of V invariant under T. The simplest possible nonzero invariant subspaces are one-dimensional. A decomposition as above in which each V_k is a one-dimensional subspace of V invariant under T is possible if and only if V has a basis consisting of eigenvectors of T (see 5.55). This happens if and only if V has an eigenspace decomposition

8.7 $$V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T),$$

where $\lambda_1, \ldots, \lambda_m$ are the distinct eigenvalues of T (see 5.55).

The spectral theorem in the previous chapter shows that if V is an inner product space, then a decomposition of the form 8.7 holds for every self-adjoint operator if $\mathbf{F} = \mathbf{R}$ and for every normal operator if $\mathbf{F} = \mathbf{C}$ because operators of those types have enough eigenvectors to form a basis of V (see 7.29 and 7.31).

However, a decomposition of the form 8.7 may not hold for more general operators, even on a complex vector space. An example was given by the operator in 5.57, which does not have enough eigenvectors for 8.7 to hold. Generalized eigenvectors and generalized eigenspaces, which we now introduce, will remedy this situation.

8.8 definition: *generalized eigenvector*

Suppose $T \in \mathcal{L}(V)$ and λ is an eigenvalue of T. A vector $v \in V$ is called a *generalized eigenvector* of T corresponding to λ if $v \neq 0$ and

$$(T - \lambda I)^k v = 0$$

for some positive integer k.

$\operatorname{null} T \cap \operatorname{range} T \neq \{0\}$. לכן $\operatorname{null} T + \operatorname{range} T$ אינו סכום ישר. בנוסף, נשים לב כי $\operatorname{null} T + \operatorname{range} T \neq \mathbf{F}^3$. בכל זאת, מתקיים $T^3(z_1, z_2, z_3) = (0, 0, 125z_3)$. לפיכך אנו רואים ש־

$$\operatorname{null} T^3 = \{(z_1, z_2, 0) : z_1, z_2 \in \mathbf{F}\} \quad \text{וגם} \quad \operatorname{range} T^3 = \{(0, 0, z_3) : z_3 \in \mathbf{F}\}.$$

לכן $\mathbf{F}^3 = \operatorname{null} T^3 \oplus \operatorname{range} T^3$, כמצופה לפי 8.4.

וקטוריים עצמיים מוכללים

קיימים אופרטורים שמספר הווקטורים העצמיים שלהם אינו מספיק כדי שנוכל לתאר את פעולתם באופן מלא. לפיכך בתת־יחידה זו נציג את המושג וקטורים עצמיים מוכללים, אשר ימלא תפקיד ראשי בתיאור מבנה האופרטור.

כדי להבין מדוע אנו זקוקים ליותר מווקטורים עצמיים, הבה נבחן את שאלת תיאור האופרטור באמצעות פירוק התחום שלו לתתי־מרחב שמורים. נְקַבֵּעַ $T \in \mathcal{L}(V)$. אנו מבקשים לתאר את T על ידי מציאת פירוק "נחמד" באמצעות סכום ישר

$$V = V_1 \oplus \cdots \oplus V_n,$$

כאשר כל V_k הוא תת־מרחב שמור של V תחת T. תתי־המרחב השמורים השונים מאפס הפשוטים ביותר האפשריים הם החד־ממדיים. פירוק כמו זה המובא לעיל שבו כל V_k הוא תת־מרחב חד־ממדי של V השמור תחת T הוא אפשרי אם־ורק־אם יש ל־V בסיס המורכב מווקטורים עצמיים של T (ראו 5.55). הדבר קורה אם־ורק־אם ניתן לפרק את V למרחבים עצמיים

$$V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T), \tag{8.7}$$

כאשר $\lambda_1, \ldots, \lambda_m$ הם הערכים העצמיים השונים של T (ראו 5.55).

משפט הפירוק הספקטרלי בפרק הקודם הראה שאם V מרחב מכפלה פנימית, אזי פירוק בצורת 8.7 מתקיים לכל אופרטור הצמוד לעצמו אם $\mathbf{F} = \mathbf{R}$ ולכל אופרטור נורמלי אם $\mathbf{F} = \mathbf{C}$ מאחר שלאופרטורים מסוגים אלו יש מספיק וקטורים עצמיים ליצירת בסיס של V (ראו 7.29 ו־7.31).

בכל זאת, ייתכן שפירוק בצורת 8.7 לא יתקיים עבור אופרטורים כלליים יותר, אפילו מעל מרחב וקטורי מרוכב. דוגמה ניתנה על ידי האופרטור ב־5.57, שאין לו מספיק וקטורים עצמיים כך ש־8.7 יתקיים. וקטורים עצמיים מוכללים ומרחבים עצמיים מוכללים, שאותם נכיר כעת, יתקנו מצב זה.

8.8 הגדרה: וקטור עצמי מוכלל

נניח ש־$T \in \mathcal{L}(V)$ ו־λ הוא ערך עצמי של T. וקטור $v \in V$ נקרא וקטור עצמי מוכלל של T השייך ל־λ אם $v \neq 0$ וגם

$$(T - \lambda I)^k v = 0$$

עבור k מספר שלם חיובי כלשהו.

A nonzero vector $v \in V$ is a generalized eigenvector of T corresponding to λ if and only if

$$(T - \lambda I)^{\dim V} v = 0,$$

as follows from applying 8.1 and 8.3 to the operator $T - \lambda I$.

Generalized eigenvalues are not defined because doing so would not lead to anything new. Reason: if $(T - \lambda I)^k$ is not injective for some positive integer k, then $T - \lambda I$ is not injective, and hence λ is an eigenvalue of T.

As we know, an operator on a complex vector space may not have enough eigenvectors to form a basis of the domain. The next result shows that on a complex vector space there are enough generalized eigenvectors to do this.

8.9 *a basis of generalized eigenvectors*

Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$. Then there is a basis of V consisting of generalized eigenvectors of T.

Proof Let $n = \dim V$. We will use induction on n. To get started, note that the desired result holds if $n = 1$ because then every nonzero vector in V is an eigenvector of T.

Now suppose $n > 1$ and the desired result holds for all smaller values of $\dim V$. Let λ be an eigenvalue of T. Applying 8.4 to $T - \lambda I$ shows that

This step is where we use the hypothesis that $\mathbf{F} = \mathbf{C}$, because if $\mathbf{F} = \mathbf{R}$ then T may not have any eigenvalues.

$$V = \operatorname{null}(T - \lambda I)^n \oplus \operatorname{range}(T - \lambda I)^n.$$

If $\operatorname{null}(T - \lambda I)^n = V$, then every nonzero vector in V is a generalized eigenvector of T, and thus in this case there is a basis of V consisting of generalized eigenvectors of T. Hence we can assume that $\operatorname{null}(T - \lambda I)^n \neq V$, which implies that $\operatorname{range}(T - \lambda I)^n \neq \{0\}$.

Also, $\operatorname{null}(T - \lambda I)^n \neq \{0\}$, because λ is an eigenvalue of T. Thus we have

$$0 < \dim \operatorname{range}(T - \lambda I)^n < n.$$

Furthermore, $\operatorname{range}(T - \lambda I)^n$ is invariant under T [by 5.18 with $p(z) = (z - \lambda)^n$]. Let $S \in \mathcal{L}(\operatorname{range}(T - \lambda I)^n)$ equal T restricted to $\operatorname{range}(T - \lambda I)^n$. Our induction hypothesis applied to the operator S implies that there is a basis of $\operatorname{range}(T - \lambda I)^n$ consisting of generalized eigenvectors of S, which of course are generalized eigenvectors of T. Adjoining that basis of $\operatorname{range}(T - \lambda I)^n$ to a basis of $\operatorname{null}(T - \lambda I)^n$ gives a basis of V consisting of generalized eigenvectors of T. ∎

If $\mathbf{F} = \mathbf{R}$ and $\dim V > 1$, then some operators on V have the property that there exists a basis of V consisting of generalized eigenvectors of the operator, and (unlike what happens when $\mathbf{F} = \mathbf{C}$) other operators do not have this property. See Exercise 11 for a necessary and sufficient condition that determines whether an operator has this property.

וקטור השונה מאפס $v \in V$ הוא וקטור עצמי מוכלל של T השייך ל־ λ אם־ורק־אם

$$(T - \lambda I)^{\dim V} v = 0,$$

כפי שנובע מהפעלת 8.1 ו־ 8.3 על האופרטור $T - \lambda I$.

ערכים עצמיים מוכללים אינם מוגדרים מאחר שהגדרתם לא תוביל לתובנות חדשות. סיבה: אם $(T - \lambda I)^k$ אינו חד־חד־ערכי עבור k מספר שלם חיובי כלשהו, אזי $T - \lambda I$ אינו חד־חד־ערכי, ולכן λ ערך עצמי של T.

כידוע, ייתכן שלאופרטור מעל מרחב וקטורי מרוכב לא יהיו די וקטורים עצמיים ליצירת בסיס של התחום. המסקנה הבאה מראה שמעל מרחב וקטורי מרוכב יש מספיק וקטורים עצמיים מוכללים לצורך כך.

8.9 בסיס של וקטורים עצמיים מוכללים

נניח ש־ $\mathbf{F} = \mathbf{C}$ ו־ $T \in \mathcal{L}(V)$. אזי קיים בסיס של V המורכב מווקטורים עצמיים מוכללים של T.

הוכחה יהי $n = \dim V$. נשתמש באינדוקציה על n. תחילה, נשים לב שהתוצאה הרצויה מתקיימת אם $n = 1$ מאחר שאז כל וקטור השונה מאפס ב־ V הוא וקטור עצמי של T.

זה השלב שבו אנו משתמשים בהנחה ש־ $\mathbf{F} = \mathbf{C}$, מאחר שאם $\mathbf{F} = \mathbf{R}$ אזי ייתכן של־ T לא יהיו ערכים עצמיים כלל.

כעת נניח ש־ $n > 1$ והתוצאה הרצויה מתקיימת לכל מרחב וקטורי שממדו קטן מ־ n. יהי λ ערך עצמי של T. הפעלת 8.4 על $T - \lambda I$ מראה לנו ש־

$$V = \operatorname{null}(T - \lambda I)^n \oplus \operatorname{range}(T - \lambda I)^n.$$

אם $\operatorname{null}(T - \lambda I)^n = V$, אזי כל וקטור השונה מאפס ב־ V הוא וקטור עצמי מוכלל של T, ולפיכך במקרה זה קיים בסיס של V המורכב מווקטורים עצמיים מוכללים של T. לכן אנו יכולים להניח ש־ $\operatorname{null}(T - \lambda I)^n \neq V$, מה שגורר ש־ $\operatorname{range}(T - \lambda I)^n \neq \{0\}$.

בנוסף, $\operatorname{null}(T - \lambda I)^n \neq \{0\}$, מאחר ש־ λ הוא ערך עצמי של T. לפיכך מתקיים

$$0 < \dim \operatorname{range}(T - \lambda I)^n < n.$$

יתר על כן, $\operatorname{range}(T - \lambda I)^n$ שמור תחת T [לפי 5.18 כאשר $p(z) = (z - \lambda)^n$]. יהי $S \in \mathcal{L}(\operatorname{range}(T - \lambda I)^n)$ שווה ל־ T מצומצם ל־ $\operatorname{range}(T - \lambda I)^n$. הנחת האינדוקציה שלנו המוחלת על האופרטור S גוררת כי ישנו בסיס של $\operatorname{range}(T - \lambda I)^n$ המורכב מווקטורים עצמיים מוכללים של S, שהם בוודאי וקטורים עצמיים מוכללים של T. צירוף בסיס זה של $\operatorname{range}(T - \lambda I)^n$ לבסיס של $\operatorname{null}(T - \lambda I)^n$ נותן לנו בסיס של V המורכב מווקטורים עצמיים מוכללים של T. ■

אם $\mathbf{F} = \mathbf{R}$ ו־ $\dim V > 1$, אזי לאופרטורים מסוימים מעל V התכונה כי קיים בסיס של V המורכב מווקטורים עצמיים מוכללים של האופרטור, ואילו (בשונה ממה שקורה כאשר $\mathbf{F} = \mathbf{C}$) לאופרטורים אחרים אין תכונה זו. ראו תרגיל 11 עבור תנאי הכרחי ומספיק הקובע האם לאופרטור יש תכונה זו.

8.10 example: *generalized eigenvectors of an operator on* $\mathbf{C}^3$

Define $T \in \mathcal{L}(\mathbf{C}^3)$ by

$$T(z_1, z_2, z_3) = (4z_2, 0, 5z_3)$$

for each $(z_1, z_2, z_3) \in \mathbf{C}^3$. A routine use of the definition of eigenvalue shows that the eigenvalues of T are 0 and 5. Furthermore, the eigenvectors corresponding to the eigenvalue 0 are the nonzero vectors of the form $(z_1, 0, 0)$, and the eigenvectors corresponding to the eigenvalue 5 are the nonzero vectors of the form $(0, 0, z_3)$. Hence this operator does not have enough eigenvectors to span its domain $\mathbf{C}^3$.

We compute that $T^3(z_1, z_2, z_3) = (0, 0, 125z_3)$. Thus 8.1 and 8.3 imply that the generalized eigenvectors of T corresponding to the eigenvalue 0 are the nonzero vectors of the form $(z_1, z_2, 0)$.

We also have $(T - 5I)^3(z_1, z_2, z_3) = (-125z_1 + 300z_2, -125z_2, 0)$. Thus the generalized eigenvectors of T corresponding to the eigenvalue 5 are the nonzero vectors of the form $(0, 0, z_3)$.

The paragraphs above show that each of the standard basis vectors of $\mathbf{C}^3$ is a generalized eigenvector of T. Thus $\mathbf{C}^3$ indeed has a basis consisting of generalized eigenvectors of T, as promised by 8.9.

If v is an eigenvector of $T \in \mathcal{L}(V)$, then the corresponding eigenvalue λ is uniquely determined by the equation $Tv = \lambda v$, which can be satisfied by only one $\lambda \in \mathbf{F}$ (because $v \neq 0$). However, if v is a generalized eigenvector of T, then it is not obvious that the equation $(T - \lambda I)^{\dim V} v = 0$ can be satisfied by only one $\lambda \in \mathbf{F}$. Fortunately, the next result tells us that all is well on this issue.

8.11 *generalized eigenvector corresponds to a unique eigenvalue*

Suppose $T \in \mathcal{L}(V)$. Then each generalized eigenvector of T corresponds to only one eigenvalue of T.

Proof Suppose $v \in V$ is a generalized eigenvector of T corresponding to eigenvalues α and λ of T. Let m be the smallest positive integer such that $(T - \alpha I)^m v = 0$. Let $n = \dim V$. Then

$$\begin{aligned} 0 &= (T - \lambda I)^n v \\ &= ((T - \alpha I) + (\alpha - \lambda)I)^n v \\ &= \sum_{k=0}^{n} b_k (\alpha - \lambda)^{n-k} (T - \alpha I)^k v, \end{aligned}$$

where $b_0 = 1$ and the values of the other binomial coefficients b_k do not matter. Apply the operator $(T - \alpha I)^{m-1}$ to both sides of the equation above, getting

$$0 = (\alpha - \lambda)^n (T - \alpha I)^{m-1} v.$$

Because $(T - \alpha I)^{m-1} v \neq 0$, the equation above implies that $\alpha = \lambda$, as desired. ∎

8.10 דוגמה: וקטורים עצמיים מוכללים של אופרטור מעל $\mathbf{C}^3$

נגדיר $T \in \mathcal{L}(\mathbf{C}^3)$ על ידי

$$T(z_1, z_2, z_3) = (4z_2, 0, 5z_3)$$

לכל $(z_1, z_2, z_3) \in \mathbf{C}^3$. שימוש שגרתי בהגדרת ערך עצמי מראה לנו שהערכים העצמיים של T הם 0 ו־ 5. יתר על כן, הווקטורים העצמיים השייכים לערך העצמי 0 הם הווקטורים השונים מאפס מהצורה $(z_1, 0, 0)$, והווקטורים העצמיים השייכים לערך העצמי 5 הם הווקטורים השונים מאפס מהצורה $(0, 0, z_3)$. לכן לאופרטור זה אין מספיק וקטורים עצמיים לפריסת תחומו $\mathbf{C}^3$.

בפעולת חישוב נמצא כי $T^3(z_1, z_2, z_3) = (0, 0, 125z_3)$. לפיכך 8.1 ו־ 8.3 גוררים שהווקטורים העצמיים המוכללים של T השייכים לערך העצמי 0 הם הווקטורים השונים מאפס מהצורה $(z_1, z_2, 0)$.

בנוסף, נקבל ש־ $(T - 5I)^3(z_1, z_2, z_3) = (-125z_1 + 300z_2, -125z_2, 0)$. לפיכך הווקטורים העצמיים המוכללים של T השייכים לערך העצמי 5 הם הווקטורים השונים מאפס מהצורה $(0, 0, z_3)$.

מהפסקאות לעיל אנו רואים כי כל וקטור מווקטורי הבסיס הסטנדרטי של $\mathbf{C}^3$ הוא וקטור עצמי מוכלל של T. לפיכך בהחלט יש ל־ $\mathbf{C}^3$ בסיס המורכב מווקטורים עצמיים מוכללים של T, כפי שהובטח לנו על ידי 8.9.

אם v הוא וקטור עצמי של $T \in \mathcal{L}(V)$, אזי הערך העצמי השייך λ נקבע באופן יחיד על ידי המשוואה $Tv = \lambda v$, שאותה יכול לקיים רק $\lambda \in \mathbf{F}$ יחיד (שהרי $v \neq 0$). לעומת זאת, אם v וקטור עצמי מוכלל של T, אזי אין זה מובן מאליו שהמשוואה $(T - \lambda I)^{\dim V} v = 0$ יכולה להתקיים על ידי $\lambda \in \mathbf{F}$ יחיד. למזלנו הרב, המסקנה הבאה קובעת כי אין סיבה לדאגה בנוגע לשאלה זו.

8.11 וקטור עצמי מוכלל שייך לערך עצמי יחיד

נניח ש־ $T \in \mathcal{L}(V)$. אזי כל וקטור עצמי מוכלל של T שייך לערך עצמי אחד בלבד של T.

הוכחה נניח ש־ $v \in V$ וקטור עצמי מוכלל של T השייך לערכים העצמיים α ו־ λ של T. יהי m המספר השלם החיובי הקטן ביותר כך ש־ $(T - \alpha I)^m v = 0$. יהי $n = \dim V$. אזי

$$\begin{aligned} 0 &= (T - \lambda I)^n v \\ &= ((T - \alpha I) + (\alpha - \lambda)I)^n v \\ &= \sum_{k=0}^{n} b_k (\alpha - \lambda)^{n-k} (T - \alpha I)^k v, \end{aligned}$$

כאשר $b_0 = 1$ והערכים של מקדמי הבינום האחרים b_k אינם רלוונטיים. הפעלת האופרטור $(T - \alpha I)^{m-1}$ על שני אגפי המשוואה לעיל נותנת

$$0 = (\alpha - \lambda)^n (T - \alpha I)^{m-1} v.$$

מאחר ש־ $(T - \alpha I)^{m-1} v \neq 0$, המשוואה לעיל גוררת ש־ $\alpha = \lambda$, כפי שרצינו. ■

We saw earlier (5.11) that eigenvectors corresponding to distinct eigenvalues are linearly independent. Now we prove a similar result for generalized eigenvectors, with a proof that roughly follows the pattern of the proof of that earlier result.

8.12 *linearly independent generalized eigenvectors*

Suppose that $T \in \mathcal{L}(V)$. Then every list of generalized eigenvectors of T corresponding to distinct eigenvalues of T is linearly independent.

Proof Suppose the desired result is false. Then there exists a smallest positive integer m such that there exists a linearly dependent list $v_1, \dots, v_m$ of generalized eigenvectors of T corresponding to distinct eigenvalues $\lambda_1, \dots, \lambda_m$ of T (note that $m \geq 2$ because a generalized eigenvector is, by definition, nonzero). Thus there exist $a_1, \dots, a_m \in \mathbf{F}$, none of which are 0 (because of the minimality of m), such that

$$a_1 v_1 + \cdots + a_m v_m = 0.$$

Let $n = \dim V$. Apply $(T - \lambda_m I)^n$ to both sides of the equation above, getting

$$a_1 (T - \lambda_m I)^n v_1 + \cdots + a_{m-1} (T - \lambda_m I)^n v_{m-1} = 0. \tag{8.13}$$

Suppose $k \in \{1, \dots, m-1\}$. Then

$$(T - \lambda_m I)^n v_k \neq 0$$

because otherwise v_k would be a generalized eigenvector of T corresponding to the distinct eigenvalues λ_k and λ_m, which would contradict 8.11. However,

$$(T - \lambda_k I)^n ((T - \lambda_m I)^n v_k) = (T - \lambda_m I)^n ((T - \lambda_k I)^n v_k) = 0.$$

Thus the last two displayed equations show that $(T - \lambda_m I)^n v_k$ is a generalized eigenvector of T corresponding to the eigenvalue λ_k. Hence

$$(T - \lambda_m I)^n v_1, \dots, (T - \lambda_m I)^n v_{m-1}$$

is a linearly dependent list (by 8.13) of $m-1$ generalized eigenvectors corresponding to distinct eigenvalues, contradicting the minimality of m. This contradiction completes the proof. ∎

Nilpotent Operators

8.14 definition: *nilpotent*

An operator is called *nilpotent* if some power of it equals 0.

Thus an operator $T \in \mathcal{L}(V)$ is nilpotent if and only if every nonzero vector in V is a generalized eigenvector of T corresponding to the eigenvalue 0.

ראינו לעיל (5.11) כי וקטורים עצמיים השייכים לערכים עצמיים שונים הם בלתי תלויים ליניארית. כעת נוכיח מסקנה דומה בנוגע לווקטורים עצמיים מוכללים, עם הוכחה הדומה באופן מסוים לתבנית ההוכחה של המסקנה ההיא.

8.12 וקטורים עצמיים מוכללים בלתי תלויים ליניארית

נניח ש־ $T \in \mathcal{L}(V)$. אזי כל רשימה של וקטורים עצמיים מוכללים של T השייכים לערכים עצמיים שונים של T היא בלתי תלויה ליניארית.

הוכחה נניח שהתוצאה הרצויה אינה מתקיימת. אזי קיים מספר שלם חיובי קטן ביותר m כך שקיימת רשימה תלויה ליניארית $v_1, \dots, v_m$ של וקטורים עצמיים מוכללים של T השייכים לערכים עצמיים שונים $\lambda_1, \dots, \lambda_m$ של T (שימו לב ש־ $m \geq 2$ מאחר שווקטור עצמי מוכלל הוא, לפי הגדרה, שונה מאפס). לפיכך קיימים $a_1, \dots, a_m \in \mathbf{F}$, שאף אחד מהם אינו 0 (לאור המינימליות של m), כך ש־

$$a_1 v_1 + \cdots + a_m v_m = 0.$$

יהי $n = \dim V$. נפעיל $(T - \lambda_m I)^n$ על שני אגפי המשוואה לעיל ונקבל

$$a_1 (T - \lambda_m I)^n v_1 + \cdots + a_{m-1} (T - \lambda_m I)^n v_{m-1} = 0. \tag{8.13}$$

נניח ש־ $k \in \{1, \dots, m-1\}$. אזי

$$(T - \lambda_m I)^n v_k \neq 0$$

אחרת v_k יהיה וקטור עצמי מוכלל של T השייך לערכים העצמיים השונים λ_k ו־ λ_m, מה שיעמוד בסתירה ל־ 8.11. עם זאת,

$$(T - \lambda_k I)^n ((T - \lambda_m I)^n v_k) = (T - \lambda_m I)^n ((T - \lambda_k I)^n v_k) = 0.$$

לפיכך שתי המשוואות האחרונות שהצגנו מראות כי $(T - \lambda_m I)^n v_k$ וקטור עצמי מוכלל של T השייך לערך עצמי λ_k. לכן

$$(T - \lambda_m I)^n v_1, \dots, (T - \lambda_m I)^n v_{m-1}$$

רשימה תלויה ליניארית (לפי 8.13) של $m-1$ וקטורים עצמיים מוכללים השייכים לערכים עצמיים שונים, בסתירה למינימליות של m. ועם סתירה זו השלמנו את ההוכחה. ■

אופרטורים נילפוטנטיים

8.14 הגדרה: נילפוטנטי

אופרטור נקרא נילפוטנטי אם חזקה כלשהי שלו שווה 0.

לפיכך אופרטור $T \in \mathcal{L}(V)$ הוא נילפוטנטי אם־ורק־אם כל וקטור השונה מאפס ב־ V הוא וקטור עצמי מוכלל של T השייך לערך העצמי 0.

8.15 example: *nilpotent operators*

(a) The operator $T \in \mathcal{L}(\mathbf{F}^4)$ defined by

$$T(z_1, z_2, z_3, z_4) = (0, 0, z_1, z_2)$$

is nilpotent because $T^2 = 0$.

(b) The operator on $\mathbf{F}^3$ whose matrix (with respect to the standard basis) is

$$\begin{pmatrix} -3 & 9 & 0 \\ -7 & 9 & 6 \\ 4 & 0 & -6 \end{pmatrix}$$

is nilpotent, as can be shown by cubing the matrix above to get the zero matrix.

(c) The operator of differentiation on $\mathcal{P}_m(\mathbf{R})$ is nilpotent because the $(m+1)^{\text{th}}$ derivative of every polynomial of degree at most m equals 0. Note that on this space of dimension $m+1$, we need to raise the nilpotent operator to the power $m+1$ to get the 0 operator.

The next result shows that when raising a nilpotent operator to a power, we never need to use a power higher than the dimension of the space. For a slightly stronger result, see Exercise 18.

*The Latin word **nil** means nothing or zero; the Latin word **potens** means having power. Thus **nilpotent** literally means having a power that is zero.*

8.16 *nilpotent operator raised to dimension of domain is* 0

Suppose $T \in \mathcal{L}(V)$ is nilpotent. Then $T^{\dim V} = 0$.

Proof Because T is nilpotent, there exists a positive integer k such that $T^k = 0$. Thus $\text{null } T^k = V$. Now 8.1 and 8.3 imply that $\text{null } T^{\dim V} = V$. Thus $T^{\dim V} = 0$. ∎

8.17 *eigenvalues of nilpotent operator*

Suppose $T \in \mathcal{L}(V)$.

(a) If T is nilpotent, then 0 is an eigenvalue of T and T has no other eigenvalues.

(b) If $\mathbf{F} = \mathbf{C}$ and 0 is the only eigenvalue of T, then T is nilpotent.

Proof

(a) To prove (a), suppose T is nilpotent. Hence there is a positive integer m such that $T^m = 0$. This implies that T is not injective. Thus 0 is an eigenvalue of T.

8.15 דוגמה: אופרטורים נילפוטנטיים

א. האופרטור $T \in \mathcal{L}(\mathbf{F}^4)$ המוגדר על ידי
$$T(z_1, z_2, z_3, z_4) = (0, 0, z_1, z_2)$$
הוא נילפוטנטי מאחר ש־ $T^2 = 0$.

ב. האופרטור מעל $\mathbf{F}^3$ שהמטריצה שלו (ביחס לבסיס הסטנדרטי) היא
$$\begin{pmatrix} -3 & 9 & 0 \\ -7 & 9 & 6 \\ 4 & 0 & -6 \end{pmatrix}$$
הוא נילפוטנטי, כפי שניתן לראות כאשר נעלה את המטריצה לעיל בחזקת שלוש ונקבל את מטריצת האפס.

ג. אופרטור הגזירה מעל $\mathcal{P}_m(\mathbf{R})$ הוא נילפוטנטי מאחר שהנגזרת ה־ $(m+1)$ של כל פולינום ממעלה m לכל היותר שווה 0. שימו לב שמעל מרחב זה מממד $m+1$, עלינו להעלות את האופרטור הנילפוטנטי בחזקת $m+1$ כדי לקבל את אופרטור האפס.

המסקנה הבאה מראה שכאשר מעלים אופרטור נילפוטנטי בחזקה, לעולם לא נזדקק להעלות בחזקה גבוהה יותר מממד המרחב. למסקנה מעט חזקה יותר, ראו תרגיל 18.

ניל בלטינית משמעותה ריק או אפס; **פוטנס** בלטינית משמעותה עוצמה (חֻזְקָה). לפיכך המשמעות המילולית של **נילפוטנטי** היא קיומה של חזקה שהיא אפס.

8.16 אופרטור נילפוטנטי המועלה בחזקת ממד התחום הוא 0

נניח ש־ $T \in \mathcal{L}(V)$ נילפוטנטי. אזי $T^{\dim V} = 0$.

הוכחה מאחר ש־ T נילפוטנטי, קיים מספר שלם חיובי k כך ש־ $T^k = 0$. לפיכך $\operatorname{null} T^k = V$. כעת, מ־ 8.1 ו־ 8.3 נובע ש־ $\operatorname{null} T^{\dim V} = V$. לפיכך $T^{\dim V} = 0$. ■

8.17 ערכים עצמיים של אופרטור נילפוטנטי

נניח ש־ $T \in \mathcal{L}(V)$.

א. אם T נילפוטנטי, אזי 0 הוא ערך עצמי של T ואין ל־ T ערכים עצמיים נוספים.

ב. אם $\mathbf{F} = \mathbf{C}$ ו־ 0 הוא הערך העצמי היחיד של T, אזי T נילפוטנטי.

הוכחה

א. להוכחת (א), נניח ש־ T נילפוטנטי. לכן קיים m מספר שלם חיובי כך ש־ $T^m = 0$. מכאן נובע ש־ T אינו חד־חד־ערכי. לפיכך 0 הוא ערך עצמי של T.

To show that T has no other eigenvalues, suppose λ is an eigenvalue of T. Then there exists a nonzero vector $v \in V$ such that

$$\lambda v = Tv.$$

Repeatedly applying T to both sides of this equation shows that

$$\lambda^m v = T^m v = 0.$$

Thus $\lambda = 0$, as desired.

(b) Suppose $\mathbf{F} = \mathbf{C}$ and 0 is the only eigenvalue of T. By 5.27(b), the minimal polynomial of T equals z^m for some positive integer m. Thus $T^m = 0$. Hence T is nilpotent. ∎

Exercise 23 shows that the hypothesis that $\mathbf{F} = \mathbf{C}$ cannot be deleted in (b) of the result above.

Given an operator on V, we want to find a basis of V such that the matrix of the operator with respect to this basis is as simple as possible, meaning that the matrix contains many 0's. The next result shows that if T is nilpotent, then we can choose a basis of V such that the matrix of T with respect to this basis has more than half of its entries equal to 0. Later in this chapter we will do even better.

8.18 *minimal polynomial and upper-triangular matrix of nilpotent operator*

Suppose $T \in \mathcal{L}(V)$. Then the following are equivalent.

(a) T is nilpotent.

(b) The minimal polynomial of T is z^m for some positive integer m.

(c) There is a basis of V with respect to which the matrix of T has the form

$$\begin{pmatrix} 0 & & * \\ & \ddots & \\ 0 & & 0 \end{pmatrix},$$

where all entries on and below the diagonal equal 0.

Proof Suppose (a) holds, so T is nilpotent. Thus there exists a positive integer n such that $T^n = 0$. Now 5.29 implies that z^n is a polynomial multiple of the minimal polynomial of T. Thus the minimal polynomial of T is z^m for some positive integer m, proving that (a) implies (b).

Now suppose (b) holds, so the minimal polynomial of T is z^m for some positive integer m. This implies, by 5.27(a), that 0 (which is the only zero of z^m) is the only eigenvalue of T. This further implies, by 5.44, that there is a basis of V with respect to which the matrix of T is upper triangular. This also implies, by 5.41, that all entries on the diagonal of this matrix are 0, proving that (b) implies (c).

Now suppose (c) holds. Then 5.40 implies that $T^{\dim V} = 0$. Thus T is nilpotent, proving that (c) implies (a). ∎

כדי להראות כי אין ל־ T ערכים עצמיים נוספים, נניח ש־ λ ערך עצמי של T. אזי קיים וקטור $v \in V$ השונה מאפס כך ש־

$$\lambda v = Tv.$$

הפעלה חוזרת ונשנית של T על שני אגפי משוואה זו מראה לנו כי

$$\lambda^m v = T^m v = 0.$$

לפיכך $\lambda = 0$, כפי שרצינו.

ב. נניח ש־ $\mathbf{F} = \mathbf{C}$ ו־ 0 הערך העצמי היחיד של T. לפי 5.27(ב), הפולינום המינימלי של T שווה ל־ z^m עבור m מספר שלם חיובי כלשהו. לפיכך $T^m = 0$. לכן T נילפוטנטי. ■

תרגיל 23 מראה כי בסעיף (ב) של המסקנה לעיל, לא ניתן לוותר על ההנחה ש־ $\mathbf{F} = \mathbf{C}$.

בהינתן אופרטור מעל V, אנו מעוניינים למצוא בסיס של V כך שהמטריצה של האופרטור ביחס לבסיס זה תהיה פשוטה ככל האפשר, כלומר שהמטריצה תכיל אפסים רבים. המסקנה הבאה מראה שאם T נילפוטנטי, אזי נוכל לבחור בסיס של V כך שביותר ממחצית איברי המטריצה של T ביחס לבסיס זה יהיו אפסים. בהמשך פרק זה נציג תוצאות טובות יותר מזו.

8.18 פולינום מינימלי ומטריצה משולשית עליונה של אופרטור נילפוטנטי

נניח ש־ $T \in \mathcal{L}(V)$. אזי התנאים הבאים שקולים.

א. T נילפוטנטי.

ב. הפולינום המינימלי של T הוא z^m עבור m מספר שלם חיובי כלשהו.

ג. קיים בסיס של V אשר ביחס אליו המטריצה של T היא מהצורה

$$\begin{pmatrix} 0 & & * \\ & \ddots & \\ 0 & & 0 \end{pmatrix},$$

כאשר כל האיברים על האלכסון ומתחתיו שווים 0.

הוכחה נניח כי (א) מתקיים, ולכן T נילפוטנטי. לפיכך קיים n מספר שלם חיובי כך ש־ $T^n = 0$. כעת 5.29 גוררת ש־ z^n כפולת פולינום של הפולינום המינימלי של T. לפיכך הפולינום המינימלי של T הוא z^m עבור m מספר שלם חיובי כלשהו, ובכך הוכחנו כי (א) גורר את (ב).

כעת נניח כי (ב) מתקיים, ולכן הפולינום המינימלי של T הוא z^m עבור m מספר שלם חיובי כלשהו. מכאן נובע, לפי 5.27(א), כי 0 (שהוא השורש היחיד של z^m) הוא הערך העצמי היחיד של T. יתרה מזו, מכאן נובע, לפי 5.44, שישנו בסיס של V אשר ביחס אליו המטריצה של T משולשית עליונה. כמו כן, מכאן נובע, לפי 5.41, שכל האיברים על האלכסון של מטריצה זו הם 0, ובכך הוכחנו כי (ב) גורר את (ג).

כעת נניח כי (ג) מתקיים. אזי 5.40 גורר ש־ $T^{\dim V} = 0$. לפיכך T נילפוטנטי ובכך הוכחנו כי (ג) גורר את (א). ■

Exercises 8A

1 Suppose $T \in \mathcal{L}(V)$. Prove that if $\dim \operatorname{null} T^4 = 8$ and $\dim \operatorname{null} T^6 = 9$, then $\dim \operatorname{null} T^m = 9$ for all integers $m \geq 5$.

2 Suppose $T \in \mathcal{L}(V)$, m is a positive integer, $v \in V$, and $T^{m-1}v \neq 0$ but $T^m v = 0$. Prove that $v, Tv, T^2v, \dots, T^{m-1}v$ is linearly independent.

The result in this exercise is used in the proof of 8.45.

3 Suppose $T \in \mathcal{L}(V)$. Prove that

$$V = \operatorname{null} T \oplus \operatorname{range} T \iff \operatorname{null} T^2 = \operatorname{null} T.$$

4 Suppose $T \in \mathcal{L}(V)$, $\lambda \in \mathbf{F}$, and m is a positive integer such that the minimal polynomial of T is a polynomial multiple of $(z - \lambda)^m$. Prove that

$$\dim \operatorname{null}(T - \lambda I)^m \geq m.$$

5 Suppose $T \in \mathcal{L}(V)$ and m is a positive integer. Prove that

$$\dim \operatorname{null} T^m \leq m \dim \operatorname{null} T.$$

Hint: Exercise 21 in Section 3B may be useful.

6 Suppose $T \in \mathcal{L}(V)$. Show that

$$V = \operatorname{range} T^0 \supseteq \operatorname{range} T^1 \supseteq \cdots \supseteq \operatorname{range} T^k \supseteq \operatorname{range} T^{k+1} \supseteq \cdots .$$

7 Suppose $T \in \mathcal{L}(V)$ and m is a nonnegative integer such that

$$\operatorname{range} T^m = \operatorname{range} T^{m+1}.$$

Prove that $\operatorname{range} T^k = \operatorname{range} T^m$ for all $k > m$.

8 Suppose $T \in \mathcal{L}(V)$. Prove that

$$\operatorname{range} T^{\dim V} = \operatorname{range} T^{\dim V + 1} = \operatorname{range} T^{\dim V + 2} = \cdots .$$

9 Suppose $T \in \mathcal{L}(V)$ and m is a nonnegative integer. Prove that

$$\operatorname{null} T^m = \operatorname{null} T^{m+1} \iff \operatorname{range} T^m = \operatorname{range} T^{m+1}.$$

10 Define $T \in \mathcal{L}(\mathbf{C}^2)$ by $T(w, z) = (z, 0)$. Find all generalized eigenvectors of T.

11 Suppose that $T \in \mathcal{L}(V)$. Prove that there is a basis of V consisting of generalized eigenvectors of T if and only if the minimal polynomial of T equals $(z - \lambda_1) \cdots (z - \lambda_m)$ for some $\lambda_1, \dots, \lambda_m \in \mathbf{F}$.

Assume $\mathbf{F} = \mathbf{R}$ *because the case* $\mathbf{F} = \mathbf{C}$ *follows from 5.27(b) and 8.9.*

This exercise states that the condition for there to be a basis of V consisting of generalized eigenvectors of T is the same as the condition for there to be a basis with respect to which T has an upper-triangular matrix (see 5.44).

Caution: *If T has an upper-triangular matrix with respect to a basis $v_1, \dots, v_n$ of V, then v_1 is an eigenvector of T but it is not necessarily true that $v_2, \dots, v_n$ are generalized eigenvectors of T.*

תרגילים חו1

1 נניח ש־ $T \in \mathcal{L}(V)$. הוכיחו כי אם $\dim \operatorname{null} T^4 = 8$ ו־ $\dim \operatorname{null} T^6 = 9$, אזי $\dim \operatorname{null} T^m = 9$ לכל המספרים השלמים $m \geq 5$.

2 נניח ש־ $T \in \mathcal{L}(V)$, m מספר שלם חיובי, $v \in V$, ו־ $T^{m-1}v \neq 0$ אך $T^m v = 0$. הוכיחו כי $v, Tv, T^2v, \dots, T^{m-1}v$ בלתי תלויה ליניארית.

הוכחת 8.45 עושה שימוש בתוצאה בתרגיל זה.

3 נניח ש־ $T \in \mathcal{L}(V)$. הוכיחו כי

$$V = \operatorname{null} T \oplus \operatorname{range} T \iff \operatorname{null} T^2 = \operatorname{null} T.$$

4 נניח ש־ $T \in \mathcal{L}(V)$, $\lambda \in \mathbf{F}$, ו־ m מספר שלם חיובי כך שהפולינום המינימלי של T הוא כפולת פולינום של $(z - \lambda)^m$. הוכיחו כי

$$\dim \operatorname{null}(T - \lambda I)^m \geq m.$$

5 נניח ש־ $T \in \mathcal{L}(V)$ ו־ m מספר שלם חיובי. הוכיחו כי

$$\dim \operatorname{null} T^m \leq m \dim \operatorname{null} T.$$

רמז: תרגיל 21 ביחידה 2ג יוכל לסייע בפתרון.

6 נניח ש־ $T \in \mathcal{L}(V)$. הראו כי

$$V = \operatorname{range} T^0 \supseteq \operatorname{range} T^1 \supseteq \cdots \supseteq \operatorname{range} T^k \supseteq \operatorname{range} T^{k+1} \supseteq \cdots.$$

7 נניח ש־ $T \in \mathcal{L}(V)$ ו־ m מספר שלם אי־שלילי כך ש־

$$\operatorname{range} T^m = \operatorname{range} T^{m+1}.$$

הוכיחו כי $\operatorname{range} T^k = \operatorname{range} T^m$ לכל $k > m$.

8 נניח ש־ $T \in \mathcal{L}(V)$. הוכיחו כי

$$\operatorname{range} T^{\dim V} = \operatorname{range} T^{\dim V + 1} = \operatorname{range} T^{\dim V + 2} = \cdots.$$

9 נניח ש־ $T \in \mathcal{L}(V)$ ו־ m מספר שלם אי־שלילי. הוכיחו כי

$$\operatorname{null} T^m = \operatorname{null} T^{m+1} \iff \operatorname{range} T^m = \operatorname{range} T^{m+1}.$$

10 נגדיר $T \in \mathcal{L}(\mathbf{C}^2)$ על ידי $T(w, z) = (z, 0)$. מצאו את כל הווקטורים העצמיים המוכללים של T.

11 נניח ש־ $T \in \mathcal{L}(V)$. הוכיחו כי קיים בסיס של V המורכב מווקטורים עצמיים מוכללים של T אם־ורק־אם הפולינום המינימלי של T שווה ל־ $(z-\lambda_1) \cdots (z-\lambda_m)$ עבור $\lambda_1, \dots, \lambda_m \in \mathbf{F}$ כלשהם.

נניח ש־ $\mathbf{F} = \mathbf{R}$ מאחר שהמקרה $\mathbf{F} = \mathbf{C}$ נובע מ־ 5.27(ב) ו־ 8.9.

תרגיל זה קובע כי התנאי לקיומו של בסיס של V המורכב מווקטורים עצמיים מוכללים של T הוא זהה לתנאי לקיומו של בסיס אשר ביחס אליו יש ל־ T מטריצה משולשית עליונה (ראו 5.44).

אזהרה: אם יש ל־ T מטריצה משולשית עליונה ביחס לבסיס $v_1, \dots, v_n$ של V, אזי v_1 הוא וקטור עצמי של T אך לא נוכל לקבוע בהכרח כי $v_2, \dots, v_n$ גם הם וקטורים עצמיים מוכללים של T.

12 Suppose $T \in \mathcal{L}(V)$ is such that every nonzero vector in V is a generalized eigenvector of T. Prove that there exists $\lambda \in \mathbf{F}$ such that $T - \lambda I$ is nilpotent.

13 Suppose $S, T \in \mathcal{L}(V)$ and ST is nilpotent. Prove that TS is nilpotent.

14 Suppose $T \in \mathcal{L}(V)$ is nilpotent and $T \neq 0$. Prove T is not diagonalizable.

15 Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$. Prove that T is diagonalizable if and only if every generalized eigenvector of T is an eigenvector of T.

For $\mathbf{F} = \mathbf{C}$, this exercise adds another equivalence to the list of conditions for diagonalizability in 5.55.

16 (a) Give an example of nilpotent operators S, T on the same vector space such that neither $S + T$ nor ST is nilpotent.
(b) Suppose $S, T \in \mathcal{L}(V)$ are nilpotent and $ST = TS$. Prove that $S + T$ and ST are nilpotent.

17 Suppose $T \in \mathcal{L}(V)$ is nilpotent and m is a positive integer such that $T^m = 0$.
(a) Prove that $I - T$ is invertible and that $(I - T)^{-1} = I + T + \cdots + T^{m-1}$.
(b) Explain how you would guess the formula above.

18 Suppose $T \in \mathcal{L}(V)$ is nilpotent. Prove that $T^{1+\dim \operatorname{range} T} = 0$.

If $\dim \operatorname{range} T < \dim V - 1$, then this exercise improves 8.16.

19 Suppose $T \in \mathcal{L}(V)$ is not nilpotent. Show that

$$V = \operatorname{null} T^{\dim V - 1} \oplus \operatorname{range} T^{\dim V - 1}.$$

For operators that are not nilpotent, this exercise improves 8.4.

20 Suppose V is an inner product space and $T \in \mathcal{L}(V)$ is normal and nilpotent. Prove that $T = 0$.

21 Suppose $T \in \mathcal{L}(V)$ is such that $\operatorname{null} T^{\dim V - 1} \neq \operatorname{null} T^{\dim V}$. Prove that T is nilpotent and that $\dim \operatorname{null} T^k = k$ for every integer k with $0 \leq k \leq \dim V$.

22 Suppose $T \in \mathcal{L}(\mathbf{C}^5)$ is such that $\operatorname{range} T^4 \neq \operatorname{range} T^5$. Prove that T is nilpotent.

23 Give an example of an operator T on a finite-dimensional real vector space such that 0 is the only eigenvalue of T but T is not nilpotent.

This exercise shows that (b) in 8.17 does not hold without the hypothesis that $\mathbf{F} = \mathbf{C}$.

24 For each item in Example 8.15, find a basis of the domain vector space such that the matrix of the nilpotent operator with respect to that basis has the upper-triangular form promised by 8.18(c).

25 Suppose that V is an inner product space and $T \in \mathcal{L}(V)$ is nilpotent. Show that there is an orthonormal basis of V with respect to which the matrix of T has the upper-triangular form promised by 8.18(c).

12 נניח ש־ $T \in \mathcal{L}(V)$ כך שכל וקטור השונה מאפס ב־ V הוא וקטור עצמי מוכלל של T. הוכיחו כי קיים $\lambda \in \mathbf{F}$ כך ש־ $T - \lambda I$ נילפוטנטי.

13 נניח ש־ $S, T \in \mathcal{L}(V)$ ו־ ST נילפוטנטי. הוכיחו כי TS נילפוטנטי.

14 נניח ש־ $T \in \mathcal{L}(V)$ נילפוטנטי ו־ $T \neq 0$. הוכיחו כי T אינו לכסין.

15 נניח ש־ $\mathbf{F} = \mathbf{C}$ ו־ $T \in \mathcal{L}(V)$. הוכיחו כי T לכסין אם־ורק־אם כל וקטור עצמי מוכלל של T הוא וקטור עצמי של T.

עבור $\mathbf{F} = \mathbf{C}$, תרגיל זה מוסיף שקילות נוספת לרשימת התנאים עבור לכסינות ב־ 5.55.

16 א. הביאו דוגמה של אופרטורים נילפוטנטיים S, T מעל אותו מרחב וקטורי כך שאף אחד מ־ $S + T$ ו־ ST אינו נילפוטנטי.

ב. נניח ש־ $S, T \in \mathcal{L}(V)$ נילפוטנטיים ו־ $ST = TS$. הוכיחו כי $S + T$ ו־ ST נילפוטנטיים.

17 נניח ש־ $T \in \mathcal{L}(V)$ נילפוטנטי ו־ m מספר שלם חיובי כך ש־ $T^m = 0$.

א. הוכיחו כי $I - T$ הפיך וכי $(I - T)^{-1} = I + T + \cdots + T^{m-1}$.

ב. הסבירו כיצד הייתם מנחשים את הנוסחה לעיל.

18 נניח ש־ $T \in \mathcal{L}(V)$ נילפוטנטי. הוכיחו כי $T^{1+\dim \operatorname{range} T} = 0$.

אם $\dim \operatorname{range} T < \dim V - 1$, אזי תרגיל זה משפר את 8.16.

19 נניח ש־ $T \in \mathcal{L}(V)$ אינו נילפוטנטי. הראו כי

$$V = \operatorname{null} T^{\dim V - 1} \oplus \operatorname{range} T^{\dim V - 1}.$$

עבור אופרטורים שאינם נילפוטנטיים, תרגיל זה משפר את 8.4.

20 נניח ש־ V מרחב מכפלה פנימית ו־ $T \in \mathcal{L}(V)$ נורמלי ונילפוטנטי. הוכיחו כי $T = 0$.

21 נניח ש־ $T \in \mathcal{L}(V)$ כך ש־ $\operatorname{null} T^{\dim V - 1} \neq \operatorname{null} T^{\dim V}$. הוכיחו כי T נילפוטנטי וכי $\dim \operatorname{null} T^k = k$ לכל k מספר שלם, כאשר $0 \leq k \leq \dim V$.

22 נניח ש־ $T \in \mathcal{L}(\mathbf{C}^5)$ כך ש־ $\operatorname{range} T^4 \neq \operatorname{range} T^5$. הוכיחו כי T נילפוטנטי.

23 הביאו דוגמה של אופרטור T מעל מרחב וקטורי ממשי מממד סופי כך ש־ 0 הוא הערך העצמי היחיד של T אך T אינו נילפוטנטי.

תרגיל זה מראה שסעיף (ב) ב־ 8.17 אינו מתקיים ללא ההנחה ש־ $\mathbf{F} = \mathbf{C}$.

24 לכל סעיף בדוגמה 8.15, מצאו בסיס של מרחב התחום כך שהמטריצה של האופרטור הנילפוטנטי ביחס לבסיס זה היא משולשית עליונה כמובטח על ידי 8.18(ג).

25 נניח ש־ V מרחב מכפלה פנימית ו־ $T \in \mathcal{L}(V)$ נילפוטנטי. הראו כי ישנו בסיס אורתונורמלי של V אשר ביחס אליו המטריצה של T משולשית עליונה כמובטח על ידי 8.18(ג).

8B Generalized Eigenspace Decomposition

Generalized Eigenspaces

8.19 definition: *generalized eigenspace*, $G(\lambda, T)$

Suppose $T \in \mathcal{L}(V)$ and $\lambda \in \mathbf{F}$. The *generalized eigenspace* of T corresponding to λ, denoted by $G(\lambda, T)$, is defined by

$$G(\lambda, T) = \{v \in V : (T - \lambda I)^k v = 0 \text{ for some positive integer } k\}.$$

Thus $G(\lambda, T)$ is the set of generalized eigenvectors of T corresponding to λ, along with the 0 vector.

Because every eigenvector of T is a generalized eigenvector of T (take $k = 1$ in the definition of generalized eigenvector), each eigenspace is contained in the corresponding generalized eigenspace. In other words, if $T \in \mathcal{L}(V)$ and $\lambda \in \mathbf{F}$, then $E(\lambda, T) \subseteq G(\lambda, T)$.

The next result implies that if $T \in \mathcal{L}(V)$ and $\lambda \in \mathbf{F}$, then the generalized eigenspace $G(\lambda, T)$ is a subspace of V (because the null space of each linear map on V is a subspace of V).

8.20 *description of generalized eigenspaces*

Suppose $T \in \mathcal{L}(V)$ and $\lambda \in \mathbf{F}$. Then $G(\lambda, T) = \operatorname{null}(T - \lambda I)^{\dim V}$.

Proof Suppose $v \in \operatorname{null}(T - \lambda I)^{\dim V}$. The definitions imply $v \in G(\lambda, T)$. Thus $G(\lambda, T) \supseteq \operatorname{null}(T - \lambda I)^{\dim V}$.

Conversely, suppose $v \in G(\lambda, T)$. Thus there is a positive integer k such that $v \in \operatorname{null}(T - \lambda I)^k$. From 8.1 and 8.3 (with $T - \lambda I$ replacing T), we get $v \in \operatorname{null}(T - \lambda I)^{\dim V}$. Thus $G(\lambda, T) \subseteq \operatorname{null}(T - \lambda I)^{\dim V}$, completing the proof. ∎

8.21 example: *generalized eigenspaces of an operator on* $\mathbf{C}^3$

Define $T \in \mathcal{L}(\mathbf{C}^3)$ by

$$T(z_1, z_2, z_3) = (4z_2, 0, 5z_3).$$

In Example 8.10, we saw that the eigenvalues of T are 0 and 5, and we found the corresponding sets of generalized eigenvectors. Taking the union of those sets with $\{0\}$, we have

$$G(0, T) = \{(z_1, z_2, 0) : z_1, z_2 \in \mathbf{C}\} \quad \text{and} \quad G(5, T) = \{(0, 0, z_3) : z_3 \in \mathbf{C}\}.$$

Note that $\mathbf{C}^3 = G(0, T) \oplus G(5, T)$.

חג פירוק למרחבים עצמיים מוכללים

מרחבים עצמיים מוכללים

8.19 הגדרה: מרחב עצמי מוכלל, $G(\lambda, T)$

נניח ש־ $T \in \mathcal{L}(V)$ ו־ $\lambda \in \mathbf{F}$. המרחב העצמי המוכלל של T השייך ל־ λ, ומסומן על ידי $G(\lambda, T)$, מוגדר על ידי

$$G(\lambda, T) = \{v \in V : (T - \lambda I)^k v = 0, \text{ שלם חיובי כלשהו } k\}.$$

לפיכך $G(\lambda, T)$ הוא קבוצת הווקטורים העצמיים המוכללים של T השייכים ל־ λ, יחד עם וקטור האפס.

מאחר שכל וקטור עצמי של T הוא וקטור עצמי מוכלל של T (הציבו $k = 1$ בהגדרה של וקטור עצמי מוכלל), כל מרחב עצמי מוכל במרחב העצמי המוכלל המתאים. במילים אחרות, אם $T \in \mathcal{L}(V)$ ו־ $\lambda \in \mathbf{F}$, אזי $E(\lambda, T) \subseteq G(\lambda, T)$.

המסקנה הבאה גוררת שאם $T \in \mathcal{L}(V)$ ו־ $\lambda \in \mathbf{F}$, אזי המרחב העצמי המוכלל $G(\lambda, T)$ הוא תת־מרחב של V (מאחר שמרחב האפס של כל העתקה ליניארית מעל V הוא תת־מרחב של V).

8.20 תיאור של מרחבים עצמיים מוכללים

נניח ש־ $T \in \mathcal{L}(V)$ ו־ $\lambda \in \mathbf{F}$. אזי $G(\lambda, T) = \operatorname{null}(T - \lambda I)^{\dim V}$.

הוכחה נניח ש־ $v \in \operatorname{null}(T - \lambda I)^{\dim V}$. מההגדרות נובע ש־ $v \in G(\lambda, T)$. לפיכך $\operatorname{null}(T - \lambda I)^{\dim V} \subseteq G(\lambda, T)$.

בכיוון ההפוך, נניח ש־ $v \in G(\lambda, T)$. לפיכך קיים מספר שלם חיובי k כך ש־ $v \in \operatorname{null}(T - \lambda I)^k$. מ־ 8.1 ו־ 8.3 (עם הצבת $T - \lambda I$ במקום T), נובע ש־ $v \in \operatorname{null}(T - \lambda I)^{\dim V}$. לפיכך $G(\lambda, T) \subseteq \operatorname{null}(T - \lambda I)^{\dim V}$, ובכך הושלמה ההוכחה. ■

8.21 דוגמה: מרחבים עצמיים מוכללים של אופרטור מעל $\mathbf{C}^3$

נגדיר $T \in \mathcal{L}(\mathbf{C}^3)$ על ידי

$$T(z_1, z_2, z_3) = (4z_2, 0, 5z_3).$$

בדוגמה 8.10, ראינו שהערכים העצמיים של T הם 0 ו־ 5, ומצאנו את קבוצות הווקטורים העצמיים המוכללים השייכים להם. ניקח את האיחוד של קבוצות אלו עם $\{0\}$ ונקבל

$$G(5, T) = \{(0, 0, z_3) : z_3 \in \mathbf{C}\} \quad \text{וגם} \quad G(0, T) = \{(z_1, z_2, 0) : z_1, z_2 \in \mathbf{C}\}.$$

שימו לב ש־ $\mathbf{C}^3 = G(0, T) \oplus G(5, T)$.

In Example 8.21, the domain space $\mathbf{C}^3$ is the direct sum of the generalized eigenspaces of the operator T in that example. Our next result shows that this behavior holds in general. Specifically, the following major result shows that if $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$, then V is the direct sum of the generalized eigenspaces of T, each of which is invariant under T and on which T is a nilpotent operator plus a scalar multiple of the identity. Thus the next result achieves our goal of decomposing V into invariant subspaces on which T has a known behavior.

As we will see, the proof follows from putting together what we have learned about generalized eigenspaces and then using our result that for each operator $T \in \mathcal{L}(V)$, there exists a basis of V consisting of generalized eigenvectors of T.

8.22 *generalized eigenspace decomposition*

Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$. Let $\lambda_1, \dots, \lambda_m$ be the distinct eigenvalues of T. Then

(a) $G(\lambda_k, T)$ is invariant under T for each $k = 1, \dots, m$;

(b) $(T - \lambda_k I)|_{G(\lambda_k, T)}$ is nilpotent for each $k = 1, \dots, m$;

(c) $V = G(\lambda_1, T) \oplus \cdots \oplus G(\lambda_m, T)$.

Proof

(a) Suppose $k \in \{1, \dots, m\}$. Then 8.20 shows that

$$G(\lambda_k, T) = \operatorname{null}(T - \lambda_k I)^{\dim V}.$$

Thus 5.18, with $p(z) = (z - \lambda_k)^{\dim V}$, implies that $G(\lambda_k, T)$ is invariant under T, proving (a).

(b) Suppose $k \in \{1, \dots, m\}$. If $v \in G(\lambda_k, T)$, then $(T - \lambda_k I)^{\dim V} v = 0$ (by 8.20). Thus $\left((T - \lambda_k I)|_{G(\lambda_k, T)}\right)^{\dim V} = 0$. Hence $(T - \lambda_k I)|_{G(\lambda_k, T)}$ is nilpotent, proving (b).

(c) To show that $G(\lambda_1, T) + \cdots + G(\lambda_m, T)$ is a direct sum, suppose

$$v_1 + \cdots + v_m = 0,$$

where each v_k is in $G(\lambda_k, T)$. Because generalized eigenvectors of T corresponding to distinct eigenvalues are linearly independent (by 8.12), this implies that each v_k equals 0. Thus $G(\lambda_1, T) + \cdots + G(\lambda_m, T)$ is a direct sum (by 1.45).

Finally, each vector in V can be written as a finite sum of generalized eigenvectors of T (by 8.9). Thus

$$V = G(\lambda_1, T) \oplus \cdots \oplus G(\lambda_m, T),$$

proving (c). ∎

For the analogous result when $\mathbf{F} = \mathbf{R}$, see Exercise 8.

בדוגמה 8.21, מרחב התחום $\mathbf{C}^3$ הוא הסכום הישר של המרחבים העצמיים המוכללים של האופרטור T בדוגמה ההיא. המסקנה הבאה מראה שתופעה זו מתקיימת באופן כללי. ליתר דיוק, המסקנה המרכזית הבאה מראה שאם $\mathbf{F} = \mathbf{C}$ ו־ $T \in \mathcal{L}(V)$, אזי V הוא הסכום הישר של המרחבים העצמיים המוכללים של T, כאשר כל אחד מהם שמור תחת T ומעליו T הוא אופרטור נילפוטנטי בתוספת כפולת סקלר של אופרטור הזהות. לפיכך המסקנה הבאה משיגה את מטרתנו לפירוק V לתתי־מרחב שמורים אשר מעליהם התנהגות T ידועה.

כפי שנראה, ההוכחה נובעת מצירוף מה שלמדנו על מרחבים עצמיים מוכללים ולאחר מכן שימוש במסקנה זו שלפיה לכל אופרטור $T \in \mathcal{L}(V)$, קיים בסיס של V המורכב מווקטורים עצמיים מוכללים של T.

8.22 פירוק למרחבים עצמיים מוכללים

נניח ש־ $\mathbf{F} = \mathbf{C}$ ו־ $T \in \mathcal{L}(V)$. יהיו $\lambda_1, \dots, \lambda_m$ הערכים העצמיים השונים של T. אזי

א. $G(\lambda_k, T)$ שמור תחת T לכל $k = 1, \dots, m$

ב. $(T - \lambda_k I)|_{G(\lambda_k, T)}$ נילפוטנטי לכל $k = 1, \dots, m$

ג. $V = G(\lambda_1, T) \oplus \cdots \oplus G(\lambda_m, T)$.

הוכחה

א. נניח ש־ $k \in \{1, \dots, m\}$. אזי מ־ 8.20 נובע ש־

$$G(\lambda_k, T) = \operatorname{null}(T - \lambda_k I)^{\dim V}.$$

לפיכך 5.18, עם $p(z) = (z - \lambda_k)^{\dim V}$, גורר ש־ $G(\lambda_k, T)$ שמור תחת T, ובכך הוכחנו את (א).

ב. נניח ש־ $k \in \{1, \dots, m\}$. אם $v \in G(\lambda_k, T)$, אזי $(T - \lambda_k I)^{\dim V} v = 0$ (לפי 8.20). לפיכך $((T - \lambda_k I)|_{G(\lambda_k, T)})^{\dim V} = 0$. לכן $(T - \lambda_k I)|_{G(\lambda_k, T)}$ נילפוטנטי, ובכך הוכחנו את (ב).

ג. כדי להראות כי $G(\lambda_1, T) + \cdots + G(\lambda_m, T)$ סכום ישר, נניח ש־

$$v_1 + \cdots + v_m = 0,$$

כאשר כל v_k ב־ $G(\lambda_k, T)$. מאחר שווקטורים עצמיים מוכללים של T השייכים לערכים עצמיים שונים בלתי תלויים ליניארית (לפי 8.12), מכאן נובע שכל v_k שווה 0. לפיכך $G(\lambda_1, T) + \cdots + G(\lambda_m, T)$ סכום ישר (לפי 1.45).

לסיום, כל וקטור ב־ V ניתן לכתיבה כסכום סופי של וקטורים עצמיים מוכללים של T (לפי 8.9). לפיכך

$$V = G(\lambda_1, T) \oplus \cdots \oplus G(\lambda_m, T),$$

ובכך הוכחנו את (ג).

למסקנה האנלוגית כאשר $\mathbf{F} = \mathbf{R}$, ראו תרגיל 8.

Multiplicity of an Eigenvalue

If V is a complex vector space and $T \in \mathcal{L}(V)$, then the decomposition of V provided by the generalized eigenspace decomposition (8.22) can be a powerful tool. The dimensions of the subspaces involved in this decomposition are sufficiently important to get a name, which is given in the next definition.

8.23 definition: *multiplicity*

- Suppose $T \in \mathcal{L}(V)$. The *multiplicity* of an eigenvalue λ of T is defined to be the dimension of the corresponding generalized eigenspace $G(\lambda, T)$.
- In other words, the multiplicity of an eigenvalue λ of T equals

$$\dim \operatorname{null}(T - \lambda I)^{\dim V}.$$

The second bullet point above holds because $G(\lambda, T) = \operatorname{null}(T - \lambda I)^{\dim V}$ (see 8.20).

8.24 example: *multiplicity of each eigenvalue of an operator*

Suppose $T \in \mathcal{L}(\mathbf{C}^3)$ is defined by

$$T(z_1, z_2, z_3) = (6z_1 + 3z_2 + 4z_3, 6z_2 + 2z_3, 7z_3).$$

The matrix of T (with respect to the standard basis) is

$$\begin{pmatrix} 6 & 3 & 4 \\ 0 & 6 & 2 \\ 0 & 0 & 7 \end{pmatrix}.$$

The eigenvalues of T are the diagonal entries 6 and 7, as follows from 5.41. You can verify that the generalized eigenspaces of T are as follows:

$$G(6, T) = \operatorname{span}((1, 0, 0), (0, 1, 0)) \quad \text{and} \quad G(7, T) = \operatorname{span}((10, 2, 1)).$$

Thus the eigenvalue 6 has multiplicity 2 and the eigenvalue 7 has multiplicity 1. The direct sum $\mathbf{C}^3 = G(6, T) \oplus G(7, T)$ is the generalized eigenspace decomposition promised by 8.22. A basis of $\mathbf{C}^3$ consisting of generalized eigenvectors of T, as promised by 8.9, is $(1, 0, 0), (0, 1, 0), (10, 2, 1)$. There does not exist a basis of $\mathbf{C}^3$ consisting of eigenvectors of this operator.

In this example, the multiplicity of each eigenvalue equals the number of times that eigenvalue appears on the diagonal of an upper-triangular matrix representing the operator. This behavior always happens, as we will see in 8.31.

In the example above, the sum of the multiplicities of the eigenvalues of T equals 3, which is the dimension of the domain of T. The next result shows that this holds for all operators on finite-dimensional complex vector spaces.

ריבוי של ערך עצמי

אם V מרחב וקטורי מרוכב ו־ $T \in \mathcal{L}(V)$, אזי הפירוק של V הניתן על ידי הפירוק למרחבים עצמיים מוכללים (8.22) יכול לשמש ככלי רב־עוצמה. ממדי תתי־המרחב הכלולים בפירוק זה חשובים דיים לקבלת שם ייעודי, הניתן להם בהגדרה הבאה.

8.23 הגדרה: ריבוי

- נניח ש־ $T \in \mathcal{L}(V)$. הריבוי של ערך עצמי λ של T מוגדר כממד של המרחב העצמי המוכלל המתאים $G(\lambda, T)$.
- במילים אחרות, הריבוי של ערך עצמי λ של T שווה ל־
$$\dim \operatorname{null}(T - \lambda I)^{\dim V}.$$

הסעיף השני בהגדרה לעיל מתקיים מאחר ש־ $G(\lambda, T) = \operatorname{null}(T - \lambda I)^{\dim V}$ (ראו 8.20).

8.24 דוגמה: ריבוי של כל ערך עצמי של אופרטור

נניח ש־ $T \in \mathcal{L}(\mathbf{C}^3)$ מוגדר על ידי
$$T(z_1, z_2, z_3) = (6z_1 + 3z_2 + 4z_3, 6z_2 + 2z_3, 7z_3).$$
המטריצה של T (ביחס לבסיס הסטנדרטי) היא
$$\begin{pmatrix} 6 & 3 & 4 \\ 0 & 6 & 2 \\ 0 & 0 & 7 \end{pmatrix}.$$
הערכים העצמיים של T הם איברי האלכסון 6 ו־ 7, כפי שנובע מ־ 5.41. תוכלו לוודא כי המרחבים העצמיים המוכללים של T הם כדלהלן:
$$G(6, T) = \operatorname{span}((1, 0, 0), (0, 1, 0)) \quad \text{וגם} \quad G(7, T) = \operatorname{span}((10, 2, 1)).$$

בדוגמה זו, הריבוי של כל ערך עצמי שווה למספר הפעמים שהערך העצמי מופיע על האלכסון של מטריצה משולשית עליונה המייצגת את האופרטור. התנהגות זו מתרחשת תמיד, כפי שנראה ב־ 8.31.

לפיכך הריבוי של הערך העצמי 6 הוא 2 והריבוי של הערך העצמי 7 הוא 1. הסכום הישר $\mathbf{C}^3 = G(6, T) \oplus G(7, T)$ הוא הפירוק למרחבים עצמיים מוכללים כפי שהובטח על ידי 8.22. בסיס של $\mathbf{C}^3$ המורכב מווקטורים עצמיים מוכללים של T, כפי שהובטח על ידי 8.9, הוא $(1, 0, 0), (0, 1, 0), (10, 2, 1)$. לא קיים בסיס של $\mathbf{C}^3$ המורכב מווקטורים עצמיים של אופרטור זה.

בדוגמה לעיל, סכום הריבויים של הערכים העצמיים של T שווה 3, שהוא ממד התחום של T. המסקנה הבאה מראה שהדבר מתקיים לכל האופרטורים מעל מרחב וקטורי מרוכב מממד סופי.

8.25 *sum of the multiplicities equals* $\dim V$

Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$. Then the sum of the multiplicities of all eigenvalues of T equals $\dim V$.

Proof The desired result follows from the generalized eigenspace decomposition (8.22) and the formula for the dimension of a direct sum (see 3.94). ∎

The terms *algebraic multiplicity* and *geometric multiplicity* are used in some books. In case you encounter this terminology, be aware that the algebraic multiplicity is the same as the multiplicity defined here and the geometric multiplicity is the dimension of the corresponding eigenspace. In other words, if $T \in \mathcal{L}(V)$ and λ is an eigenvalue of T, then

$$\text{algebraic multiplicity of } \lambda = \dim \operatorname{null}(T - \lambda I)^{\dim V} = \dim G(\lambda, T),$$

$$\text{geometric multiplicity of } \lambda = \dim \operatorname{null}(T - \lambda I) = \dim E(\lambda, T).$$

Note that as defined above, the algebraic multiplicity also has a geometric meaning as the dimension of a certain null space. The definition of multiplicity given here is cleaner than the traditional definition that involves determinants; 9.62 implies that these definitions are equivalent.

If V is an inner product space, $T \in \mathcal{L}(V)$ is normal, and λ is an eigenvalue of T, then the algebraic multiplicity of λ equals the geometric multiplicity of λ, as can be seen from applying Exercise 27 in Section 7A to the normal operator $T - \lambda I$. As a special case, the singular values of $S \in \mathcal{L}(V, W)$ (here V and W are both finite-dimensional inner product spaces) depend on the multiplicities (either algebraic or geometric) of the eigenvalues of the self-adjoint operator S^*S.

The next definition associates a monic polynomial with each operator on a finite-dimensional complex vector space.

8.26 definition: *characteristic polynomial*

Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$. Let $\lambda_1, \dots, \lambda_m$ denote the distinct eigenvalues of T, with multiplicities $d_1, \dots, d_m$. The polynomial

$$(z - \lambda_1)^{d_1} \cdots (z - \lambda_m)^{d_m}$$

is called the *characteristic polynomial* of T.

8.27 example: *the characteristic polynomial of an operator*

Suppose $T \in \mathcal{L}(\mathbf{C}^3)$ is defined as in Example 8.24. Because the eigenvalues of T are 6, with multiplicity 2, and 7, with multiplicity 1, we see that the characteristic polynomial of T is $(z - 6)^2(z - 7)$.

8.25 סכום של הריבויים שווה ל־ $\dim V$

נניח ש־ $\mathbf{F} = \mathbf{C}$ ו־ $T \in \mathcal{L}(V)$. אזי סכום הריבויים של כל הערכים העצמיים של T שווה ל־ $\dim V$.

הוכחה התוצאה הרצויה נובעת מהפירוק למרחבים עצמיים מוכללים (8.22) ומהנוסחה עבור הממד של סכום ישר (ראו 3.94).

ישנם ספרים המשתמשים במונחים ריבוי אלגברי וריבוי גאומטרי. אם נתקלתם במונחים אלו, שימו לב כי הריבוי האלגברי הוא זהה לריבוי המוגדר כאן והריבוי הגאומטרי הוא הממד של המרחב העצמי המתאים. במילים אחרות, אם $T \in \mathcal{L}(V)$ ו־ λ ערך עצמי של T, אזי

$$\text{ריבוי אלגברי של } \lambda = \dim \operatorname{null}(T - \lambda I)^{\dim V} = \dim G(\lambda, T),$$

$$\text{ריבוי גאומטרי של } \lambda = \dim \operatorname{null}(T - \lambda I) = \dim E(\lambda, T).$$

שימו לב כי, כפי שהוגדר לעיל, לריבוי האלגברי יש בנוסף משמעות גאומטרית כממד של מרחב אפס מסוים. ההגדרה של ריבוי שניתנה כאן היא אלגנטית יותר מההגדרה המסורתית המערבת דטרמיננטות; 9.62 גורר שהגדרות אלו שקולות.

אם V מרחב מכפלה פנימית, $T \in \mathcal{L}(V)$ נורמלי, ו־ λ ערך עצמי של T, אזי הריבוי האלגברי של λ שווה לריבוי הגאומטרי של λ, כפי שניתן לראות מהפעלת תרגיל 27 ביחידה 7ז על האופרטור הנורמלי $T - \lambda I$. כמקרה מיוחד, הערכים הסינגולריים של $S \in \mathcal{L}(V, W)$ (כאן V ו־ W שניהם מרחבי מכפלה פנימית מממד סופי) תלויים בריבויים (האלגברי או הגאומטרי) של הערכים העצמיים של האופרטור הצמוד לעצמו S^*S.

ההגדרה הבאה משייכת פולינום מתוקן לכל אופרטור מעל מרחב וקטורי מרוכב מממד סופי.

8.26 הגדרה: פולינום אופייני

נניח ש־ $\mathbf{F} = \mathbf{C}$ ו־ $T \in \mathcal{L}(V)$. נסמן ב־ $\lambda_1, \dots, \lambda_m$ את הערכים העצמיים השונים של T, עם הריבויים $d_1, \dots, d_m$. הפולינום

$$(z - \lambda_1)^{d_1} \cdots (z - \lambda_m)^{d_m}$$

נקרא הפולינום האופייני של T.

8.27 דוגמה: הפולינום האופייני של אופרטור

נניח ש־ $T \in \mathcal{L}(\mathbf{C}^3)$ מוגדר כמו בדוגמה 8.24. מאחר שהערכים העצמיים של T הם 6, עם ריבוי 2, ו־ 7, עם ריבוי 1, אנו רואים כי הפולינום האופייני של T הוא $(z - 6)^2(z - 7)$.

8.28 *degree and zeros of characteristic polynomial*

Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$. Then

(a) the characteristic polynomial of T has degree $\dim V$;

(b) the zeros of the characteristic polynomial of T are the eigenvalues of T.

Proof Our result about the sum of the multiplicities (8.25) implies (a). The definition of the characteristic polynomial implies (b). ∎

Most texts define the characteristic polynomial using determinants (the two definitions are equivalent by 9.62). The approach taken here, which is considerably simpler, leads to the following nice proof of the Cayley–Hamilton theorem.

8.29 *Cayley–Hamilton theorem*

Suppose $\mathbf{F} = \mathbf{C}$, $T \in \mathcal{L}(V)$, and q is the characteristic polynomial of T. Then $q(T) = 0$.

Proof Let $\lambda_1, \dots, \lambda_m$ be the distinct eigenvalues of T, and let $d_k = \dim G(\lambda_k, T)$. For each $k \in \{1, \dots, m\}$, we know that $(T - \lambda_k I)|_{G(\lambda_k, T)}$ is nilpotent. Thus we have

$$(T - \lambda_k I)^{d_k}|_{G(\lambda_k, T)} = 0$$

(by 8.16) for each $k \in \{1, \dots, m\}$.

Arthur Cayley (1821–1895) published three mathematics papers before completing his undergraduate degree.

The generalized eigenspace decomposition (8.22) states that every vector in V is a sum of vectors in $G(\lambda_1, T), \dots, G(\lambda_m, T)$. Thus to prove that $q(T) = 0$, we only need to show that $q(T)|_{G(\lambda_k, T)} = 0$ for each k.

Fix $k \in \{1, \dots, m\}$. We have

$$q(T) = (T - \lambda_1 I)^{d_1} \cdots (T - \lambda_m I)^{d_m}.$$

The operators on the right side of the equation above all commute, so we can move the factor $(T - \lambda_k I)^{d_k}$ to be the last term in the expression on the right. Because $(T - \lambda_k I)^{d_k}|_{G(\lambda_k, T)} = 0$, we have $q(T)|_{G(\lambda_k, T)} = 0$, as desired. ∎

The next result implies that if the minimal polynomial of an operator $T \in \mathcal{L}(V)$ has degree $\dim V$ (as happens almost always—see the paragraphs following 5.24), then the characteristic polynomial of T equals the minimal polynomial of T.

8.30 *characteristic polynomial is a multiple of minimal polynomial*

Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$. Then the characteristic polynomial of T is a polynomial multiple of the minimal polynomial of T.

Proof The desired result follows immediately from the Cayley–Hamilton theorem (8.29) and 5.29. ∎

8.28 מעלה ושורשים של פולינום אופייני

נניח ש־ $\mathbf{F} = \mathbf{C}$ ו־ $T \in \mathcal{L}(V)$. אזי

א. מעלת הפולינום האופייני של T היא $\dim V$

ב. השורשים של הפולינום האופייני של T הם הערכים העצמיים של T.

הוכחה מסקנתנו בנוגע לסכום הריבויים (8.25) גוררת את (א). הגדרת הפולינום האופייני גוררת את (ב). ■

ברוב ספרי הלימוד הפולינום האופייני מוגדר תוך שימוש בדטרמיננטות (שתי ההגדרות שקולות לפי 9.62). הגישה הננקטת כאן, שהיא במידה משמעותית פשוטה יותר, מובילה אל ההוכחה הנחמדה הבאה של משפט קיילי־המילטון.

8.29 משפט קיילי־המילטון

נניח ש־ $\mathbf{F} = \mathbf{C}$, $T \in \mathcal{L}(V)$, ו־ q הפולינום האופייני של T. אזי $q(T) = 0$.

הוכחה יהיו $\lambda_1, \dots, \lambda_m$ הערכים העצמיים השונים של T, ויהי $d_k = \dim G(\lambda_k, T)$. לכל $k \in \{1, \dots, m\}$, אנו יודעים כי $(T - \lambda_k I)|_{G(\lambda_k, T)}$ נילפוטנטי. לפיכך מתקיים

$$(T - \lambda_k I)^{d_k}|_{G(\lambda_k, T)} = 0$$

(לפי 8.16) לכל $k \in \{1, \dots, m\}$.

ארתור קיילי (1895–1821) פרסם שלושה מאמרים במתמטיקה טרם סיים את התואר הראשון שלו.

הפירוק למרחבים עצמיים מוכללים (8.22) קובע כי כל וקטור ב־ V הוא סכום של וקטורים ב־ $G(\lambda_1, T), \dots, G(\lambda_m, T)$. לפיכך כדי להוכיח כי $q(T) = 0$, כל שעלינו להראות הוא כי $q(T)|_{G(\lambda_k, T)} = 0$ לכל k. נְקַבֵּעַ $k \in \{1, \dots, m\}$ ונקבל

$$q(T) = (T - \lambda_1 I)^{d_1} \cdots (T - \lambda_m I)^{d_m}.$$

האופרטורים באגף הימני של המשוואה לעיל כולם קומוטטיביים, כך שאנו יכולים להזיז את הגורם $(T - \lambda_k I)^{d_k}$ כך שיהיה האיבר האחרון בביטוי באגף הימני. מאחר ש־ $(T - \lambda_k I)^{d_k}|_{G(\lambda_k, T)} = 0$, מתקיים $q(T)|_{G(\lambda_k, T)} = 0$, כפי שרצינו. ■

המסקנה הבאה גוררת שאם מעלת הפולינום המינימלי של אופרטור $T \in \mathcal{L}(V)$ היא $\dim V$ (כפי שקורה ברוב המכריע של המקרים—ראו הפסקאות שאחרי 5.24), אזי הפולינום האופייני של T שווה לפולינום המינימלי של T.

8.30 הפולינום האופייני הוא כפולה של הפולינום המינימלי

נניח ש־ $\mathbf{F} = \mathbf{C}$ ו־ $T \in \mathcal{L}(V)$. אזי הפולינום האופייני של T הוא כפולת פולינום של הפולינום המינימלי של T.

הוכחה התוצאה הרצויה נובעת ישירות ממשפט קיילי־המילטון (8.29) ו־ 5.29. ■

Now we can prove that the result suggested by Example 8.24 holds for all operators on finite-dimensional complex vector spaces.

8.31 *multiplicity of an eigenvalue equals number of times on diagonal*

Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$. Suppose $v_1, \dots, v_n$ is a basis of V such that $\mathcal{M}(T, (v_1, \dots, v_n))$ is upper triangular. Then the number of times that each eigenvalue λ of T appears on the diagonal of $\mathcal{M}(T, (v_1, \dots, v_n))$ equals the multiplicity of λ as an eigenvalue of T.

Proof Let $A = \mathcal{M}(T, (v_1, \dots, v_n))$. Thus A is an upper-triangular matrix. Let $\lambda_1, \dots, \lambda_n$ denote the entries on the diagonal of A. Thus for each $k \in \{1, \dots, n\}$, we have

$$Tv_k = u_k + \lambda_k v_k$$

for some $u_k \in \operatorname{span}(v_1, \dots, v_{k-1})$. Hence if $k \in \{1, \dots, n\}$ and $\lambda_k \neq 0$, then Tv_k is not a linear combination of $Tv_1, \dots, Tv_{k-1}$. The linear dependence lemma (2.19) now implies that the list of those Tv_k such that $\lambda_k \neq 0$ is linearly independent.

Let d denote the number of indices $k \in \{1, \dots, n\}$ such that $\lambda_k = 0$. The conclusion of the previous paragraph implies that

$$\dim \operatorname{range} T \geq n - d.$$

Because $n = \dim V = \dim \operatorname{null} T + \dim \operatorname{range} T$, the inequality above implies that

8.32 $$\dim \operatorname{null} T \leq d.$$

The matrix of the operator T^n with respect to the basis $v_1, \dots, v_n$ is the upper-triangular matrix A^n, which has diagonal entries $\lambda_1^n, \dots, \lambda_n^n$ [see Exercise 2(b) in Section 5C]. Because $\lambda_k^n = 0$ if and only if $\lambda_k = 0$, the number of times that 0 appears on the diagonal of A^n equals d. Thus applying 8.32 with T replaced with T^n, we have

8.33 $$\dim \operatorname{null} T^n \leq d.$$

For λ an eigenvalue of T, let m_λ denote the multiplicity of λ as an eigenvalue of T and let d_λ denote the number of times that λ appears on the diagonal of A. Replacing T in 8.33 with $T - \lambda I$, we see that

8.34 $$m_\lambda \leq d_\lambda$$

for each eigenvalue λ of T. The sum of the multiplicities m_λ over all eigenvalues λ of T equals n, the dimension of V (by 8.25). The sum of the numbers d_λ over all eigenvalues λ of T also equals n, because the diagonal of A has length n.

Thus summing both sides of 8.34 over all eigenvalues λ of T produces an equality. Hence 8.34 must actually be an equality for each eigenvalue λ of T. Thus the multiplicity of λ as an eigenvalue of T equals the number of times that λ appears on the diagonal of A, as desired. ∎

כעת נוכל להוכיח כי התוצאה המוצעת בדוגמה 8.24 מתקיימת לכל האופרטורים מעל מרחבים וקטוריים מרוכבים מממד סופי.

8.31 ריבוי של ערך עצמי שווה למספר המופעים על האלכסון

נניח ש־ $\mathbf{F} = \mathbf{C}$ ו־ $T \in \mathcal{L}(V)$. נניח ש־ $v_1, \dots, v_n$ בסיס של V כך ש־ $\mathcal{M}(T, (v_1, \dots, v_n))$ משולשית עליונה. אזי מספר הפעמים שכל ערך עצמי λ של T מופיע על האלכסון של $\mathcal{M}(T, (v_1, \dots, v_n))$ שווה לריבוי של λ כערך עצמי של T.

הוכחה יהי $A = \mathcal{M}(T, (v_1, \dots, v_n))$. לפיכך A מטריצה משולשית עליונה. נסמן ב־ $\lambda_1, \dots, \lambda_n$ את איברי האלכסון של A. לפיכך לכל $k \in \{1, \dots, n\}$, נקבל

$$Tv_k = u_k + \lambda_k v_k$$

עבור $u_k \in \operatorname{span}(v_1, \dots, v_{k-1})$ כלשהו. לכן אם $k \in \{1, \dots, n\}$ ו־ $\lambda_k \neq 0$, אזי Tv_k אינו צירוף ליניארי של $Tv_1, \dots, Tv_{k-1}$. למת התלות הליניארית (2.19) כעת גוררת שהרשימה של Tv_k אלו כאשר $\lambda_k \neq 0$ בלתי תלויה ליניארית.

נסמן ב־ d את מספר האינדקסים $k \in \{1, \dots, n\}$ כך ש־ $\lambda_k = 0$. המסקנה בפסקה הקודמת גוררת ש־

$$\dim \operatorname{range} T \geq n - d.$$

מאחר ש־ $n = \dim V = \dim \operatorname{null} T + \dim \operatorname{range} T$, האי־שוויון לעיל גורר ש־

$$\dim \operatorname{null} T \leq d. \tag{8.32}$$

המטריצה של האופרטור T^n ביחס אל הבסיס $v_1, \dots, v_n$ היא מטריצה משולשית עליונה A^n, שאברי האלכסון שלה הם $\lambda_1^n, \dots, \lambda_n^n$ [ראו תרגיל 2(ב) ביחידה ה3]. מאחר ש־ $\lambda_k^n = 0$ אם־ורק־אם $\lambda_k = 0$, מספר הפעמים ש־ 0 מופיע על האלכסון של A^n שווה d. לפיכך הפעלת 8.32 כאשר T מוחלף ב־ T^n נותנת לנו

$$\dim \operatorname{null} T^n \leq d. \tag{8.33}$$

עבור λ ערך עצמי של T, נסמן ב־ m_λ את הריבוי של λ כערך עצמי של T ונסמן ב־ d_λ את מספר הפעמים ש־ λ מופיע על האלכסון של A. כשנחליף T ב־ 8.33 ב־ $T - \lambda I$, אנו רואים כי

$$m_\lambda \leq d_\lambda \tag{8.34}$$

לכל ערך עצמי λ של T. סכום הריבויים m_λ של כל הערכים העצמיים λ של T שווה n, הממד של V (לפי 8.25). סכום המספרים d_λ של הערכים העצמיים λ של T גם הוא שווה n, מאחר שאורך האלכסון של A הוא n.

לפיכך סכימת שני האגפים של 8.34 מעל כל הערכים העצמיים λ של T מניבה שוויון. לכן 8.34 חייב להוות שוויון לכל ערך עצמי λ של T. לפיכך הריבוי של λ כערך עצמי של T שווה למספר הפעמים ש־ λ מופיע על האלכסון של A, כפי שרצינו. ■

Block Diagonal Matrices

Often we can understand a matrix better by thinking of it as composed of smaller matrices.

To interpret our results in matrix form, we make the following definition, generalizing the notion of a diagonal matrix. If each matrix A_k in the definition below is a 1-by-1 matrix, then we actually have a diagonal matrix.

8.35 definition: *block diagonal matrix*

A *block diagonal matrix* is a square matrix of the form

$$\begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_m \end{pmatrix},$$

where $A_1, \dots, A_m$ are square matrices lying along the diagonal and all other entries of the matrix equal 0.

8.36 example: *a block diagonal matrix*

The 5-by-5 matrix

$$A = \begin{pmatrix} \begin{pmatrix} 4 \end{pmatrix} & 0 & 0 & 0 & 0 \\ 0 & \begin{pmatrix} 2 & -3 \\ 0 & 2 \end{pmatrix} & & 0 & 0 \\ 0 & & & 0 & 0 \\ 0 & 0 & 0 & \begin{pmatrix} 1 & 7 \\ 0 & 1 \end{pmatrix} & \\ 0 & 0 & 0 & & \end{pmatrix}$$

is a block diagonal matrix with

$$A = \begin{pmatrix} A_1 & & 0 \\ & A_2 & \\ 0 & & A_3 \end{pmatrix},$$

where

$$A_1 = \begin{pmatrix} 4 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 2 & -3 \\ 0 & 2 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 1 & 7 \\ 0 & 1 \end{pmatrix}.$$

Here the inner matrices in the 5-by-5 matrix above are blocked off to show how we can think of it as a block diagonal matrix.

Note that in the example above, each of A_1, A_2, A_3 is an upper-triangular matrix whose diagonal entries are all equal. The next result shows that with respect to an appropriate basis, every operator on a finite-dimensional complex vector space has a matrix of this form. Note that this result gives us many more zeros in the matrix than are needed to make it upper triangular.

מטריצות בלוקים אלכסוניות

לעיתים קרובות נוכל להבין מטריצה טוב יותר בכך שנחשוב עליה כמורכבת ממטריצות קטנות יותר.

כדי להעניק פרשנות למסקנות שלנו בתבנית של מטריצה, אנו מנסחים את ההגדרה להלן, המכלילה את רעיון המטריצה האלכסונית. אם כל מטריצה A_k בהגדרה להלן היא מטריצה 1×1, אזי מתקבלת בפועל מטריצה אלכסונית.

8.35 הגדרה: מטריצת בלוקים אלכסונית

מטריצת בלוקים אלכסונית היא מטריצה ריבועית מהצורה

$$\begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_m \end{pmatrix},$$

כאשר $A_1, \dots, A_m$ הן מטריצות ריבועיות המוצבות לאורך האלכסון וכל שאר איברי המטריצה שווים 0.

8.36 דוגמה: מטריצת בלוקים אלכסונית

המטריצה 5×5

$$A = \begin{pmatrix} \begin{pmatrix} 4 \end{pmatrix} & \begin{matrix} 0 & 0 \end{matrix} & \begin{matrix} 0 & 0 \\ 0 & 0 \end{matrix} \\ \begin{matrix} 0 \\ 0 \\ 0 \end{matrix} & \begin{pmatrix} 2 & -3 \\ 0 & 2 \end{pmatrix} & \begin{matrix} 0 & 0 \\ 0 & 0 \end{matrix} \\ \begin{matrix} 0 \\ 0 \end{matrix} & \begin{matrix} 0 & 0 \\ 0 & 0 \end{matrix} & \begin{pmatrix} 1 & 7 \\ 0 & 1 \end{pmatrix} \end{pmatrix}$$

היא מטריצת בלוקים אלכסונית עם

$$A = \begin{pmatrix} A_1 & & 0 \\ & A_2 & \\ 0 & & A_3 \end{pmatrix},$$

כאשר

$$A_1 = \begin{pmatrix} 4 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 2 & -3 \\ 0 & 2 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 1 & 7 \\ 0 & 1 \end{pmatrix}.$$

במקרה שלנו, תחמנו את המטריצות הפנימיות במטריצה 5×5 לעיל בבלוקים כדי להמחיש כיצד ניתן לחשוב עליה כמטריצת בלוקים אלכסונית.

שימו לב שבדוגמה לעיל, כל אחת מהמטריצות A_1, A_2, A_3 היא מטריצה משולשית עליונה שאברי האלכסון שלה כולם שווים. המסקנה הבאה מראה שביחס לבסיס מתאים, לכל אופרטור מעל מרחב וקטורי מרוכב מממד סופי יש מטריצה מצורה זו. שימו לב שכתוצאה מכך במטריצה יש יותר אפסים מאלו הנדרשים לקיומה של מטריצה משולשית עליונה.

8.37 *block diagonal matrix with upper-triangular blocks*

Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$. Let $\lambda_1, \dots, \lambda_m$ be the distinct eigenvalues of T, with multiplicities $d_1, \dots, d_m$. Then there is a basis of V with respect to which T has a block diagonal matrix of the form

$$\begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_m \end{pmatrix},$$

where each A_k is a d_k-by-d_k upper-triangular matrix of the form

$$A_k = \begin{pmatrix} \lambda_k & & * \\ & \ddots & \\ 0 & & \lambda_k \end{pmatrix}.$$

Proof Each $(T - \lambda_k I)|_{G(\lambda_k, T)}$ is nilpotent (see 8.22). For each k, choose a basis of $G(\lambda_k, T)$, which is a vector space of dimension d_k, such that the matrix of $(T - \lambda_k I)|_{G(\lambda_k, T)}$ with respect to this basis is as in 8.18(c). Thus with respect to this basis, the matrix of $T|_{G(\lambda_k, T)}$, which equals $(T - \lambda_k I)|_{G(\lambda_k, T)} + \lambda_k I|_{G(\lambda_k, T)}$, looks like the desired form shown above for A_k.

The generalized eigenspace decomposition (8.22) shows that putting together the bases of the $G(\lambda_k, T)$'s chosen above gives a basis of V. The matrix of T with respect to this basis has the desired form. ∎

8.38 example: *block diagonal matrix via generalized eigenvectors*

Let $T \in \mathcal{L}(\mathbf{C}^3)$ be defined by $T(z_1, z_2, z_3) = (6z_1 + 3z_2 + 4z_3, 6z_2 + 2z_3, 7z_3)$. The matrix of T (with respect to the standard basis) is

$$\begin{pmatrix} 6 & 3 & 4 \\ 0 & 6 & 2 \\ 0 & 0 & 7 \end{pmatrix},$$

which is an upper-triangular matrix but is not of the form promised by 8.37.

As we saw in Example 8.24, the eigenvalues of T are 6 and 7; also,

$$G(6, T) = \operatorname{span}((1, 0, 0), (0, 1, 0)) \quad \text{and} \quad G(7, T) = \operatorname{span}((10, 2, 1)).$$

We also saw that a basis of $\mathbf{C}^3$ consisting of generalized eigenvectors of T is

$$(1, 0, 0), (0, 1, 0), (10, 2, 1).$$

The matrix of T with respect to this basis is

$$\begin{pmatrix} \begin{pmatrix} 6 & 3 \\ 0 & 6 \end{pmatrix} & \begin{matrix} 0 \\ 0 \end{matrix} \\ \begin{matrix} 0 & 0 \end{matrix} & \begin{pmatrix} 7 \end{pmatrix} \end{pmatrix},$$

which is a matrix of the block diagonal form promised by 8.37.

8.37 מטריצת בלוקים אלכסונית עם בלוקי מטריצות משולשיות עליונות

נניח ש־ $\mathbf{F} = \mathbf{C}$ ו־ $T \in \mathcal{L}(V)$. יהיו $\lambda_1, \dots, \lambda_m$ הערכים העצמיים השונים של T, עם הריבויים $d_1, \dots, d_m$. אזי קיים בסיס של V אשר ביחס אליו יש ל־ T מטריצת בלוקים אלכסונית במבנה

$$\begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_m \end{pmatrix},$$

כאשר כל A_k הוא מטריצה משולשית עליונה $d_k \times d_k$ במבנה

$$A_k = \begin{pmatrix} \lambda_k & & * \\ & \ddots & \\ 0 & & \lambda_k \end{pmatrix}.$$

הוכחה כל $(T - \lambda_k I)|_{G(\lambda_k, T)}$ הוא נילפוטנטי (ראו 8.22). לכל k, נבחר בסיס של $G(\lambda_k, T)$, שהוא מרחב וקטורי מממד d_k, כך שהמטריצה של $(T - \lambda_k I)|_{G(\lambda_k, T)}$ ביחס לבסיס זה היא כמו ב־ 8.18(ג). לפיכך ביחס לבסיס זה, המטריצה של $T|_{G(\lambda_k, T)}$, השווה ל־ $(T - \lambda_k I)|_{G(\lambda_k, T)} + \lambda_k I|_{G(\lambda_k, T)}$, נראית כמו הצורה הרצויה המוצגת לעיל עבור A_k.

הפירוק למרחבים עצמיים מוכללים (8.22) מראה שצירוף של כל הבסיסים של ה־ $G(\lambda_k, T)$־ים שנבחרו לעיל ייתן בסיס של V. למטריצה של T ביחס לבסיס זה יש את המבנה המבוקש. ■

8.38 דוגמה: מטריצת בלוקים אלכסונית באמצעות וקטורים עצמיים מוכללים

יהי $T \in \mathcal{L}(\mathbf{C}^3)$ מוגדר על ידי $T(z_1, z_2, z_3) = (6z_1 + 3z_2 + 4z_3, 6z_2 + 2z_3, 7z_3)$. המטריצה של T (ביחס לבסיס הסטנדרטי) היא

$$\begin{pmatrix} 6 & 3 & 4 \\ 0 & 6 & 2 \\ 0 & 0 & 7 \end{pmatrix},$$

שהיא מטריצה משולשית עליונה אך אינה מהמבנה שהובטח על ידי 8.37.

כפי שראינו בדוגמה 8.24, הערכים העצמיים של T הם 6 ו־ 7, ובנוסף

$$G(6, T) = \operatorname{span}((1, 0, 0), (0, 1, 0)) \quad \text{וגם} \quad G(7, T) = \operatorname{span}((10, 2, 1)).$$

ראינו בנוסף כי בסיס של $\mathbf{C}^3$ המורכב מווקטורים עצמיים מוכללים של T הוא

$$(1, 0, 0), (0, 1, 0), (10, 2, 1).$$

המטריצה של T ביחס לבסיס זה היא

$$\begin{pmatrix} \begin{pmatrix} 6 & 3 \\ 0 & 6 \end{pmatrix} & \begin{matrix} 0 \\ 0 \end{matrix} \\ \begin{matrix} 0 & 0 \end{matrix} & \begin{pmatrix} 7 \end{pmatrix} \end{pmatrix},$$

שהיא מטריצת בלוקים אלכסונית כפי שהובטח על ידי 8.37.

Exercises 8B

1 Define $T \in \mathcal{L}(\mathbf{C}^2)$ by $T(w, z) = (-z, w)$. Find the generalized eigenspaces corresponding to the distinct eigenvalues of T.

2 Suppose $T \in \mathcal{L}(V)$ is invertible. Prove that $G(\lambda, T) = G\left(\frac{1}{\lambda}, T^{-1}\right)$ for every $\lambda \in \mathbf{F}$ with $\lambda \neq 0$.

3 Suppose $T \in \mathcal{L}(V)$. Suppose $S \in \mathcal{L}(V)$ is invertible. Prove that T and $S^{-1}TS$ have the same eigenvalues with the same multiplicities.

4 Suppose $\dim V \geq 2$ and $T \in \mathcal{L}(V)$ is such that $\operatorname{null} T^{\dim V - 2} \neq \operatorname{null} T^{\dim V - 1}$. Prove that T has at most two distinct eigenvalues.

5 Suppose $T \in \mathcal{L}(V)$ and 3 and 8 are eigenvalues of T. Let $n = \dim V$. Prove that $V = (\operatorname{null} T^{n-2}) \oplus (\operatorname{range} T^{n-2})$.

6 Suppose $T \in \mathcal{L}(V)$ and λ is an eigenvalue of T. Explain why the exponent of $z - \lambda$ in the factorization of the minimal polynomial of T is the smallest positive integer m such that $(T - \lambda I)^m|_{G(\lambda, T)} = 0$.

7 Suppose $T \in \mathcal{L}(V)$ and λ is an eigenvalue of T with multiplicity d. Prove that $G(\lambda, T) = \operatorname{null}(T - \lambda I)^d$.

If $d < \dim V$, then this exercise improves 8.20.

8 Suppose $T \in \mathcal{L}(V)$ and $\lambda_1, \ldots, \lambda_m$ are the distinct eigenvalues of T. Prove that

$$V = G(\lambda_1, T) \oplus \cdots \oplus G(\lambda_m, T)$$

if and only if the minimal polynomial of T equals $(z - \lambda_1)^{k_1} \cdots (z - \lambda_m)^{k_m}$ for some positive integers $k_1, \ldots, k_m$.

The case $\mathbf{F} = \mathbf{C}$ follows immediately from 5.27(b) and the generalized eigenspace decomposition (8.22); thus this exercise is interesting only when $\mathbf{F} = \mathbf{R}$.

9 Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$. Prove that there exist $D, N \in \mathcal{L}(V)$ such that $T = D + N$, the operator D is diagonalizable, N is nilpotent, and $DN = ND$.

10 Suppose V is a complex inner product space, $e_1, \ldots, e_n$ is an orthonormal basis of V, and $T \in \mathcal{L}(V)$. Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of T, each included as many times as its multiplicity. Prove that

$$|\lambda_1|^2 + \cdots + |\lambda_n|^2 \leq \|Te_1\|^2 + \cdots + \|Te_n\|^2.$$

See the comment after Exercise 5 in Section 7A.

11 Give an example of an operator on $\mathbf{C}^4$ whose characteristic polynomial equals $(z - 7)^2(z - 8)^2$.

תרגילים חג

1 נגדיר $T \in \mathcal{L}(\mathbf{C}^2)$ על ידי $T(w,z) = (-z,w)$. מצאו את המרחבים העצמיים המוכללים השייכים לערכים העצמיים השונים של T.

2 נניח ש־ $T \in \mathcal{L}(V)$ הפיך. הוכיחו כי $G(\lambda, T) = G\left(\frac{1}{\lambda}, T^{-1}\right)$ לכל $\lambda \in \mathbf{F}$ עם $\lambda \neq 0$.

3 נניח ש־ $T \in \mathcal{L}(V)$. נניח ש־ $S \in \mathcal{L}(V)$ הפיך. הוכיחו כי ל־ T ול־ $S^{-1}TS$ יש אותם ערכים עצמיים עם אותם ריבויים.

4 נניח ש־ $\dim V \geq 2$ ו־ $T \in \mathcal{L}(V)$ כך ש־ $\operatorname{null} T^{\dim V - 2} \neq \operatorname{null} T^{\dim V - 1}$. הוכיחו כי יש ל־ T שני ערכים עצמיים שונים לכל היותר.

5 נניח ש־ $T \in \mathcal{L}(V)$ ו־ 3 ו־ 8 הם ערכים עצמיים של T. יהי $n = \dim V$. הוכיחו כי $V = (\operatorname{null} T^{n-2}) \oplus (\operatorname{range} T^{n-2})$.

6 נניח ש־ $T \in \mathcal{L}(V)$ ו־ λ ערך עצמי של T. הסבירו מדוע המעריך של $z - \lambda$ בפירוק לגורמים של הפולינום המינימלי של T הוא המספר השלם החיובי m הקטן ביותר כך ש־ $(T - \lambda I)^m|_{G(\lambda,T)} = 0$.

7 נניח ש־ $T \in \mathcal{L}(V)$ ו־ λ ערך עצמי של T עם ריבוי d. הוכיחו כי $G(\lambda, T) = \operatorname{null}(T - \lambda I)^d$.

אם $d < \dim V$, אזי תרגיל זה משפר את 8.20.

8 נניח ש־ $T \in \mathcal{L}(V)$ ו־ $\lambda_1, \ldots, \lambda_m$ הערכים העצמיים השונים של T. הוכיחו כי

$$V = G(\lambda_1, T) \oplus \cdots \oplus G(\lambda_m, T)$$

אם־ורק־אם הפולינום המינימלי של T שווה ל־ $(z - \lambda_1)^{k_1} \cdots (z - \lambda_m)^{k_m}$ עבור מספרים שלמים חיוביים כלשהם $k_1, \ldots, k_m$.

המקרה של $\mathbf{F} = \mathbf{C}$ נובע ישירות מ־ 5.27(ב) והפירוק למרחבים עצמיים מוכללים (8.22); לפיכך תרגיל זה מעניין רק כאשר $\mathbf{F} = \mathbf{R}$.

9 נניח ש־ $\mathbf{F} = \mathbf{C}$ ו־ $T \in \mathcal{L}(V)$. הוכיחו כי קיימים $D, N \in \mathcal{L}(V)$ כך ש־ $T = D + N$, האופרטור D לכסין, N נילפוטנטי, ו־ $DN = ND$.

10 נניח ש־ V מרחב מכפלה פנימית מרוכב, $e_1, \ldots, e_n$ בסיס אורתונורמלי של V, ו־ $T \in \mathcal{L}(V)$. יהיו $\lambda_1, \ldots, \lambda_n$ הערכים העצמיים של T, כאשר כל אחד נכלל מספר פעמים לפי ריבויו. הוכיחו כי

$$|\lambda_1|^2 + \cdots + |\lambda_n|^2 \leq \|Te_1\|^2 + \cdots + \|Te_n\|^2.$$

ראו ההערה המופיעה אחרי תרגיל 5 ביחידה זו.

11 הביאו דוגמה של אופרטור מעל $\mathbf{C}^4$ שהפולינום האופייני שלו שווה $(z-7)^2(z-8)^2$.

12 Give an example of an operator on $\mathbf{C}^4$ whose characteristic polynomial equals $(z-1)(z-5)^3$ and whose minimal polynomial equals $(z-1)(z-5)^2$.

13 Give an example of an operator on $\mathbf{C}^4$ whose characteristic and minimal polynomials both equal $z(z-1)^2(z-3)$.

14 Give an example of an operator on $\mathbf{C}^4$ whose characteristic polynomial equals $z(z-1)^2(z-3)$ and whose minimal polynomial equals $z(z-1)(z-3)$.

15 Let T be the operator on $\mathbf{C}^4$ defined by $T(z_1, z_2, z_3, z_4) = (0, z_1, z_2, z_3)$. Find the characteristic polynomial and the minimal polynomial of T.

16 Let T be the operator on $\mathbf{C}^6$ defined by

$$T(z_1, z_2, z_3, z_4, z_5, z_6) = (0, z_1, z_2, 0, z_4, 0).$$

Find the characteristic polynomial and the minimal polynomial of T.

17 Suppose $\mathbf{F} = \mathbf{C}$ and $P \in \mathcal{L}(V)$ is such that $P^2 = P$. Prove that the characteristic polynomial of P is $z^m(z-1)^n$, where $m = \dim \operatorname{null} P$ and $n = \dim \operatorname{range} P$.

18 Suppose $T \in \mathcal{L}(V)$ and λ is an eigenvalue of T. Explain why the following four numbers equal each other.

(a) The exponent of $z - \lambda$ in the factorization of the minimal polynomial of T.

(b) The smallest positive integer m such that $(T - \lambda I)^m|_{G(\lambda, T)} = 0$.

(c) The smallest positive integer m such that

$$\operatorname{null}(T - \lambda I)^m = \operatorname{null}(T - \lambda I)^{m+1}.$$

(d) The smallest positive integer m such that

$$\operatorname{range}(T - \lambda I)^m = \operatorname{range}(T - \lambda I)^{m+1}.$$

19 Suppose $\mathbf{F} = \mathbf{C}$ and $S \in \mathcal{L}(V)$ is a unitary operator. Prove that the constant term in the characteristic polynomial of S has absolute value 1.

20 Suppose that $\mathbf{F} = \mathbf{C}$ and $V_1, \ldots, V_m$ are nonzero subspaces of V such that

$$V = V_1 \oplus \cdots \oplus V_m.$$

Suppose $T \in \mathcal{L}(V)$ and each V_k is invariant under T. For each k, let p_k denote the characteristic polynomial of $T|_{V_k}$. Prove that the characteristic polynomial of T equals $p_1 \cdots p_m$.

21 Suppose $p, q \in \mathcal{P}(\mathbf{C})$ are monic polynomials with the same zeros and q is a polynomial multiple of p. Prove that there exists $T \in \mathcal{L}(\mathbf{C}^{\deg q})$ such that the characteristic polynomial of T is q and the minimal polynomial of T is p.

This exercise implies that every monic polynomial is the characteristic polynomial of some operator.

12 הביאו דוגמה של אופרטור מעל $\mathbf{C}^4$ שהפולינום האופייני שלו שווה $(z-1)(z-5)^3$ והפולינום המינימלי שלו שווה $(z-1)(z-5)^2$.

13 הביאו דוגמה של אופרטור מעל $\mathbf{C}^4$ שהפולינומים האופייני והמינימלי שלו שניהם שווים $z(z-1)^2(z-3)$.

14 הביאו דוגמה של אופרטור מעל $\mathbf{C}^4$ שהפולינום האופייני שלו שווה $z(z-1)^2(z-3)$ והפולינום המינימלי שלו שווה $z(z-1)(z-3)$.

15 יהי T האופרטור מעל $\mathbf{C}^4$ המוגדר על ידי $T(z_1,z_2,z_3,z_4) = (0,z_1,z_2,z_3)$. מצאו את הפולינום האופייני והפולינום המינימלי של T.

16 יהי T האופרטור מעל $\mathbf{C}^6$ המוגדר על ידי

$$T(z_1,z_2,z_3,z_4,z_5,z_6) = (0,z_1,z_2,0,z_4,0).$$

מצאו את הפולינום האופייני והפולינום המינימלי של T.

17 נניח ש־ $\mathbf{F} = \mathbf{C}$ ו־ $P \in \mathcal{L}(V)$ כך ש־ $P^2 = P$. הוכיחו כי הפולינום האופייני של P הוא $z^m(z-1)^n$, כאשר $m = \dim \operatorname{null} P$ ו־ $n = \dim \operatorname{range} P$.

18 נניח ש־ $T \in \mathcal{L}(V)$ ו־ λ ערך עצמי של T. הסבירו מדוע ארבעת המספרים הבאים שווים.

א. המעריך של $z-\lambda$ בפירוק לגורמים של הפולינום המינימלי של T.

ב. המספר השלם החיובי m הקטן ביותר כך ש־ $(T-\lambda I)^m|_{G(\lambda,T)} = 0$.

ג. המספר השלם החיובי m הקטן ביותר כך ש־

$$\operatorname{null}(T-\lambda I)^m = \operatorname{null}(T-\lambda I)^{m+1}.$$

ד. המספר השלם החיובי m הקטן ביותר כך ש־

$$\operatorname{range}(T-\lambda I)^m = \operatorname{range}(T-\lambda I)^{m+1}.$$

19 נניח ש־ $\mathbf{F} = \mathbf{C}$ ו־ $S \in \mathcal{L}(V)$ אופרטור אוניטרי. הוכיחו כי הערך המוחלט של האיבר הקבוע בפולינום האופייני של S הוא 1.

20 נניח ש־ $\mathbf{F} = \mathbf{C}$ ו־ $V_1, \dots, V_m$ תתי־מרחב השונים מאפס של V כך ש־

$$V = V_1 \oplus \cdots \oplus V_m.$$

נניח ש־ $T \in \mathcal{L}(V)$ וכל V_k שמור תחת T. לכל k, נסמן ב־ p_k את הפולינום האופייני של $T|_{V_k}$. הוכיחו כי הפולינום האופייני של T שווה ל־ $p_1 \cdots p_m$.

21 נניח ש־ $p, q \in \mathcal{P}(\mathbf{C})$ פולינומים מתוקנים עם אותם שורשים ו־ q כפולת פולינום של p. הוכיחו כי קיים $T \in \mathcal{L}(\mathbf{C}^{\deg q})$ כך שהפולינום האופייני של T הוא q והפולינום המינימלי של T הוא p.

מתרגיל זה נובע שכל פולינום מתוקן הוא פולינום אופייני של אופרטור כלשהו.

אלגברה ליניארית בדרך הישר, מהדורה רביעית, מחבר: שלדון אקסלר
Linear Algebra Done Right, fourth edition, by Sheldon Axler

22 Suppose A and B are block diagonal matrices of the form

$$A = \begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_m \end{pmatrix}, \quad B = \begin{pmatrix} B_1 & & 0 \\ & \ddots & \\ 0 & & B_m \end{pmatrix},$$

where A_k and B_k are square matrices of the same size for each $k = 1, \dots, m$. Show that AB is a block diagonal matrix of the form

$$AB = \begin{pmatrix} A_1B_1 & & 0 \\ & \ddots & \\ 0 & & A_mB_m \end{pmatrix}.$$

23 Suppose $\mathbf{F} = \mathbf{R}$, $T \in \mathcal{L}(V)$, and $\lambda \in \mathbf{C}$.

(a) Show that $u + iv \in G(\lambda, T_{\mathbf{C}})$ if and only if $u - iv \in G(\overline{\lambda}, T_{\mathbf{C}})$.

(b) Show that the multiplicity of λ as an eigenvalue of $T_{\mathbf{C}}$ equals the multiplicity of $\overline{\lambda}$ as an eigenvalue of $T_{\mathbf{C}}$.

(c) Use (b) and the result about the sum of the multiplicities (8.25) to show that if $\dim V$ is an odd number, then $T_{\mathbf{C}}$ has a real eigenvalue.

(d) Use (c) and the result about real eigenvalues of $T_{\mathbf{C}}$ (Exercise 17 in Section 5A) to show that if $\dim V$ is an odd number, then T has an eigenvalue (thus giving an alternative proof of 5.34).

See Exercise 33 in Section 3B for the definition of the complexification $T_{\mathbf{C}}$.

22 נניח ש־ A ו־ B מטריצות בלוקים אלכסוניות מהמבנה

$$A = \begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_m \end{pmatrix}, \quad B = \begin{pmatrix} B_1 & & 0 \\ & \ddots & \\ 0 & & B_m \end{pmatrix},$$

כאשר A_k ו־ B_k מטריצות ריבועיות מאותו הגודל לכל $k = 1, \dots, m$. הראו כי AB היא מטריצת בלוקים אלכסונית במבנה

$$AB = \begin{pmatrix} A_1B_1 & & 0 \\ & \ddots & \\ 0 & & A_mB_m \end{pmatrix}.$$

23 נניח ש־ $T \in \mathcal{L}(V)$, ו־ $\mathbf{F} = \mathbf{R}$, $\lambda \in \mathbf{C}$.

א. הראו כי $u + iv \in G(\lambda, T_{\mathbf{C}})$ אם־ורק־אם $u - iv \in G(\overline{\lambda}, T_{\mathbf{C}})$.

ב. הראו כי הריבוי של λ כערך עצמי של $T_{\mathbf{C}}$ שווה לריבוי של $\overline{\lambda}$ כערך עצמי של $T_{\mathbf{C}}$.

ג. השתמשו ב־ (ב) ובמסקנה לגבי הסכום של הריבויים (8.25) כדי להראות כי אם $\dim V$ מספר אי־זוגי, אזי יש ל־ $T_{\mathbf{C}}$ ערך עצמי ממשי.

ד. השתמשו ב־ (ג) ובמסקנה לגבי ערכים עצמיים ממשיים של $T_{\mathbf{C}}$ (תרגיל 17 ביחידה ה1) כדי להראות כי אם $\dim V$ מספר אי־זוגי, אזי יש ל־ T ערך עצמי (לפיכך זו תהיה הוכחה חלופית של 5.34).

ראו תרגיל 33 ביחידה ג2 עבור ההגדרה של המירכוב $T_{\mathbf{C}}$.

8C Consequences of Generalized Eigenspace Decomposition

Square Roots of Operators

Recall that a square root of an operator $T \in \mathcal{L}(V)$ is an operator $R \in \mathcal{L}(V)$ such that $R^2 = T$ (see 7.36). Every complex number has a square root, but not every operator on a complex vector space has a square root. For example, the operator on $\mathbf{C}^3$ defined by $T(z_1, z_2, z_3) = (z_2, z_3, 0)$ does not have a square root, as you are asked to show in Exercise 1. The noninvertibility of that operator is no accident, as we will soon see. We begin by showing that the identity plus any nilpotent operator has a square root.

8.39 *identity plus nilpotent has a square root*

Suppose $T \in \mathcal{L}(V)$ is nilpotent. Then $I + T$ has a square root.

Proof Consider the Taylor series for the function $\sqrt{1+x}$:

$$\sqrt{1+x} = 1 + a_1 x + a_2 x^2 + \cdots . \tag{8.40}$$

We do not find an explicit formula for the coefficients or worry about whether the infinite sum converges because we use this equation only as motivation.

Because $a_1 = \frac{1}{2}$, the formula above implies that $1 + \frac{x}{2}$ is a good estimate for $\sqrt{1+x}$ when x is small.

Because T is nilpotent, $T^m = 0$ for some positive integer m. In 8.40, suppose we replace x with T and 1 with I. Then the infinite sum on the right side becomes a finite sum (because $T^k = 0$ for all $k \geq m$). Thus we guess that there is a square root of $I + T$ of the form

$$I + a_1 T + a_2 T^2 + \cdots + a_{m-1} T^{m-1}.$$

Having made this guess, we can try to choose $a_1, a_2, \ldots, a_{m-1}$ such that the operator above has its square equal to $I + T$. Now

$$\begin{aligned}(I + a_1 T + a_2 T^2 + a_3 T^3 + \cdots + a_{m-1} T^{m-1})^2 \\ = I + 2a_1 T + (2a_2 + a_1^2) T^2 + (2a_3 + 2a_1 a_2) T^3 + \cdots \\ + (2a_{m-1} + \text{terms involving } a_1, \ldots, a_{m-2}) T^{m-1}.\end{aligned}$$

We want the right side of the equation above to equal $I + T$. Hence choose a_1 such that $2a_1 = 1$ (thus $a_1 = 1/2$). Next, choose a_2 such that $2a_2 + a_1^2 = 0$ (thus $a_2 = -1/8$). Then choose a_3 such that the coefficient of T^3 on the right side of the equation above equals 0 (thus $a_3 = 1/16$). Continue in this fashion for each $k = 4, \ldots, m-1$, at each step solving for a_k so that the coefficient of T^k on the right side of the equation above equals 0. Actually we do not care about the explicit formula for the a_k's. We only need to know that some choice of the a_k's gives a square root of $I + T$. ∎

3ח מסקנות מהפירוק למרחבים עצמיים מוכללים

שורשים ריבועיים של אופרטורים

נזכיר כי שורש ריבועי של אופרטור $T \in \mathcal{L}(V)$ הוא אופרטור $R \in \mathcal{L}(V)$ כך ש־ $R^2 = T$ (ראו 7.36). לכל מספר מרוכב יש שורש ריבועי, אך לא לכל אופרטור מעל מרחב וקטורי מרוכב יש שורש ריבועי. לדוגמה, לאופרטור מעל $\mathbf{C}^3$ המוגדר על ידי $T(z_1, z_2, z_3) = (z_2, z_3, 0)$ אין שורש ריבועי, כפי שנדרשתם להראות בתרגיל 1. האי־הפיכות של אופרטור זה אינה מקרית, כפי שנראה בקרוב. אנו פותחים בכך שנראה כי לאופרטור הזהות בתוספת כל אופרטור נילפוטנטי יש שורש ריבועי.

8.39 לַזֶהוּת בתוספת נילפוטנטי יש שורש ריבועי

נניח ש־ $T \in \mathcal{L}(V)$ נילפוטנטי. אזי ל־ $I + T$ יש שורש ריבועי.

הוכחה נתבונן בטור טיילור עבור הפונקציה $\sqrt{1+x}$:

$$\sqrt{1+x} = 1 + a_1 x + a_2 x^2 + \cdots. \tag{8.40}$$

אינני מוצאים נוסחה מפורשת למקדמים או מוטרדים מהשאלה האם הסכום האינסופי מתכנס מאחר שאנו משתמשים במשוואה זו רק כמוטיבציה.

מאחר ש־ $a_1 = \frac{1}{2}$, הנוסחה לעיל גוררת ש־ $1 + \frac{x}{2}$ הוא הערכה טובה עבור $\sqrt{1+x}$ כאשר x קטן.

מאחר ש־ T נילפוטנטי, $T^m = 0$ עבור m מספר שלם חיובי כלשהו. ב־ 8.40, נניח שאנו מחליפים את x ב־ T ואת 1 ב־ I. אזי הסכום האינסופי באגף הימני הופך לסכום סופי (מאחר ש־ $T^k = 0$ לכל $k \geq m$). לפיכך אנו מנחשים שקיים שורש ריבועי של $I + T$ מהצורה

$$I + a_1 T + a_2 T^2 + \cdots + a_{m-1} T^{m-1}.$$

אחרי שניחשנו זאת, נוכל לנסות לבחור $a_1, a_2, \ldots, a_{m-1}$ כך שריבוע האופרטור לעיל שווה ל־ $I + T$. כעת

$$\begin{aligned}(I + a_1 T + a_2 T^2 + a_3 T^3 + \cdots + a_{m-1} T^{m-1})^2 \\ = I + 2a_1 T + (2a_2 + a_1^2) T^2 + (2a_3 + 2a_1 a_2) T^3 + \cdots \\ + (2a_{m-1} + a_1, \ldots, a_{m-2} \text{ איברים הכוללים}) T^{m-1}.\end{aligned}$$

אנו רוצים שהאגף הימני של המשוואה לעיל יהיה שווה ל־ $I + T$. לכן נבחר a_1 כך ש־ $2a_1 = 1$ (לפיכך $a_1 = 1/2$). בשלב הבא, נבחר a_2 כך ש־ $2a_2 + a_1^2 = 0$ (לפיכך $a_2 = -1/8$). לאחר מכן נבחר a_3 כך שהמקדם של T^3 באגף הימני של המשוואה לעיל שווה 0 (לפיכך $a_3 = 1/16$). נמשיך בדרך זו עבור כל $k = 4, \ldots, m-1$, בכל שלב נפתור עבור a_k כך שהמקדם של T^k באגף הימני של המשוואה לעיל שווה 0. למעשה, אין לנו עניין בנוסחה מפורשת עבור ה־ a_k־ים. כל מה שאנו מבקשים לדעת הוא כי בחירה כלשהי של ה־ a_k־ים מספקת שורש ריבועי של $I + T$. ∎

The previous lemma is valid on real and complex vector spaces. However, the result below holds only on complex vector spaces. For example, the operator of multiplication by -1 on the one-dimensional real vector space $\mathbf{R}$ has no square root.

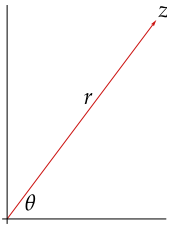
For the proof below, we need to know that every $z \in \mathbf{C}$ has a square root in $\mathbf{C}$. To show this, write

$$z = r(\cos\theta + i\sin\theta),$$

where r is the length of the line segment in the complex plane from the origin to z and θ is the angle of that line segment with the positive horizontal axis. Then

$$\sqrt{r}\left(\cos\tfrac{\theta}{2} + i\sin\tfrac{\theta}{2}\right)$$

is a square root of z, as you can verify by showing that the square of the complex number above equals z.



Representation of a complex number with polar coordinates.

8.41 *over* $\mathbf{C}$*, invertible operators have square roots*

Suppose V is a complex vector space and $T \in \mathcal{L}(V)$ is invertible. Then T has a square root.

Proof Let $\lambda_1, \ldots, \lambda_m$ be the distinct eigenvalues of T. For each k, there exists a nilpotent operator $T_k \in \mathcal{L}(G(\lambda_k, T))$ such that $T|_{G(\lambda_k,T)} = \lambda_k I + T_k$ [see 8.22(b)]. Because T is invertible, none of the λ_k's equals 0, so we can write

$$T|_{G(\lambda_k,T)} = \lambda_k\left(I + \frac{T_k}{\lambda_k}\right)$$

for each k. Because T_k/λ_k is nilpotent, $I + T_k/\lambda_k$ has a square root (by 8.39). Multiplying a square root of the complex number λ_k by a square root of $I + T_k/\lambda_k$, we obtain a square root R_k of $T|_{G(\lambda_k,T)}$.

By the generalized eigenspace decomposition (8.22), a typical vector $v \in V$ can be written uniquely in the form

$$v = u_1 + \cdots + u_m,$$

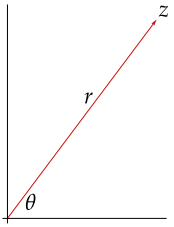
where each u_k is in $G(\lambda_k, T)$. Using this decomposition, define an operator $R \in \mathcal{L}(V)$ by

$$Rv = R_1u_1 + \cdots + R_mu_m.$$

You should verify that this operator R is a square root of T, completing the proof. ∎

By imitating the techniques in this subsection, you should be able to prove that if V is a complex vector space and $T \in \mathcal{L}(V)$ is invertible, then T has a kth root for every positive integer k.

הלמה הקודמת מתקיימת מעל מרחבים וקטוריים ממשיים ומרוכבים. עם זאת, המסקנה להלן מתקיימת רק מעל מרחבים וקטוריים מרוכבים. לדוגמה, לאופרטור של כפל ב־ -1 מעל המרחב הווקטורי הממשי החד־ממדי $\mathbf{R}$ אין שורש ריבועי.



הצגה קוטבית של מספר מרוכב.

עבור ההוכחה להלן, עלינו לדעת כי לכל $z \in \mathbf{C}$ יש שורש ריבועי ב־ $\mathbf{C}$. כדי להראות זאת, נכתוב

$$z = r(\cos\theta + i\sin\theta),$$

כאשר r הוא אורכו של הקטע הישר במישור המרוכב מראשית הצירים אל z ו־ θ היא הזווית של אותו קטע ישר עם הציר האופקי החיובי. אזי

$$\sqrt{r}\left(\cos\tfrac{\theta}{2} + i\sin\tfrac{\theta}{2}\right)$$

הוא שורש ריבועי של z, כפי שתוכלו לוודא בכך שתראו כי הריבוע של מספר מרוכב זה שווה ל־ z.

8.41 מעל $\mathbf{C}$, לאופרטורים הפיכים יש שורשים ריבועיים

נניח ש־ V מרחב וקטורי מרוכב ו־ $T \in \mathcal{L}(V)$ הפיך. אזי יש ל־ T שורש ריבועי.

הוכחה יהיו $\lambda_1, \dots, \lambda_m$ הערכים העצמיים השונים של T. לכל k, קיים אופרטור נילפוטנטי $T_k \in \mathcal{L}(G(\lambda_k, T))$ כך ש־ $T|_{G(\lambda_k,T)} = \lambda_k I + T_k$ [ראו 8.22(ב)]. מאחר ש־ T הפיך, כל ה־ λ_k־ים שונים מ־ 0, לכן נוכל לכתוב

$$T|_{G(\lambda_k,T)} = \lambda_k\left(I + \frac{T_k}{\lambda_k}\right)$$

לכל k. מאחר ש־ T_k/λ_k נילפוטנטי, יש ל־ $I + T_k/\lambda_k$ שורש ריבועי (לפי 8.39). כשנכפיל שורש ריבועי של המספר המרוכב λ_k בשורש ריבועי של $I + T_k/\lambda_k$, נקבל שורש ריבועי R_k של $T|_{G(\lambda_k,T)}$.

לפי הפירוק למרחבים עצמיים מוכללים (8.22), וקטור טיפוסי $v \in V$ ניתן לכתיבה באופן יחיד בצורה

$$v = u_1 + \cdots + u_m,$$

כאשר כל u_k הוא ב־ $G(\lambda_k, T)$. בשימוש בפירוק זה, נגדיר אופרטור $R \in \mathcal{L}(V)$ על ידי

$$Rv = R_1u_1 + \cdots + R_mu_m.$$

עליכם לוודא כי אופרטור R זה הוא שורש ריבועי של T, ובכך הושלמה ההוכחה. ■

באמצעות חיקוי הטכניקות בתת־יחידה זו של הפרק, ביכולתכם להוכיח כי אם V מרחב וקטורי מרוכב ו־ $T \in \mathcal{L}(V)$ הפיך, אזי יש ל־ T שורש k^{th} לכל k מספר שלם חיובי.

Jordan Form

We know that if V is a complex vector space, then for every $T \in \mathcal{L}(V)$ there is a basis of V with respect to which T has a nice upper-triangular matrix (see 8.37). In this subsection we will see that we can do even better—there is a basis of V with respect to which the matrix of T contains 0's everywhere except possibly on the diagonal and the line directly above the diagonal.

We begin by looking at two examples of nilpotent operators.

8.42 example: *nilpotent operator with nice matrix*

Let T be the operator on $\mathbf{C}^4$ defined by

$$T(z_1, z_2, z_3, z_4) = (0, z_1, z_2, z_3).$$

Then $T^4 = 0$; thus T is nilpotent. If $v = (1, 0, 0, 0)$, then T^3v, T^2v, Tv, v is a basis of $\mathbf{C}^4$. The matrix of T with respect to this basis is

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

The next example of a nilpotent operator has more complicated behavior than the example above.

8.43 example: *nilpotent operator with slightly more complicated matrix*

Let T be the operator on $\mathbf{C}^6$ defined by

$$T(z_1, z_2, z_3, z_4, z_5, z_6) = (0, z_1, z_2, 0, z_4, 0).$$

Then $T^3 = 0$; thus T is nilpotent. In contrast to the nice behavior of the nilpotent operator of the previous example, for this nilpotent operator there does not exist a vector $v \in \mathbf{C}^6$ such that $T^5v, T^4v, T^3v, T^2v, Tv, v$ is a basis of $\mathbf{C}^6$. However, if we take $v_1 = (1, 0, 0, 0, 0, 0)$, $v_2 = (0, 0, 0, 1, 0, 0)$, and $v_3 = (0, 0, 0, 0, 0, 1)$, then $T^2v_1, Tv_1, v_1, Tv_2, v_2, v_3$ is a basis of $\mathbf{C}^6$. The matrix of T with respect to this basis is

$$\begin{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} & \begin{matrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{matrix} & \begin{matrix} 0 \\ 0 \\ 0 \end{matrix} \\ \begin{matrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{matrix} & \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} & \begin{matrix} 0 \\ 0 \end{matrix} \\ \begin{matrix} 0 & 0 & 0 \end{matrix} & \begin{matrix} 0 & 0 \end{matrix} & \begin{pmatrix} 0 \end{pmatrix} \end{pmatrix}.$$

Here the inner matrices are blocked off to show that we can think of the 6-by-6 matrix above as a block diagonal matrix consisting of a 3-by-3 block with 1's on the line above the diagonal and 0's elsewhere, a 2-by-2 block with 1 above the diagonal and 0's elsewhere, and a 1-by-1 block containing 0.

צורת ז'ורדן

אנו יודעים כי אם V הוא מרחב וקטורי מרוכב, אזי לכל $T \in \mathcal{L}(V)$ קיים בסיס של V אשר ביחס אליו יש ל־ T מטריצה משולשית עליונה נחמדה (ראו 8.37). בתת־יחידה זו של הפרק נראה שאנו יכולים לעשות אף טוב יותר—קיים בסיס של V אשר ביחס אליו המטריצה של T מכילה אפסים בכל מקום מלבד אולי על האלכסון הראשי ולאורך הקו שמעליו. נפתח באמצעות התבוננות בשתי דוגמאות של אופרטורים נילפוטנטיים.

8.42 דוגמה: אופרטור נילפוטנטי עם מטריצה נחמדה

יהי T האופרטור מעל $\mathbf{C}^4$ המוגדר על ידי

$$T(z_1, z_2, z_3, z_4) = (0, z_1, z_2, z_3).$$

אזי $T^4 = 0$; לפיכך T נילפוטנטי. אם $v = (1, 0, 0, 0)$, אזי $T^3 v, T^2 v, Tv, v$ בסיס של $\mathbf{C}^4$. המטריצה של T ביחס לבסיס זה היא

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

התנהגות האופרטור הנילפוטנטי בדוגמה הבאה מורכבת יותר מזו שבדוגמה לעיל.

8.43 דוגמה: אופרטור נילפוטנטי עם מטריצה מעט יותר מורכבת

יהי T האופרטור מעל $\mathbf{C}^6$ המוגדר על ידי

$$T(z_1, z_2, z_3, z_4, z_5, z_6) = (0, z_1, z_2, 0, z_4, 0).$$

אזי $T^3 = 0$; לפיכך T נילפוטנטי. בניגוד להתנהגותו הנחמדה של האופרטור הנילפוטנטי מהדוגמה הקודמת, לא קיים עבור אופרטור נילפוטנטי זה וקטור $v \in \mathbf{C}^6$ כך ש־ $T^5 v, T^4 v, T^3 v, T^2 v, Tv, v$ בסיס של $\mathbf{C}^6$. עם זאת, אם ניקח $v_1 = (1, 0, 0, 0, 0, 0)$, $v_2 = (0, 0, 0, 1, 0, 0)$ ו־ $v_3 = (0, 0, 0, 0, 0, 1)$, אזי $T^2 v_1, Tv_1, v_1, Tv_2, v_2, v_3$ בסיס של $\mathbf{C}^6$. המטריצה של T ביחס לבסיס זה היא

$$\begin{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} & \begin{matrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{matrix} & \begin{matrix} 0 \\ 0 \\ 0 \end{matrix} \\ \begin{matrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{matrix} & \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} & \begin{matrix} 0 \\ 0 \\ 0 \end{matrix} \\ \begin{matrix} 0 & 0 & 0 \end{matrix} & \begin{matrix} 0 & 0 \end{matrix} & \begin{pmatrix} 0 \end{pmatrix} \end{pmatrix}.$$

המטריצות הפנימיות כאן סומנו כבלוקים כדי להראות כי נוכל לחשוב על המטריצה 6×6 לעיל כמטריצת בלוקים אלכסונית המורכבת מבלוק 3×3 עם 1־ים על הקו מעל האלכסון ואפסים בשאר האיברים, בלוק 2×2 עם 1 מעל האלכסון ואפסים בשאר האיברים, ובלוק 1×1 המכיל 0.

אלגברה ליניארית בדרך הישר, מהדורה רביעית, מחבר: שלדון אקסלר
Linear Algebra Done Right, fourth edition, by Sheldon Axler

Our next goal is to show that every nilpotent operator $T \in \mathcal{L}(V)$ behaves similarly to the operator in the previous example. Specifically, there is a finite collection of vectors $v_1, \dots, v_n \in V$ such that there is a basis of V consisting of the vectors of the form $T^j v_k$, as k varies from 1 to n and j varies (in reverse order) from 0 to the largest nonnegative integer m_k such that $T^{m_k} v_k \neq 0$. With respect to this basis, the matrix of T looks like the matrix in the previous example. More specifically, T has a block diagonal matrix with respect to this basis, with each block a square matrix that is 0 everywhere except on the line above the diagonal.

In the next definition, the diagonal of each A_k is filled with some eigenvalue λ_k of T, the line directly above the diagonal of A_k is filled with 1's, and all other entries in A_k are 0 (to understand why each λ_k is an eigenvalue of T, see 5.41). The λ_k's need not be distinct. Also, A_k may be a 1-by-1 matrix (λ_k) containing just an eigenvalue of T. If each λ_k is 0, then the next definition captures the behavior described in the paragraph above (recall that if T is nilpotent, then 0 is the only eigenvalue of T).

8.44 definition: *Jordan basis*

Suppose $T \in \mathcal{L}(V)$. A basis of V is called a *Jordan basis* for T if with respect to this basis T has a block diagonal matrix

$$\begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_p \end{pmatrix}$$

in which each A_k is an upper-triangular matrix of the form

$$A_k = \begin{pmatrix} \lambda_k & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_k \end{pmatrix}.$$

Most of the work in proving that every operator on a finite-dimensional complex vector space has a Jordan basis occurs in proving the special case below of nilpotent operators. This special case holds on real vector spaces as well as complex vector spaces.

8.45 *every nilpotent operator has a Jordan basis*

Suppose $T \in \mathcal{L}(V)$ is nilpotent. Then there is a basis of V that is a Jordan basis for T.

Proof We will prove this result by induction on $\dim V$. To get started, note that the desired result holds if $\dim V = 1$ (because in that case, the only nilpotent operator is the 0 operator). Now assume that $\dim V > 1$ and that the desired result holds on all vector spaces of smaller dimension.

מטרתנו הבאה היא להראות כי כל אופרטור נילפוטנטי $T \in \mathcal{L}(V)$ מתנהג באופן דומה לאופרטור בדוגמה לעיל. ליתר דיוק, קיים אוסף סופי של וקטורים $v_1, \dots, v_n \in V$ כך שקיים בסיס של V המורכב מהווקטורים מהצורה $T^j v_k$, כאשר k משתנה מ־ 1 ל־ n ו־ j משתנה (בסדר הפוך) מ־ 0 למספר השלם האי־שלילי m_k הגדול ביותר כך ש־ $T^{m_k} v_k \neq 0$. ביחס לבסיס זה, המטריצה של T נראית כמו המטריצה בדוגמה לעיל. ביתר פירוט, יש ל־ T מטריצת בלוקים אלכסונית ביחס לבסיס זה, כאשר כל בלוק הוא מטריצה ריבועית עם 0 בכל מקום מלבד על הקו מעל האלכסון.

בהגדרה הבאה, באלכסון של כל A_k משובץ ערך עצמי λ_k של T, הקו שמיד מעל האלכסון של A_k מלא ב־ 1־ים, וכל שאר האיברים של A_k הם 0 (להבין מדוע כל λ_k הוא ערך עצמי של T, ראו 5.41). ה־ λ_k־ים לא חייבים להיות שונים. בנוסף, ייתכן ש־ A_k היא מטריצה 1×1 המכילה רק (λ_k) שהוא ערך עצמי של T. אם כל λ_k הוא 0, אזי ההגדרה הבאה לוכדת את ההתנהגות שתוארה בפסקה לעיל (נזכיר כי אם T נילפוטנטי, אזי 0 הוא הערך העצמי היחיד של T).

8.44 הגדרה: בסיס ז'ורדן

נניח ש־ $T \in \mathcal{L}(V)$. בסיס של V נקרא בסיס ז'ורדן עבור T אם ביחס לבסיס זה יש ל־ T מטריצת בלוקים אלכסונית

$$\begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_p \end{pmatrix}$$

כאשר כל A_k היא מטריצה משולשית עליונה מהמבנה

$$A_k = \begin{pmatrix} \lambda_k & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_k \end{pmatrix}.$$

רוב מלאכת ההוכחה לכך שלכל אופרטור מעל מרחב וקטורי מרוכב מממד סופי יש בסיס ז'ורדן מתבצעת במהלך הוכחת המקרה הפרטי להלן של אופרטורים נילפוטנטיים. מקרה פרטי זה מתקיים מעל מרחבים וקטוריים ממשיים ומרוכבים.

8.45 לכל אופרטור נילפוטנטי יש בסיס ז'ורדן

נניח ש־ $T \in \mathcal{L}(V)$ נילפוטנטי. אזי קיים בסיס של V שהוא בסיס ז'ורדן עבור T.

הוכחה אנו נוכיח מסקנה זו באמצעות אינדוקציה על $\dim V$. בתחילה, נשים לב שהתוצאה הרצויה מתקיימת אם $\dim V = 1$ (מאחר שבמקרה זה, האופרטור הנילפוטנטי היחידי הוא אופרטור האפס). כעת נניח ש־ $\dim V > 1$ וכי התוצאה הרצויה מתקיימת מעל כל המרחבים הווקטוריים מממד קטן יותר.

Let m be the smallest positive integer such that $T^m = 0$. Thus there exists $u \in V$ such that $T^{m-1}u \neq 0$. Let

$$U = \operatorname{span}(u, Tu, \dots, T^{m-1}u).$$

The list $u, Tu, \dots, T^{m-1}u$ is linearly independent (see Exercise 2 in Section 8A). If $U = V$, then writing this list in reverse order gives a Jordan basis for T and we are done. Thus we can assume that $U \neq V$.

Note that U is invariant under T. By our induction hypothesis, there is a basis of U that is a Jordan basis for $T|_U$. The strategy of our proof is that we will find a subspace W of V such that W is also invariant under T and $V = U \oplus W$. Again by our induction hypothesis, there will be a basis of W that is a Jordan basis for $T|_W$. Putting together the Jordan bases for $T|_U$ and $T|_W$, we will have a Jordan basis for T.

Let $\varphi \in V'$ be such that $\varphi(T^{m-1}u) \neq 0$. Let

$$W = \big\{v \in V : \varphi(T^k v) = 0 \text{ for each } k = 0, \dots, m-1\big\}.$$

Then W is a subspace of V that is invariant under T $\big($the invariance holds because if $v \in W$ then $\varphi(T^k(Tv)) = 0$ for $k = 0, \dots, m-1$, where the case $k = m-1$ holds because $T^m = 0\big)$. We will show that $V = U \oplus W$, which by the previous paragraph will complete the proof.

To show that $U + W$ is a direct sum, suppose $v \in U \cap W$ with $v \neq 0$. Because $v \in U$, there exist $c_0, \dots, c_{m-1} \in \mathbf{F}$ such that

$$v = c_0 u + c_1 Tu + \cdots + c_{m-1}T^{m-1}u.$$

Let j be the smallest index such that $c_j \neq 0$. Apply T^{m-j-1} to both sides of the equation above, getting

$$T^{m-j-1}v = c_j T^{m-1}u,$$

where we have used the equation $T^m = 0$. Now apply φ to both sides of the equation above, getting

$$\varphi(T^{m-j-1}v) = c_j \varphi(T^{m-1}u) \neq 0.$$

The equation above shows that $v \notin W$. Hence we have proved that $U \cap W = \{0\}$, which implies that $U + W$ is a direct sum (see 1.46).

To show that $U \oplus W = V$, define $S\colon V \to \mathbf{F}^m$ by

$$Sv = \big(\varphi(v), \varphi(Tv), \dots, \varphi(T^{m-1}v)\big).$$

Thus $\operatorname{null} S = W$. Hence

$$\dim W = \dim \operatorname{null} S = \dim V - \dim \operatorname{range} S \geq \dim V - m,$$

where the second equality comes from the fundamental theorem of linear maps (3.21). Using the inequality above, we have

$$\dim(U \oplus W) = \dim U + \dim W \geq m + (\dim V - m) = \dim V.$$

Thus $U \oplus W = V$ (by 2.39), completing the proof. ∎

יהי m המספר השלם החיובי הקטן ביותר כך ש־ $T^m = 0$. לפיכך קיים $u \in V$ כך ש־ $T^{m-1}u \neq 0$. יהי

$$U = \operatorname{span}(u, Tu, \dots, T^{m-1}u).$$

הרשימה $u, Tu, \dots, T^{m-1}u$ בלתי תלויה ליניארית (ראו תרגיל 2 ביחידה 1ח). אם $U = V$, אזי כתיבת רשימה זו בסדר הפוך נותנת לנו בסיס ז׳ורדן עבור T וסיימנו. לפיכך אנו יכולים להניח ש־ $U \neq V$.

שימו לב ש־ U שמור תחת T. באמצעות הנחת האינדוקציה, קיים בסיס של U שהוא בסיס ז׳ורדן עבור $T|_U$. אסטרטגיית ההוכחה שלנו היא למצוא תת־מרחב W של V כך שגם W שמור תחת T ו־ $V = U \oplus W$. שוב, באמצעות הנחת האינדוקציה, יהיה בסיס של W שהוא בסיס ז׳ורדן עבור $T|_W$. צירופם יחד של בסיסי ז׳ורדן של $T|_U$ ושל $T|_W$ יעניק לנו בסיס ז׳ורדן עבור T.

יהי $\varphi \in V'$ כך ש־ $\varphi(T^{m-1}u) \neq 0$. יהי

$$W = \{v \in V : \varphi(T^k v) = 0 \text{ לכל } k = 0, \dots, m-1\}.$$

אזי W תת־מרחב של V השמור תחת T (השימור מתקיים מאחר שאם $v \in W$ אזי $\varphi(T^k(Tv)) = 0$ עבור $k = 0, \dots, m-1$, כאשר $k = m-1$ מתקיים מאחר ש־ $T^m = 0$). אנו נראה כי $V = U \oplus W$, מהלך שעל פי הפסקה לעיל ישלים את ההוכחה.

כדי להראות כי $U + W$ סכום ישר, נניח ש־ $v \in U \cap W$ עם $v \neq 0$. מאחר ש־ $v \in U$, קיימים $c_0, \dots, c_{m-1} \in \mathbf{F}$ כך ש־

$$v = c_0 u + c_1 Tu + \cdots + c_{m-1}T^{m-1}u.$$

יהי j האינדקס הנמוך ביותר כך ש־ $c_j \neq 0$. נפעיל את T^{m-j-1} על שני אגפי המשוואה לעיל ונקבל

$$T^{m-j-1}v = c_j T^{m-1}u,$$

כאשר השתמשנו במשוואה $T^m = 0$. כעת נפעיל את φ על שני אגפי המשוואה לעיל ונקבל

$$\varphi(T^{m-j-1}v) = c_j \varphi(T^{m-1}u) \neq 0.$$

המשוואה לעיל מראה ש־ $v \notin W$. לכן הוכחנו כי $U \cap W = \{0\}$, מה שגורר ש־ $U + W$ סכום ישר (ראו 1.46).

כדי להראות כי $U \oplus W = V$, נגדיר $S : V \to \mathbf{F}^m$ על ידי

$$Sv = \big(\varphi(v), \varphi(Tv), \dots, \varphi(T^{m-1}v)\big).$$

לפיכך $W = \operatorname{null} S$. לכן

$$\dim W = \dim \operatorname{null} S = \dim V - \dim \operatorname{range} S \geq \dim V - m,$$

כאשר השוויון השני נובע מהמשפט היסודי של ההעתקות הליניאריות (3.21). בשימוש באי־שוויון לעיל, נקבל

$$\dim(U \oplus W) = \dim U + \dim W \geq m + (\dim V - m) = \dim V.$$

לפיכך $U \oplus W = V$ (לפי 2.39), ובכך הושלמה ההוכחה. ■

Camille Jordan (1838–1922) published a proof of 8.46 in 1870.

Now the generalized eigenspace decomposition allows us to extend the previous result to operators that may not be nilpotent. Doing this requires that we deal with complex vector spaces.

8.46 *Jordan form*

Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$. Then there is a basis of V that is a Jordan basis for T.

Proof Let $\lambda_1, \dots, \lambda_m$ be the distinct eigenvalues of T. The generalized eigenspace decomposition states that

$$V = G(\lambda_1, T) \oplus \cdots \oplus G(\lambda_m, T),$$

where each $(T - \lambda_k I)|_{G(\lambda_k, T)}$ is nilpotent (see 8.22). Thus 8.45 implies that some basis of each $G(\lambda_k, T)$ is a Jordan basis for $(T - \lambda_k I)|_{G(\lambda_k, T)}$. Put these bases together to get a basis of V that is a Jordan basis for T. ∎

Exercises 8C

1 Suppose $T \in \mathcal{L}(\mathbf{C}^3)$ is the operator defined by $T(z_1, z_2, z_3) = (z_2, z_3, 0)$. Prove that T does not have a square root.

2 Define $T \in \mathcal{L}(\mathbf{F}^5)$ by $T(x_1, x_2, x_3, x_4, x_5) = (2x_2, 3x_3, -x_4, 4x_5, 0)$.

(a) Show that T is nilpotent.

(b) Find a square root of $I + T$.

3 Suppose V is a complex vector space. Prove that every invertible operator on V has a cube root.

4 Suppose V is a real vector space. Prove that the operator $-I$ on V has a square root if and only if $\dim V$ is an even number.

5 Suppose $T \in \mathcal{L}(\mathbf{C}^2)$ is the operator defined by $T(w, z) = (-w - z, 9w + 5z)$. Find a Jordan basis for T.

6 Find a basis of $\mathcal{P}_4(\mathbf{R})$ that is a Jordan basis for the differentiation operator D on $\mathcal{P}_4(\mathbf{R})$ defined by $Dp = p'$.

7 Suppose $T \in \mathcal{L}(V)$ is nilpotent and $v_1, \dots, v_n$ is a Jordan basis for T. Prove that the minimal polynomial of T is z^{m+1}, where m is the length of the longest consecutive string of 1's that appears on the line directly above the diagonal in the matrix of T with respect to $v_1, \dots, v_n$.

8 Suppose $T \in \mathcal{L}(V)$ and $v_1, \dots, v_n$ is a basis of V that is a Jordan basis for T. Describe the matrix of T^2 with respect to this basis.

כעת הפירוק למרחבים עצמיים מוכללים מאפשר לנו להרחיב את המסקנה הקודמת לאופרטורים שייתכן שאינם נילפוטנטיים. למשימה זו נידרש לעסוק במרחבים וקטוריים מרוכבים.

קאמי ז'ורדן (1922–1838) פרסם הוכחה של 8.46 בשנת 1870.

8.46 צורת ז'ורדן

נניח ש־ $\mathbf{F} = \mathbf{C}$ ו־ $T \in \mathcal{L}(V)$. אזי ישנו בסיס של V שהוא בסיס ז'ורדן עבור T.

הוכחה יהיו $\lambda_1, \dots, \lambda_m$ הערכים העצמיים השונים של T. הפירוק למרחבים עצמיים מוכללים קובע כי

$$V = G(\lambda_1, T) \oplus \cdots \oplus G(\lambda_m, T),$$

כאשר כל $(T - \lambda_k I)|_{G(\lambda_k, T)}$ נילפוטנטי (ראו 8.22). לפיכך 8.45 גורר שבסיס כלשהו של כל $G(\lambda_k, T)$ הוא בסיס ז'ורדן עבור $(T - \lambda_k I)|_{G(\lambda_k, T)}$. צירוף בסיסים אלו יחדיו ייתן לנו בסיס של V שהוא בסיס ז'ורדן עבור T. ∎

תרגילים חג

1 נניח ש־ $T \in \mathcal{L}(\mathbf{C}^3)$ הוא האופרטור המוגדר על ידי $T(z_1, z_2, z_3) = (z_2, z_3, 0)$. הוכיחו כי אין ל־ T שורש ריבועי.

2 נגדיר $T \in \mathcal{L}(\mathbf{F}^5)$ על ידי

$$T(x_1, x_2, x_3, x_4, x_5) = (2x_2, 3x_3, -x_4, 4x_5, 0).$$

א. הראו כי T נילפוטנטי.

ב. מצאו שורש ריבועי של $I + T$.

3 נניח ש־ V מרחב וקטורי מרוכב. הוכיחו כי לכל אופרטור הפיך מעל V יש שורש שלישי.

4 נניח ש־ V מרחב וקטורי ממשי. הוכיחו כי לאופרטור $-I$ מעל V יש שורש ריבועי אם־ורק־אם $\dim V$ מספר זוגי.

5 נניח ש־ $T \in \mathcal{L}(\mathbf{C}^2)$ הוא האופרטור המוגדר על ידי $T(w, z) = (-w - z, 9w + 5z)$. מצאו בסיס ז'ורדן עבור T.

6 מצאו בסיס של $\mathcal{P}_4(\mathbf{R})$ שהוא בסיס ז'ורדן עבור אופרטור הגזירה D מעל $\mathcal{P}_4(\mathbf{R})$ המוגדר על ידי $Dp = p'$.

7 נניח ש־ $T \in \mathcal{L}(V)$ נילפוטנטי ו־ $v_1, \dots, v_n$ בסיס ז'ורדן עבור T. הוכיחו כי הפולינום המינימלי של T הוא z^{m+1}, כאשר m הוא אורכה של המחרוזת הארוכה ביותר של 1־ים רצופים המופיעים על הקו שמעל האלכסון במטריצה של T ביחס ל־ $v_1, \dots, v_n$.

8 נניח ש־ $T \in \mathcal{L}(V)$ ו־ $v_1, \dots, v_n$ בסיס של V שהוא בסיס ז'ורדן עבור T. תארו את המטריצה של T^2 ביחס לבסיס זה.

9 Suppose $T \in \mathcal{L}(V)$ is nilpotent. Explain why there exist $v_1, \dots, v_n \in V$ and nonnegative integers $m_1, \dots, m_n$ such that (a) and (b) below both hold.
(a) $T^{m_1}v_1, \dots, Tv_1, v_1, \dots, T^{m_n}v_n, \dots, Tv_n, v_n$ is a basis of V.
(b) $T^{m_1+1}v_1 = \cdots = T^{m_n+1}v_n = 0$.

10 Suppose $T \in \mathcal{L}(V)$ and $v_1, \dots, v_n$ is a basis of V that is a Jordan basis for T. Describe the matrix of T with respect to the basis $v_n, \dots, v_1$ obtained by reversing the order of the v's.

11 Suppose $T \in \mathcal{L}(V)$. Explain why every vector in each Jordan basis for T is a generalized eigenvector of T.

12 Suppose $T \in \mathcal{L}(V)$ is diagonalizable. Show that $\mathcal{M}(T)$ is a diagonal matrix with respect to every Jordan basis for T.

13 Suppose $T \in \mathcal{L}(V)$ is nilpotent. Prove that if $v_1, \dots, v_n$ are vectors in V and $m_1, \dots, m_n$ are nonnegative integers such that

$$T^{m_1}v_1, \dots, Tv_1, v_1, \dots, T^{m_n}v_n, \dots, Tv_n, v_n \text{ is a basis of } V$$

and

$$T^{m_1+1}v_1 = \cdots = T^{m_n+1}v_n = 0,$$

then $T^{m_1}v_1, \dots, T^{m_n}v_n$ is a basis of $\operatorname{null} T$.

This exercise shows that $n = \dim \operatorname{null} T$. Thus the positive integer n that appears above depends only on T and not on the specific Jordan basis chosen for T.

14 Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$. Prove that there does not exist a direct sum decomposition of V into two nonzero subspaces invariant under T if and only if the minimal polynomial of T is of the form $(z - \lambda)^{\dim V}$ for some $\lambda \in \mathbf{C}$.

9 נניח ש־ $T \in \mathcal{L}(V)$ נילפוטנטי. הסבירו מדוע קיימים $v_1, \dots, v_n \in V$ ומספרים שלמים אי־שליליים $m_1, \dots, m_n$ כך ש־ (א) ו־ (ב) להלן שניהם מתקיימים.

א. $T^{m_1}v_1, \dots, Tv_1, v_1, \dots, T^{m_n}v_n, \dots, Tv_n, v_n$ בסיס של V.

ב. $T^{m_1+1}v_1 = \cdots = T^{m_n+1}v_n = 0$.

10 נניח ש־ $T \in \mathcal{L}(V)$ ו־ $v_1, \dots, v_n$ בסיס של V שהוא בסיס ז'ורדן עבור T. תארו את המטריצה של T ביחס לבסיס $v_n, \dots, v_1$ המושג באמצעות היפוך הסדר של ה־ v־ים.

11 נניח ש־ $T \in \mathcal{L}(V)$. הסבירו מדוע כל וקטור בכל בסיס ז'ורדן עבור T הוא וקטור עצמי מוכלל של T.

12 נניח ש־ $T \in \mathcal{L}(V)$ לכסין. הראו כי $\mathcal{M}(T)$ מטריצה אלכסונית ביחס לכל בסיס ז'ורדן עבור T.

13 נניח ש־ $T \in \mathcal{L}(V)$ נילפוטנטי. הוכיחו כי אם $v_1, \dots, v_n$ וקטורים ב־ V ו־ $m_1, \dots, m_n$ מספרים שלמים אי־שליליים כך ש־

$$T^{m_1}v_1, \dots, Tv_1, v_1, \dots, T^{m_n}v_n, \dots, Tv_n, v_n$$

בסיס של V, וגם

$$T^{m_1+1}v_1 = \cdots = T^{m_n+1}v_n = 0,$$

אזי $T^{m_1}v_1, \dots, T^{m_n}v_n$ בסיס של $\operatorname{null} T$.

תרגיל זה מראה כי $n = \dim \operatorname{null} T$. לפיכך המספר השלם החיובי n המופיע לעיל תלוי רק ב־ T ולא בבסיס ז'ורדן ספציפי שנבחר עבור T.

14 נניח ש־ $\mathbf{F} = \mathbf{C}$ ו־ $T \in \mathcal{L}(V)$. הוכיחו כי לא קיים פירוק סכום ישר של V לשני תתי־מרחב השונים מאפס השמורים תחת T אם־ורק־אם הפולינום המינימלי של T הוא מהצורה $(z - \lambda)^{\dim V}$ עבור $\lambda \in \mathbf{C}$ כלשהו.

8D Trace: A Connection Between Matrices and Operators

We begin this section by defining the trace of a square matrix. After developing some properties of the trace of a square matrix, we will use this concept to define the trace of an operator.

8.47 definition: *trace of a matrix*

Suppose A is a square matrix with entries in $\mathbf{F}$. The *trace* of A, denoted $\operatorname{tr} A$, is defined to be the sum of the diagonal entries of A.

8.48 example: *trace of a 3-by-3 matrix*

Suppose

$$A = \begin{pmatrix} 3 & -1 & -2 \\ 5 & 2 & -3 \\ 1 & 6 & 0 \end{pmatrix}.$$

The diagonal entries of A, which are shown in red above, are 3, 2, and 0. Thus $\operatorname{tr} A = 3 + 2 + 0 = 5$.

Matrix multiplication is not commutative, but the next result shows that the order of matrix multiplication does not matter to the trace.

8.49 *trace of AB equals trace of BA*

Suppose A is an m-by-n matrix and B is an n-by-m matrix. Then

$$\operatorname{tr}(AB) = \operatorname{tr}(BA).$$

Proof Suppose

$$A = \begin{pmatrix} A_{1,1} & \cdots & A_{1,n} \\ \vdots & & \vdots \\ A_{m,1} & \cdots & A_{m,n} \end{pmatrix}, \quad B = \begin{pmatrix} B_{1,1} & \cdots & B_{1,m} \\ \vdots & & \vdots \\ B_{n,1} & \cdots & B_{n,m} \end{pmatrix}.$$

The jth term on the diagonal of the m-by-m matrix AB equals $\sum_{k=1}^{n} A_{j,k}B_{k,j}$. Thus

$$\begin{aligned} \operatorname{tr}(AB) &= \sum_{j=1}^{m} \sum_{k=1}^{n} A_{j,k}B_{k,j} \\ &= \sum_{k=1}^{n} \sum_{j=1}^{m} B_{k,j}A_{j,k} \\ &= \sum_{k=1}^{n} (k^{\text{th}} \text{ term on diagonal of the } n\text{-by-}n \text{ matrix } BA) \\ &= \operatorname{tr}(BA), \end{aligned}$$

as desired. ∎

ח4 עקבה: הקשר בין מטריצות ואופרטורים

אנו פותחים יחידה זו בהגדרת העקבה של מטריצה ריבועית. אחרי שנפתח מספר תכונות של עקבת מטריצה ריבועית, נשתמש ברעיון זה כדי להגדיר עקבה של אופרטור.

8.47 הגדרה: עקבת מטריצה

נניח ש־ A מטריצה ריבועית עם איברים ב־ $\mathbf{F}$. העקבה של A, המסומנת $\operatorname{tr} A$, מוגדרת כסכום איברי האלכסון של A.

8.48 דוגמה: עקבה של מטריצה 3×3

נניח ש־

$$A = \begin{pmatrix} 3 & -1 & -2 \\ 5 & 2 & -3 \\ 1 & 6 & 0 \end{pmatrix}.$$

איברי האלכסון של A, המוצגים באדום לעיל, הם 3, 2 ו־ 0. לפיכך $\operatorname{tr} A = 3+2+0 = 5$.

כפל מטריצות אינו קומוטטיבי, אך המסקנה הבאה מראה שהסדר של כפל מטריצות אינו משפיע על העקבה.

8.49 עקבה של AB שווה לעקבה של BA

נניח ש־ A מטריצה $m \times n$ ו־ B מטריצה $n \times m$. אזי

$$\operatorname{tr}(AB) = \operatorname{tr}(BA).$$

הוכחה נניח ש־

$$A = \begin{pmatrix} A_{1,1} & \cdots & A_{1,n} \\ \vdots & & \vdots \\ A_{m,1} & \cdots & A_{m,n} \end{pmatrix}, \quad B = \begin{pmatrix} B_{1,1} & \cdots & B_{1,m} \\ \vdots & & \vdots \\ B_{n,1} & \cdots & B_{n,m} \end{pmatrix}.$$

האיבר ה־ j^{th} על האלכסון של המטריצה AB, שהיא מטריצה $m \times m$, שווה ל־ $\sum_{k=1}^{n} A_{j,k}B_{k,j}$. לפיכך

$$\begin{aligned} \operatorname{tr}(AB) &= \sum_{j=1}^{m} \sum_{k=1}^{n} A_{j,k}B_{k,j} \\ &= \sum_{k=1}^{n} \sum_{j=1}^{m} B_{k,j}A_{j,k} \\ &= \sum_{k=1}^{n} (\text{האיבר ה־ } k^{\text{th}} \text{ על אלכסון } BA \text{ מטריצה } n \times n) \\ &= \operatorname{tr}(BA), \end{aligned}$$

כפי שרצינו. ■

We want to define the trace of an operator $T \in \mathcal{L}(V)$ to be the trace of the matrix of T with respect to some basis of V. However, this definition should not depend on the choice of basis. The following result will make this possible.

8.50 *trace of matrix of operator does not depend on basis*

Suppose $T \in \mathcal{L}(V)$. Suppose $u_1, \dots, u_n$ and $v_1, \dots, v_n$ are bases of V. Then

$$\operatorname{tr} \mathcal{M}(T, (u_1, \dots, u_n)) = \operatorname{tr} \mathcal{M}(T, (v_1, \dots, v_n)).$$

Proof Let $A = \mathcal{M}(T, (u_1, \dots, u_n))$ and $B = \mathcal{M}(T, (v_1, \dots, v_n))$. The change-of-basis formula tells us that there exists an invertible n-by-n matrix C such that $A = C^{-1}BC$ (see 3.84). Thus

$$\begin{aligned} \operatorname{tr} A &= \operatorname{tr}\big((C^{-1}B)C\big) \\ &= \operatorname{tr}\big(C(C^{-1}B)\big) \\ &= \operatorname{tr}\big((CC^{-1})B\big) \\ &= \operatorname{tr} B, \end{aligned}$$

where the second line comes from 8.49. ∎

Because of 8.50, the following definition now makes sense.

8.51 definition: *trace of an operator*

Suppose $T \in \mathcal{L}(V)$. The *trace* of T, denoted $\operatorname{tr} T$, is defined by

$$\operatorname{tr} T = \operatorname{tr} \mathcal{M}(T, (v_1, \dots, v_n)),$$

where $v_1, \dots, v_n$ is any basis of V.

Suppose $T \in \mathcal{L}(V)$ and λ is an eigenvalue of T. Recall that we defined the multiplicity of λ to be the dimension of the generalized eigenspace $G(\lambda, T)$ (see 8.23); we proved that this multiplicity equals $\dim \operatorname{null}(T - \lambda I)^{\dim V}$ (see 8.20). Recall also that if V is a complex vector space, then the sum of the multiplicities of all eigenvalues of T equals $\dim V$ (see 8.25).

In the following result, the sum of the eigenvalues "with each eigenvalue included as many times as its multiplicity" means that if $\lambda_1, \dots, \lambda_m$ are the distinct eigenvalues of T with multiplicities $d_1, \dots, d_m$, then the sum is

$$d_1\lambda_1 + \cdots + d_m\lambda_m.$$

Or if you prefer to work with a list of not-necessarily-distinct eigenvalues, with each eigenvalue included as many times as its multiplicity, then the eigenvalues could be denoted by $\lambda_1, \dots, \lambda_n$ (where n equals $\dim V$) and the sum is

$$\lambda_1 + \cdots + \lambda_n.$$

ברצוננו להגדיר את עקבת האופרטור $T \in \mathcal{L}(V)$ כעקבת המטריצה של T ביחס לבסיס כלשהו של V. עם זאת, הגדרה זו צריכה להיות בלתי תלויה בבסיס הנבחר. המסקנה הבאה מאפשרת לנו להשיג זאת.

8.50 עקבת מטריצה של אופרטור אינה תלויה בבסיס

נניח ש־ $T \in \mathcal{L}(V)$. נניח ש־ $u_1, \dots, u_n$ ו־ $v_1, \dots, v_n$ בסיסים של V. אזי
$$\operatorname{tr} \mathcal{M}(T, (u_1, \dots, u_n)) = \operatorname{tr} \mathcal{M}(T, (v_1, \dots, v_n)).$$

הוכחה יהיו $A = \mathcal{M}(T, (u_1, \dots, u_n))$ ו־ $B = \mathcal{M}(T, (v_1, \dots, v_n))$. הנוסחה לשינוי בסיס אומרת לנו שקיימת $n \times n$ מטריצה C הפיכה כך ש־ $A = C^{-1}BC$ (ראו 3.84). לפיכך

$$\begin{aligned}
\operatorname{tr} A &= \operatorname{tr}\big((C^{-1}B)C\big) \\
&= \operatorname{tr}\big(C(C^{-1}B)\big) \\
&= \operatorname{tr}\big((CC^{-1})B\big) \\
&= \operatorname{tr} B,
\end{aligned}$$

כאשר השורה השנייה נובעת מ־ 8.49. ■

לאור 8.50, ההגדרה הבאה כעת הגיונית.

8.51 הגדרה: עקבת אופרטור

נניח ש־ $T \in \mathcal{L}(V)$. העקבה של T, המסומנת ב־ $\operatorname{tr} T$, מוגדרת על ידי
$$\operatorname{tr} T = \operatorname{tr} \mathcal{M}(T, (v_1, \dots, v_n)),$$
כאשר $v_1, \dots, v_n$ הוא בסיס כלשהו של V.

נניח ש־ $T \in \mathcal{L}(V)$ ו־ λ ערך עצמי של T. ניזכר כי הגדרנו את הריבוי של λ כממד של המרחב העצמי המוכלל $G(\lambda, T)$ (ראו 8.23); הוכחנו כי ריבוי זה שווה ל־ $\dim \operatorname{null}(T - \lambda I)^{\dim V}$ (ראו 8.20). נזכיר בנוסף כי אם V מרחב וקטורי מרוכב, אזי סכום הריבויים של כל הערכים העצמיים של T שווה ל־ $\dim V$ (ראו 8.25).

במסקנה להלן, הסכום של הערכים העצמיים "כאשר מספר המופעים של כל ערך עצמי שווה לריבויו" משמעותו היא שאם $\lambda_1, \dots, \lambda_m$ הם הערכים העצמיים השונים של T עם הריבויים $d_1, \dots, d_m$, אזי הסכום הוא
$$d_1\lambda_1 + \cdots + d_m\lambda_m.$$
לחילופין, אם נעדיף לעבוד עם רשימה של ערכים עצמיים שאינם בהכרח שונים, כאשר מספר המופעים של כל ערך עצמי שווה לריבוי, אזי נוכל לסמן את הערכים העצמיים על ידי $\lambda_1, \dots, \lambda_n$ (כאשר n שווה ל־ $\dim V$) והסכום הוא
$$\lambda_1 + \cdots + \lambda_n.$$

8.52 *on complex vector spaces, trace equals sum of eigenvalues*

Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$. Then $\operatorname{tr} T$ equals the sum of the eigenvalues of T, with each eigenvalue included as many times as its multiplicity.

Proof There is a basis of V with respect to which T has an upper-triangular matrix with the diagonal entries of the matrix consisting of the eigenvalues of T, with each eigenvalue included as many times as its multiplicity—see 8.37. Thus the definition of the trace of an operator along with 8.50, which allows us to use a basis of our choice, implies that $\operatorname{tr} T$ equals the sum of the eigenvalues of T, with each eigenvalue included as many times as its multiplicity. ∎

8.53 example: *trace of an operator on* $\mathbf{C}^3$

Suppose $T \in \mathcal{L}(\mathbf{C}^3)$ is defined by

$$T(z_1, z_2, z_3) = (3z_1 - z_2 - 2z_3, 3z_1 + 2z_2 - 3z_3, z_1 + 2z_2).$$

Then the matrix of T with respect to the standard basis of $\mathbf{C}^3$ is

$$\begin{pmatrix} 3 & -1 & -2 \\ 3 & 2 & -3 \\ 1 & 2 & 0 \end{pmatrix}.$$

Adding up the diagonal entries of this matrix, we see that $\operatorname{tr} T = 5$.

The eigenvalues of T are 1, $2 + 3i$, and $2 - 3i$, each with multiplicity 1, as you can verify. The sum of these eigenvalues, each included as many times as its multiplicity, is $1 + (2 + 3i) + (2 - 3i)$, which equals 5, as expected by 8.52.

The trace has a close connection with the characteristic polynomial. Suppose $\mathbf{F} = \mathbf{C}$, $T \in \mathcal{L}(V)$, and $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of T, with each eigenvalue included as many times as its multiplicity. Then by definition (see 8.26), the characteristic polynomial of T equals

$$(z - \lambda_1) \cdots (z - \lambda_n).$$

Expanding the polynomial above, we can write the characteristic polynomial of T in the form

$$z^n - (\lambda_1 + \cdots + \lambda_n) z^{n-1} + \cdots + (-1)^n (\lambda_1 \cdots \lambda_n).$$

The expression above immediately leads to the next result. Also see 9.65, which does not require the hypothesis that $\mathbf{F} = \mathbf{C}$.

8.54 *trace and characteristic polynomial*

Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$. Let $n = \dim V$. Then $\operatorname{tr} T$ equals the negative of the coefficient of z^{n-1} in the characteristic polynomial of T.

8.52 מעל מרחבים וקטוריים מרוכבים, העקבה שווה לסכום הערכים העצמיים

נניח ש־ $\mathbf{F} = \mathbf{C}$ ו־ $T \in \mathcal{L}(V)$. אזי $\operatorname{tr} T$ שווה לסכום הערכים העצמיים של T, כאשר מספר המופעים של כל ערך עצמי שווה לריבוי.

הוכחה קיים בסיס של V אשר ביחס אליו יש ל־ T מטריצה משולשית עליונה כאשר איברי אלכסון המטריצה מורכבים מהערכים העצמיים של T, כאשר מספר המופעים של כל ערך עצמי שווה לריבויו—ראו 8.37. לפיכך ההגדרה של עקבת אופרטור יחד עם 8.50, המאפשרת לנו להשתמש בבסיס לפי בחירה, גוררת ש־ $\operatorname{tr} T$ שווה לסכום הערכים העצמיים של T, כאשר מספר המופעים של כל ערך עצמי שווה לריבוי. ■

8.53 דוגמה: עקבת אופרטור מעל $\mathbf{C}^3$

נניח ש־ $T \in \mathcal{L}(\mathbf{C}^3)$ מוגדר על ידי

$$T(z_1, z_2, z_3) = (3z_1 - z_2 - 2z_3, 3z_1 + 2z_2 - 3z_3, z_1 + 2z_2).$$

אזי המטריצה של T ביחס לבסיס הסטנדרטי של $\mathbf{C}^3$ היא

$$\begin{pmatrix} 3 & -1 & -2 \\ 3 & 2 & -3 \\ 1 & 2 & 0 \end{pmatrix}.$$

חיבור איברי האלכסון של מטריצה זו מראה לנו כי $\operatorname{tr} T = 5$.

הערכים העצמיים של T הם 1, $2 + 3i$ ו־ $2 - 3i$, כל אחד עם ריבוי 1, כפי שתוכלו לוודא. הסכום של ערכים עצמיים אלו, כאשר מספר המופעים של כל ערך עצמי שווה לריבויו, הוא $1 + (2 + 3i) + (2 - 3i)$, השווה 5, כמצופה לפי 8.52.

לעקבה יש קשר הדוק עם הפולינום האופייני. נניח ש־ $T \in \mathcal{L}(V)$, $\mathbf{F} = \mathbf{C}$, ו־ $\lambda_1, \dots, \lambda_n$ הערכים העצמיים של T, כאשר מספר המופעים של כל ערך עצמי שווה לריבויו. אזי לפי ההגדרה (ראו 8.26), הפולינום האופייני של T שווה ל־

$$(z - \lambda_1) \cdots (z - \lambda_n).$$

בפתיחת סוגריים של הפולינום לעיל, נוכל לכתוב את הפולינום האופייני של T באופן הבא

$$z^n - (\lambda_1 + \cdots + \lambda_n) z^{n-1} + \cdots + (-1)^n (\lambda_1 \cdots \lambda_n).$$

ביטוי זה המובא לעיל מוביל באופן ישיר למסקנה הבאה. בנוסף, ראו מסקנה 9.65, שאינה דורשת את ההנחה כי $\mathbf{F} = \mathbf{C}$.

8.54 עקבה ופולינום אופייני

נניח ש־ $\mathbf{F} = \mathbf{C}$ ו־ $T \in \mathcal{L}(V)$. יהי $n = \dim V$. אזי $\operatorname{tr} T$ שווה לנגדי של המקדם של z^{n-1} בפולינום האופייני של T.

The next result gives a nice formula for the trace of an operator on an inner product space.

8.55 *trace on an inner product space*

Suppose V is an inner product space, $T \in \mathcal{L}(V)$, and $e_1, \dots, e_n$ is an orthonormal basis of V. Then

$$\operatorname{tr} T = \langle Te_1, e_1 \rangle + \cdots + \langle Te_n, e_n \rangle.$$

Proof The desired formula follows from the observation that the entry in row k, column k of $\mathcal{M}(T, (e_1, \dots, e_n))$ equals $\langle Te_k, e_k \rangle$ [use 6.30(a) with $v = Te_k$]. ∎

The algebraic properties of the trace as defined on square matrices translate to algebraic properties of the trace as defined on operators, as shown in the next result.

8.56 *trace is linear*

The function $\operatorname{tr} \colon \mathcal{L}(V) \to \mathbf{F}$ is a linear functional on $\mathcal{L}(V)$ such that

$$\operatorname{tr}(ST) = \operatorname{tr}(TS)$$

for all $S, T \in \mathcal{L}(V)$.

Proof Choose a basis of V. All matrices of operators in this proof will be with respect to that basis. Suppose $S, T \in \mathcal{L}(V)$.

If $\lambda \in \mathbf{F}$, then

$$\operatorname{tr}(\lambda T) = \operatorname{tr} \mathcal{M}(\lambda T) = \operatorname{tr}(\lambda \mathcal{M}(T)) = \lambda \operatorname{tr} \mathcal{M}(T) = \lambda \operatorname{tr} T,$$

where the first and last equalities come from the definition of the trace of an operator, the second equality comes from 3.38, and the third equality follows from the definition of the trace of a square matrix.

Also,

$$\operatorname{tr}(S+T) = \operatorname{tr} \mathcal{M}(S+T) = \operatorname{tr}(\mathcal{M}(S) + \mathcal{M}(T)) = \operatorname{tr} \mathcal{M}(S) + \operatorname{tr} \mathcal{M}(T) = \operatorname{tr} S + \operatorname{tr} T,$$

where the first and last equalities come from the definition of the trace of an operator, the second equality comes from 3.35, and the third equality follows from the definition of the trace of a square matrix. The two paragraphs above show that $\operatorname{tr} \colon \mathcal{L}(V) \to \mathbf{F}$ is a linear functional on $\mathcal{L}(V)$.

Furthermore,

$$\operatorname{tr}(ST) = \operatorname{tr} \mathcal{M}(ST) = \operatorname{tr}(\mathcal{M}(S)\mathcal{M}(T)) = \operatorname{tr}(\mathcal{M}(T)\mathcal{M}(S)) = \operatorname{tr} \mathcal{M}(TS) = \operatorname{tr}(TS),$$

where the second and fourth equalities come from 3.43 and the crucial third equality comes from 8.49. ∎

The equations $\operatorname{tr}(ST) = \operatorname{tr}(TS)$ and $\operatorname{tr} I = \dim V$ uniquely characterize the trace among the linear functionals on $\mathcal{L}(V)$—see Exercise 10.

המסקנה הבאה מספקת נוסחה נחמדה עבור עקבת האופרטור מעל מרחב מכפלה פנימית.

8.55 עקבה מעל מרחב מכפלה פנימית

נניח ש־ V מרחב מכפלה פנימית, $T \in \mathcal{L}(V)$, ו־ $e_1, \dots, e_n$ בסיס אורתונורמלי של V. אזי

$$\operatorname{tr} T = \langle Te_1, e_1 \rangle + \cdots + \langle Te_n, e_n \rangle.$$

הוכחה הנוסחה המבוקשת נובעת מההבחנה כי האיבר בשורה k, עמודה k של $\mathcal{M}(T, (e_1, \dots, e_n))$ שווה ל־ $\langle Te_k, e_k \rangle$ [השתמשו ב־ 6.30(א) כאשר $v = Te_k$].

התכונות האלגבריות של העקבה כפי שהיא מוגדרת מעל מטריצה ריבועית מתורגמות לתכונות אלגבריות של העקבה כפי שהיא מוגדרת מעל אופרטורים, כפי שנראה במסקנה הבאה.

8.56 העקבה היא ליניארית

הפונקציה $\operatorname{tr} : \mathcal{L}(V) \to \mathbf{F}$ היא פונקציונל ליניארי מעל $\mathcal{L}(V)$ כך ש־

$$\operatorname{tr}(ST) = \operatorname{tr}(TS)$$

לכל $S, T \in \mathcal{L}(V)$.

הוכחה נבחר בסיס של V. כל המטריצות של האופרטורים בהוכחה זו תהיינה ביחס לבסיס זה. נניח ש־ $S, T \in \mathcal{L}(V)$.

אם $\lambda \in \mathbf{F}$, אזי

$$\operatorname{tr}(\lambda T) = \operatorname{tr} \mathcal{M}(\lambda T) = \operatorname{tr}(\lambda \mathcal{M}(T)) = \lambda \operatorname{tr} \mathcal{M}(T) = \lambda \operatorname{tr} T,$$

כאשר השוויונות הראשון והאחרון נובעים מהגדרת עקבת האופרטור, השוויון השני נובע מ־ 3.38, והשוויון השלישי נובע מהגדרת העקבה של מטריצה ריבועית.

בנוסף,

$$\operatorname{tr}(S+T) = \operatorname{tr} \mathcal{M}(S+T) = \operatorname{tr}(\mathcal{M}(S) + \mathcal{M}(T)) = \operatorname{tr} \mathcal{M}(S) + \operatorname{tr} \mathcal{M}(T) = \operatorname{tr} S + \operatorname{tr} T,$$

כאשר השוויונות הראשון והאחרון נובעים מהגדרת עקבת האופרטור, השוויון השני נובע מ־ 3.35, והשוויון השלישי נובע מהגדרת העקבה של מטריצה ריבועית. שתי הפסקאות לעיל הראו כי $\operatorname{tr} : \mathcal{L}(V) \to \mathbf{F}$ פונקציונל ליניארי מעל $\mathcal{L}(V)$.

יתר על כן,

$$\operatorname{tr}(ST) = \operatorname{tr} \mathcal{M}(ST) = \operatorname{tr}(\mathcal{M}(S)\mathcal{M}(T)) = \operatorname{tr}(\mathcal{M}(T)\mathcal{M}(S)) = \operatorname{tr} \mathcal{M}(TS) = \operatorname{tr}(TS),$$

כאשר השוויונות השני והרביעי נובעים מ־ 3.43 והשוויון השלישי החשוב נובע מ־ 8.49.

המשוואות $\operatorname{tr}(ST) = \operatorname{tr}(TS)$ ו־ $\operatorname{tr} I = \dim V$ מאפיינות באופן יחיד את העקבה מבין כל הפונקציונלים הליניאריים מעל $\mathcal{L}(V)$—ראו תרגיל 10.

The equation $\operatorname{tr}(ST) = \operatorname{tr}(TS)$ leads to our next result, which does not hold on infinite-dimensional vector spaces (see Exercise 13). However, additional hypotheses on S, T, and V lead to an infinite-dimensional generalization of the result below, with important applications to quantum theory.

The statement of the next result does not involve traces, but the short proof uses traces. When something like this happens in mathematics, then usually a good definition lurks in the background.

8.57 *identity operator is not the difference of ST and TS*

There do not exist operators $S, T \in \mathcal{L}(V)$ such that $ST - TS = I$.

Proof Suppose $S, T \in \mathcal{L}(V)$. Then

$$\operatorname{tr}(ST - TS) = \operatorname{tr}(ST) - \operatorname{tr}(TS) = 0,$$

where both equalities come from 8.56. The trace of I equals $\dim V$, which is not 0. Because $ST - TS$ and I have different traces, they cannot be equal. ∎

Exercises 8D

1 Suppose V is an inner product space and $v, w \in V$. Define an operator $T \in \mathcal{L}(V)$ by $Tu = \langle u, v\rangle w$. Find a formula for $\operatorname{tr} T$.

2 Suppose $P \in \mathcal{L}(V)$ satisfies $P^2 = P$. Prove that

$$\operatorname{tr} P = \dim \operatorname{range} P.$$

3 Suppose $T \in \mathcal{L}(V)$ and $T^5 = T$. Prove that the real and imaginary parts of $\operatorname{tr} T$ are both integers.

4 Suppose V is an inner product space and $T \in \mathcal{L}(V)$. Prove that

$$\operatorname{tr} T^* = \overline{\operatorname{tr} T}.$$

5 Suppose V is an inner product space. Suppose $T \in \mathcal{L}(V)$ is a positive operator and $\operatorname{tr} T = 0$. Prove that $T = 0$.

6 Suppose V is an inner product space and $P, Q \in \mathcal{L}(V)$ are orthogonal projections. Prove that $\operatorname{tr}(PQ) \geq 0$.

7 Suppose $T \in \mathcal{L}(\mathbf{C}^3)$ is the operator whose matrix is

$$\begin{pmatrix} 51 & -12 & -21 \\ 60 & -40 & -28 \\ 57 & -68 & 1 \end{pmatrix}.$$

Someone tells you (accurately) that -48 and 24 are eigenvalues of T. Without using a computer or writing anything down, find the third eigenvalue of T.

המשוואה $\operatorname{tr}(ST) = \operatorname{tr}(TS)$ מובילה אל מסקנתנו הבאה, אשר אינה מתקיימת מעל מרחבים וקטוריים מממד אינסופי (ראו תרגיל 13). בכל זאת, הנחות נוספות לגבי S, T ו־ V מובילות להכללה של המסקנה להלן לממד אינסופי, עם יישומים חשובים בתורת הקוונטים.

הניסוח של המסקנה הבאה אינו כולל עקבות, אז ההוכחה הקצרה משתמשת בעקבות. כאשר תופעה כזו מתרחשת במתמטיקה, בדרך כלל ברקע אורבת לה הגדרה טובה.

8.57 אופרטור הזהות אינו ההפרש של ST ו־ TS

לא קיימים אופרטורים $S, T \in \mathcal{L}(V)$ כך ש־ $ST - TS = I$.

הוכחה נניח ש־ $S, T \in \mathcal{L}(V)$. אזי

$$\operatorname{tr}(ST - TS) = \operatorname{tr}(ST) - \operatorname{tr}(TS) = 0,$$

כאשר שני השוויונות נובעים מ־ 8.56. העקבה של I שווה ל־ $\dim V$, השונה מ־ 0. מאחר שהעקבות של $ST - TS$ ו־ I שונות, האופרטורים $ST - TS$ ו־ I אינם יכולים להיות שווים. ∎

תרגילים 8ח

1 נניח ש־ V מרחב מכפלה פנימית ו־ $v, w \in V$. נגדיר אופרטור $T \in \mathcal{L}(V)$ על ידי $Tu = \langle u, v \rangle w$. מצאו נוסחה עבור $\operatorname{tr} T$.

2 נניח ש־ $P \in \mathcal{L}(V)$ מקיים $P^2 = P$. הוכיחו כי

$$\operatorname{tr} P = \dim \operatorname{range} P.$$

3 נניח ש־ $T \in \mathcal{L}(V)$ ו־ $T^5 = T$. הוכיחו כי החלקים הממשי והמדומה של $\operatorname{tr} T$ שניהם מספרים שלמים.

4 נניח ש־ V מרחב מכפלה פנימית ו־ $T \in \mathcal{L}(V)$. הוכיחו כי

$$\operatorname{tr} T^* = \overline{\operatorname{tr} T}.$$

5 נניח ש־ V מרחב מכפלה פנימית. נניח ש־ $T \in \mathcal{L}(V)$ אופרטור חיובי ו־ $\operatorname{tr} T = 0$. הוכיחו כי $T = 0$.

6 נניח ש־ V מרחב מכפלה פנימית ו־ $P, Q \in \mathcal{L}(V)$ היטלים אורתוגונליים. הוכיחו כי $\operatorname{tr}(PQ) \geq 0$.

7 נניח ש־ $T \in \mathcal{L}(\mathbf{C}^3)$ הוא האופרטור שהמטריצה שלו היא

$$\begin{pmatrix} 51 & -12 & -21 \\ 60 & -40 & -28 \\ 57 & -68 & 1 \end{pmatrix}.$$

מישהו אומר לכם (במדויק) כי 48− ו־ 24 הם ערכים עצמיים של T. בלי שתשתמשו במחשב או בכלי כתיבה כלשהם, מצאו את הערך העצמי השלישי של T.

8 Prove or give a counterexample: If $S, T \in \mathcal{L}(V)$, then $\operatorname{tr}(ST) = (\operatorname{tr} S)(\operatorname{tr} T)$.

9 Suppose $T \in \mathcal{L}(V)$ is such that $\operatorname{tr}(ST) = 0$ for all $S \in \mathcal{L}(V)$. Prove that $T = 0$.

10 Prove that the trace is the only linear functional $\tau : \mathcal{L}(V) \to \mathbf{F}$ such that

$$\tau(ST) = \tau(TS)$$

for all $S, T \in \mathcal{L}(V)$ and $\tau(I) = \dim V$.

Hint: Suppose that $v_1, \dots, v_n$ is a basis of V. For $j, k \in \{1, \dots, n\}$, define $P_{j,k} \in \mathcal{L}(V)$ by $P_{j,k}(a_1 v_1 + \cdots + a_n v_n) = a_k v_j$. Prove that

$$\tau(P_{j,k}) = \begin{cases} 1 & \text{if } j = k, \\ 0 & \text{if } j \neq k. \end{cases}$$

Then for $T \in \mathcal{L}(V)$, use the equation $T = \sum_{k=1}^{n} \sum_{j=1}^{n} \mathcal{M}(T)_{j,k} P_{j,k}$ to show that $\tau(T) = \operatorname{tr} T$.

11 Suppose V and W are inner product spaces and $T \in \mathcal{L}(V, W)$. Prove that if $e_1, \dots, e_n$ is an orthonormal basis of V and $f_1, \dots, f_m$ is an orthonormal basis of W, then

$$\operatorname{tr}(T^* T) = \sum_{k=1}^{n} \sum_{j=1}^{m} |\langle T e_k, f_j \rangle|^2.$$

The numbers $\langle T e_k, f_j \rangle$ are the entries of the matrix of T with respect to the orthonormal bases $e_1, \dots, e_n$ and $f_1, \dots, f_m$. These numbers depend on the bases, but $\operatorname{tr}(T^* T)$ does not depend on a choice of bases. Thus this exercise shows that the sum of the squares of the absolute values of the matrix entries does not depend on which orthonormal bases are used.

12 Suppose V and W are finite-dimensional inner product spaces.

(a) Prove that $\langle S, T \rangle = \operatorname{tr}(T^* S)$ defines an inner product on $\mathcal{L}(V, W)$.

(b) Suppose $e_1, \dots, e_n$ is an orthonormal basis of V and $f_1, \dots, f_m$ is an orthonormal basis of W. Show that the inner product on $\mathcal{L}(V, W)$ from (a) is the same as the standard inner product on $\mathbf{F}^{mn}$, where we identify each element of $\mathcal{L}(V, W)$ with its matrix (with respect to the bases just mentioned) and then with an element of $\mathbf{F}^{mn}$.

*Caution: The norm of a linear map $T \in \mathcal{L}(V, W)$ as defined by 7.86 is not the same as the norm that comes from the inner product in (a) above. Unless explicitly stated otherwise, always assume that $\|T\|$ refers to the norm as defined by 7.86. The norm that comes from the inner product in (a) is called the **Frobenius norm** or the **Hilbert–Schmidt norm**.*

13 Find $S, T \in \mathcal{L}(\mathcal{P}(\mathbf{F}))$ such that $ST - TS = I$.

Hint: Make an appropriate modification of the operators in Example 3.9.

This exercise shows that additional hypotheses are needed on S and T to extend 8.57 to the setting of infinite-dimensional vector spaces.

8 הוכיחו או הביאו דוגמה נגדית: אם $S, T \in \mathcal{L}(V)$, אזי $\operatorname{tr}(ST) = (\operatorname{tr} S)(\operatorname{tr} T)$.

9 נניח ש־ $T \in \mathcal{L}(V)$ כך ש־ $\operatorname{tr}(ST) = 0$ לכל $S \in \mathcal{L}(V)$. הוכיחו כי $T = 0$.

10 הוכיחו כי העקבה היא הפונקציונל הלינארי היחיד $\tau : \mathcal{L}(V) \to \mathbf{F}$ כך ש־

$$\tau(ST) = \tau(TS)$$

לכל $S, T \in \mathcal{L}(V)$ ו־ $\tau(I) = \dim V$.

רמז: נניח ש־ $v_1, \dots, v_n$ בסיס של V. עבור $j, k \in \{1, \dots, n\}$, נגדיר $P_{j,k} \in \mathcal{L}(V)$ על ידי $P_{j,k}(a_1 v_1 + \cdots + a_n v_n) = a_k v_j$. הוכיחו כי

$$\tau(P_{j,k}) = \begin{cases} 1 & \text{if } j = k, \\ 0 & \text{if } j \neq k. \end{cases}$$

ואז עבור $T \in \mathcal{L}(V)$, השתמשו בכך ש־ $T = \sum_{k=1}^{n} \sum_{j=1}^{n} \mathcal{M}(T)_{j,k} P_{j,k}$ כדי להראות כי $\tau(T) = \operatorname{tr} T$.

11 נניח ש־ V ו־ W מרחבי מכפלה פנימית ו־ $T \in \mathcal{L}(V, W)$. הוכיחו כי אם $e_1, \dots, e_n$ בסיס אורתונורמלי של V ו־ $f_1, \dots, f_m$ בסיס אורתונורמלי של W, אזי

$$\operatorname{tr}(T^* T) = \sum_{k=1}^{n} \sum_{j=1}^{m} |\langle T e_k, f_j \rangle|^2.$$

המספרים $\langle T e_k, f_j \rangle$ הם איברי המטריצה של T ביחס אל הבסיסים האורתונורמליים $e_1, \dots, e_n$ ו־ $f_1, \dots, f_m$. מספרים אלו תלויים בבסיסים, אך $\operatorname{tr}(T^* T)$ אינה תלויה בבחירה של בסיסים. לפיכך תרגיל זה מראה כי הסכום של הריבועים של הערכים המוחלטים של איברי המטריצה אינו תלוי בבסיסים האורתונורמליים שבשימוש.

12 נניח ש־ V ו־ W מרחבי מכפלה פנימית מממד סופי.

א. הוכיחו כי $\langle S, T \rangle = \operatorname{tr}(T^* S)$ מגדיר מכפלה פנימית מעל $\mathcal{L}(V, W)$.

ב. נניח ש־ $e_1, \dots, e_n$ בסיס אורתונורמלי של V ו־ $f_1, \dots, f_m$ בסיס אורתונורמלי של W. הראו כי המכפלה הפנימית מעל $\mathcal{L}(V, W)$ מסעיף (א) היא זהה למכפלה הפנימית הסטנדרטית מעל $\mathbf{F}^{mn}$, כאשר אנו מזהים כל איבר של $\mathcal{L}(V, W)$ עם המטריצה שלו (ביחס אל הבסיסים שזה עתה הוזכרו) ואז עם איבר של $\mathbf{F}^{mn}$.

אזהרה: הנורמה של העתקה ליניארית $T \in \mathcal{L}(V, W)$ המוגדרת ב־ 7.86 אינה זהה לנורמה הנובעת מהמכפלה הפנימית בסעיף (א) לעיל. אם לא צוין במפורש אחרת, תמיד נניח ש־ $\|T\|$ מתייחסת לנורמה כפי שהוגדרה ב־ 7.86. הנורמה הנובעת מהמכפלה הפנימית בסעיף (א) נקראת **נורמת פרובניוס** או **נורמת הילברט־שמידט**.

13 מצאו $S, T \in \mathcal{L}(\mathcal{P}(\mathbf{F}))$ כך ש־ $ST - TS = I$.

רמז: בצעו את השינויים המתאימים לאופרטורים בדוגמה 3.9.

תרגיל זה מראה שנדרשות הנחות נוספות לגבי S ו־ T כדי להרחיב את 8.57 למסגרת של מרחבים וקטוריים מממד אינסופי.

אלגברה ליניארית בדרך הישר, מהדורה רביעית, מחבר: שלדון אקסלר
Linear Algebra Done Right, fourth edition, by Sheldon Axler

Chapter 9

Multilinear Algebra and Determinants

We begin this chapter by investigating bilinear forms and quadratic forms on a vector space. Then we will move on to multilinear forms. We will show that the vector space of alternating n-linear forms has dimension one on a vector space of dimension n. This result will allow us to give a clean basis-free definition of the determinant of an operator.

This approach to the determinant via alternating multilinear forms leads to straightforward proofs of key properties of determinants. For example, we will see that the determinant is multiplicative, meaning that $\det(ST) = (\det S)(\det T)$ for all operators S and T on the same vector space. We will also see that T is invertible if and only if $\det T \neq 0$. Another important result states that the determinant of an operator on a complex vector space equals the product of the eigenvalues of the operator, with each eigenvalue included as many times as its multiplicity.

The chapter concludes with an introduction to tensor products.

standing assumptions for this chapter

- **F** denotes **R** or **C**.
- V and W denote finite-dimensional nonzero vector spaces over **F**.

Daniel Schwen CC BY-SA

The Mathematical Institute at the University of Göttingen. This building opened in 1930, when Emmy Noether (1882–1935) had already been a research mathematician and faculty member at the university for 15 years (the first eight years without salary). Noether was fired by the Nazi government in 1933. By then Noether and her collaborators had created many of the foundations of modern algebra, including an abstract algebra viewpoint that contributed to the development of linear algebra.

פרק ט

אלגברה מולטי־ליניארית ודטרמיננטות

את הפרק אנו פותחים בחקירת תבניות ביליניאריות ותבניות ריבועיות מעל מרחב וקטורי. לאחר מכן אנו פונים אל תבניות מולטי־ליניאריות. אנו נראה כי המרחב הווקטורי של תבניות n־ליניאריות מתחלפות הוא חד־ממדי מעל מרחב וקטורי מממד n. מסקנה זו תאפשר לנו לספק הגדרה נקייה שאינה תלויה בבחירת בסיס עבור הדטרמיננטה של אופרטור.

גישה זו לדטרמיננטה באמצעות תבניות מולטי־ליניאריות מתחלפות מובילה להוכחות פשוטות של תכונות מפתח של דטרמיננטות. לדוגמה, אנו נראה כי הדטרמיננטה היא כפלית, כלומר $\det(ST) = (\det S)(\det T)$ לכל האופרטורים S ו־ T מעל אותו מרחב וקטורי. בנוסף נראה כי T הפיך אם־ורק־אם $\det T \neq 0$. מסקנה חשובה נוספת קובעת כי הדטרמיננטה של אופרטור מעל מרחב וקטורי מרוכב שווה למכפלת הערכים העצמיים של האופרטור, כאשר כל ערך עצמי נכלל מספר פעמים השווה לריבויו.

את הפרק אנו חותמים בהיכרות עם מכפלות טנזוריות.

הנחות המוצא לפרק זה

- **F** מסמן **R** או **C**.
- V ו־ W מסמנים מרחבים וקטוריים מממד סופי השונים מאפס מעל **F**.

Daniel Schwen CC BY-SA

המכון למתמטיקה באוניברסיטת גטינגן. בניין זה נפתח בשנת 1930, כאשר אֶמִי נֶתֶר (1882–1935) כבר הייתה מתמטיקאית העוסקת במחקר וחברת סגל באוניברסיטה מזה 15 שנים (בשמונה השנים הראשונות ללא שכר). ב־1933, לאחר עליית הנאצים לשלטון, נתר, בדומה למרצים יהודים אחרים, פוטרה מאוניברסיטת גטינגן משום היותה יהודייה. בשלב זה נתר ועמיתיה כבר יצרו רבים מיסודות האלגברה המודרנית, כולל נקודת מבט של אלגברה מופשטת שתרמה לפיתוחה של אלגברה ליניארית.

9A Bilinear Forms and Quadratic Forms

Bilinear Forms

A bilinear form on V is a function from $V \times V$ to $\mathbf{F}$ that is linear in each slot separately, meaning that if we hold either slot fixed then we have a linear function in the other slot. Here is the formal definition.

9.1 definition: *bilinear form*

A *bilinear form* on V is a function $\beta : V \times V \to \mathbf{F}$ such that

$$v \mapsto \beta(v, u) \quad \text{and} \quad v \mapsto \beta(u, v)$$

are both linear functionals on V for every $u \in V$.

*Recall that the term **linear functional**, used in the definition above, means a linear function that maps into the scalar field* $\mathbf{F}$. *Thus the term **bilinear functional** would be more consistent terminology than **bilinear form**, which unfortunately has become standard.*

For example, if V is a real inner product space, then the function that takes an ordered pair $(u, v) \in V \times V$ to $\langle u, v\rangle$ is a bilinear form on V. If V is a nonzero complex inner product space, then this function is not a bilinear form because the inner product is not linear in the second slot (complex scalars come out of the second slot as their complex conjugates).

If $\mathbf{F} = \mathbf{R}$, then a bilinear form differs from an inner product in that an inner product requires symmetry [meaning that $\beta(v, w) = \beta(w, v)$ for all $v, w \in V$] and positive definiteness [meaning that $\beta(v, v) > 0$ for all $v \in V \backslash \{0\}$], but these properties are not required for a bilinear form.

9.2 example: *bilinear forms*

- The function $\beta : \mathbf{F}^3 \times \mathbf{F}^3 \to \mathbf{F}$ defined by

$$\beta((x_1, x_2, x_3), (y_1, y_2, y_3)) = x_1 y_2 - 5x_2 y_3 + 2x_3 y_1$$

  is a bilinear form on $\mathbf{F}^3$.

- Suppose A is an n-by-n matrix with $A_{j,k} \in \mathbf{F}$ in row j, column k. Define a bilinear form β_A on $\mathbf{F}^n$ by

$$\beta_A((x_1, \dots, x_n), (y_1, \dots, y_n)) = \sum_{k=1}^{n} \sum_{j=1}^{n} A_{j,k} x_j y_k.$$

  The first bullet point is a special case of this bullet point with $n = 3$ and

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & -5 \\ 2 & 0 & 0 \end{pmatrix}.$$

טו תבניות ביליניאריות ותבניות ריבועיות

תבניות ביליניאריות

תבנית בילינארית מעל V היא פונקציה מ־ $V \times V$ אל $\mathbf{F}$ שהיא לינארית בכל רכיב בנפרד, משמעות הדבר היא כי בקיבוע כל אחד משני הרכיבים מתקבלת פונקציה לינארית ברכיב האחר. להלן ההגדרה הרשמית.

9.1 הגדרה: תבנית ביליניארית

תבנית ביליניארית מעל V היא פונקציה $\beta : V \times V \to \mathbf{F}$ כך ש־
$v \mapsto \beta(v,u)$ וגם $v \mapsto \beta(u,v)$
שניהם פונקציונלים ליניאריים מעל V לכל $u \in V$.

לדוגמה, אם V מרחב מכפלה פנימית ממשי, אזי הפונקציה המעתיקה זוג סדור $(u,v) \in V \times V$ אל $\langle u,v\rangle$ היא תבנית בילינארית מעל V. אם V מרחב מכפלה פנימית מרוכב השונה מאפס, אזי פונקציה זו אינה תבנית בילינארית מאחר שהמכפלה הפנימית אינה לינארית ברכיב השני (סקלרים מרוכבים נשלפים מהרכיב השני כצמוד המרוכב שלהם).

נזכיר כי המונח **פונקציונל ליניארי**, המופיע בהגדרה לעיל, משמעותו פונקציה ליניארית המעתיקה אל שדה הסקלרים $\mathbf{F}$. לפיכך המונח **פונקציונל ביליניארי** היה מעיד מיד עקבי יותר מאשר **תבנית ביליניארית**, שלמרבה המצער הפך לסטנדרט.

אם $\mathbf{F} = \mathbf{R}$, אזי תבנית בילינארית שונה ממכפלה פנימית בכך שבעוד מכפלה פנימית דורשת סימטריה [כלומר $\beta(v,w) = \beta(w,v)$ לכל $v,w \in V$] חיוביות לחלוטין [כלומר $\beta(v,v) > 0$ לכל $v \in V\backslash\{0\}$], תכונות אלו אינן נדרשות עבור תבנית בילינארית.

9.2 דוגמה: תבניות ביליניאריות

- הפונקציה $\beta : \mathbf{F}^3 \times \mathbf{F}^3 \to \mathbf{F}$ המוגדרת על ידי
$$\beta((x_1,x_2,x_3),(y_1,y_2,y_3)) = x_1y_2 - 5x_2y_3 + 2x_3y_1$$
היא תבנית בילינארית מעל $\mathbf{F}^3$.

- נניח ש־ A מטריצה $n \times n$ עם $A_{j,k} \in \mathbf{F}$ בשורה j, עמודה k. נגדיר תבנית בילינארית β_A מעל $\mathbf{F}^n$ על ידי
$$\beta_A((x_1,\dots,x_n),(y_1,\dots,y_n)) = \sum_{k=1}^{n}\sum_{j=1}^{n} A_{j,k}x_jy_k.$$
הסעיף הראשון הוא מקרה פרטי של סעיף זה עם $n = 3$ ו־
$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & -5 \\ 2 & 0 & 0 \end{pmatrix}.$$

- Suppose V is a real inner product space and $T \in \mathcal{L}(V)$. Then the function $\beta\colon V \times V \to \mathbf{R}$ defined by
$$\beta(u, v) = \langle u, Tv\rangle$$
is a bilinear form on V.
- If n is a positive integer, then the function $\beta\colon \mathcal{P}_n(\mathbf{R}) \times \mathcal{P}_n(\mathbf{R}) \to \mathbf{R}$ defined by
$$\beta(p, q) = p(2) \cdot q'(3)$$
is a bilinear form on $\mathcal{P}_n(\mathbf{R})$.
- Suppose $\varphi, \tau \in V'$. Then the function $\beta\colon V \times V \to \mathbf{F}$ defined by
$$\beta(u, v) = \varphi(u) \cdot \tau(v)$$
is a bilinear form on V.
- More generally, suppose that $\varphi_1, \dots, \varphi_n, \tau_1, \dots, \tau_n \in V'$. Then the function $\beta\colon V \times V \to \mathbf{F}$ defined by
$$\beta(u, v) = \varphi_1(u) \cdot \tau_1(v) + \cdots + \varphi_n(u) \cdot \tau_n(v)$$
is a bilinear form on V.

A bilinear form on V is a function from $V \times V$ to $\mathbf{F}$. Because $V \times V$ is a vector space, this raises the question of whether a bilinear form can also be a linear map from $V \times V$ to $\mathbf{F}$. Note that none of the bilinear forms in 9.2 are linear maps except in some special cases in which the bilinear form is the zero map. Exercise 3 shows that a bilinear form β on V is a linear map on $V \times V$ only if $\beta = 0$.

9.3 definition: $V^{(2)}$

The set of bilinear forms on V is denoted by $V^{(2)}$.

With the usual operations of addition and scalar multiplication of functions, $V^{(2)}$ is a vector space.

For T an operator on an n-dimensional vector space V and a basis $e_1, \dots, e_n$ of V, we used an n-by-n matrix to provide information about T. We now do the same thing for bilinear forms on V.

9.4 definition: *matrix of a bilinear form*, $\mathcal{M}(\beta)$

Suppose β is a bilinear form on V and $e_1, \dots, e_n$ is a basis of V. The *matrix* of β with respect to this basis is the n-by-n matrix $\mathcal{M}(\beta)$ whose entry $\mathcal{M}(\beta)_{j,k}$ in row j, column k is given by
$$\mathcal{M}(\beta)_{j,k} = \beta(e_j, e_k).$$
If the basis $e_1, \dots, e_n$ is not clear from the context, then the notation $\mathcal{M}(\beta, (e_1, \dots, e_n))$ is used.

- נניח ש־ V מרחב מכפלה פנימית ממשי ו־ $T \in \mathcal{L}(V)$. אזי הפונקציה $\beta : V \times V \to \mathbf{R}$ המוגדרת על ידי
$$\beta(u, v) = \langle u, Tv \rangle$$
היא תבנית בילינארית מעל V.

- אם n מספר שלם חיובי, אזי הפונקציה $\beta : \mathcal{P}_n(\mathbf{R}) \times \mathcal{P}_n(\mathbf{R}) \to \mathbf{R}$ המוגדרת על ידי
$$\beta(p, q) = p(2) \cdot q'(3)$$
היא תבנית בילינארית מעל $\mathcal{P}_n(\mathbf{R})$.

- נניח ש־ $\varphi, \tau \in V'$. אזי הפונקציה $\beta : V \times V \to \mathbf{F}$ המוגדרת על ידי
$$\beta(u, v) = \varphi(u) \cdot \tau(v)$$
היא תבנית בילינארית מעל V.

- באופן כללי יותר, נניח ש־ $\varphi_1, \dots, \varphi_n, \tau_1, \dots, \tau_n \in V'$. אזי הפונקציה $\beta : V \times V \to \mathbf{F}$ המוגדרת על ידי
$$\beta(u, v) = \varphi_1(u) \cdot \tau_1(v) + \cdots + \varphi_n(u) \cdot \tau_n(v)$$
היא תבנית בילינארית מעל V.

תבנית בילינארית מעל V היא פונקציה מ־ $V \times V$ אל $\mathbf{F}$. מאחר ש־ $V \times V$ מרחב וקטורי, הדבר מעורר את השאלה האם תבנית בילינארית יכולה גם היא להיות העתקה לינארית מ־ $V \times V$ אל $\mathbf{F}$. נשים לב שאף אחת מהתבניות הבילינאריות ב־ 9.2 אינה העתקה לינארית למעט מקרים פרטיים מסוימים כאשר התבנית הבילינארית היא העתקת האפס. תרגיל 3 מראה כי תבנית בילינארית β מעל V היא העתקה לינארית מעל $V \times V$ רק אם $\beta = 0$.

9.3 הגדרה: $V^{(2)}$

קבוצת התבניות הבילינאריות מעל V מסומנת על ידי $V^{(2)}$.

עם הפעולות הרגילות של חיבור וכפל בסקלר של פונקציות, $V^{(2)}$ הוא מרחב וקטורי.

עבור T אופרטור מעל מרחב וקטורי n־ממדי V ובסיס $e_1, \dots, e_n$ של V, השתמשנו במטריצה $n \times n$ כדי לספק מידע לגבי T. כעת אנו עושים את אותו הדבר עבור תבניות בילינאריות מעל V.

9.4 הגדרה: מטריצה של תבנית בילינארית, $\mathcal{M}(\beta)$

נניח ש־ β תבנית בילינארית מעל V ו־ $e_1, \dots, e_n$ בסיס של V. המטריצה של β ביחס לבסיס זה היא $\mathcal{M}(\beta)$ מטריצה $n \times n$ שהאיבר $\mathcal{M}(\beta)_{j,k}$ בשורה j, עמודה k נתון על ידי
$$\mathcal{M}(\beta)_{j,k} = \beta(e_j, e_k).$$
אם הבסיס $e_1, \dots, e_n$ אינו ברור מההקשר, אזי נעשה שימוש בסימון $\mathcal{M}(\beta, (e_1, \dots, e_n))$.

Recall that $\mathbf{F}^{n,n}$ denotes the vector space of n-by-n matrices with entries in $\mathbf{F}$ and that $\dim \mathbf{F}^{n,n} = n^2$ (see 3.39 and 3.40).

9.5 $\dim V^{(2)} = (\dim V)^2$

Suppose $e_1, \dots, e_n$ is a basis of V. Then the map $\beta \mapsto \mathcal{M}(\beta)$ is an isomorphism of $V^{(2)}$ onto $\mathbf{F}^{n,n}$. Furthermore, $\dim V^{(2)} = (\dim V)^2$.

Proof The map $\beta \mapsto \mathcal{M}(\beta)$ is clearly a linear map of $V^{(2)}$ into $\mathbf{F}^{n,n}$. For $A \in \mathbf{F}^{n,n}$, define a bilinear form β_A on V by

$$\beta_A(x_1e_1 + \cdots + x_ne_n, y_1e_1 + \cdots + y_ne_n) = \sum_{k=1}^{n} \sum_{j=1}^{n} A_{j,k}x_jy_k$$

for $x_1, \dots, x_n, y_1, \dots, y_n \in \mathbf{F}$ (if $V = \mathbf{F}^n$ and $e_1, \dots, e_n$ is the standard basis of $\mathbf{F}^n$, this β_A is the same as the bilinear form β_A in the second bullet point of Example 9.2).

The linear map $\beta \mapsto \mathcal{M}(\beta)$ from $V^{(2)}$ to $\mathbf{F}^{n,n}$ and the linear map $A \mapsto \beta_A$ from $\mathbf{F}^{n,n}$ to $V^{(2)}$ are inverses of each other because $\beta_{\mathcal{M}(\beta)} = \beta$ for all $\beta \in V^{(2)}$ and $\mathcal{M}(\beta_A) = A$ for all $A \in \mathbf{F}^{n,n}$, as you should verify.

Thus both maps are isomorphisms and the two spaces that they connect have the same dimension. Hence $\dim V^{(2)} = \dim \mathbf{F}^{n,n} = n^2 = (\dim V)^2$. ∎

Recall that C^t denotes the transpose of a matrix C. The matrix C^t is obtained by interchanging the rows and the columns of C.

9.6 *composition of a bilinear form and an operator*

Suppose β is a bilinear form on V and $T \in \mathcal{L}(V)$. Define bilinear forms α and ρ on V by

$$\alpha(u, v) = \beta(u, Tv) \quad \text{and} \quad \rho(u, v) = \beta(Tu, v).$$

Let $e_1, \dots, e_n$ be a basis of V. Then

$$\mathcal{M}(\alpha) = \mathcal{M}(\beta)\mathcal{M}(T) \quad \text{and} \quad \mathcal{M}(\rho) = \mathcal{M}(T)^t\mathcal{M}(\beta).$$

Proof If $j, k \in \{1, \dots, n\}$, then

$$\begin{aligned}
\mathcal{M}(\alpha)_{j,k} &= \alpha(e_j, e_k) \\
&= \beta(e_j, Te_k) \\
&= \beta\Big(e_j, \sum_{m=1}^{n} \mathcal{M}(T)_{m,k}\, e_m\Big) \\
&= \sum_{m=1}^{n} \beta(e_j, e_m)\mathcal{M}(T)_{m,k} \\
&= (\mathcal{M}(\beta)\mathcal{M}(T))_{j,k}.
\end{aligned}$$

Thus $\mathcal{M}(\alpha) = \mathcal{M}(\beta)\mathcal{M}(T)$. The proof that $\mathcal{M}(\rho) = \mathcal{M}(T)^t\mathcal{M}(\beta)$ is similar. ∎

נזכיר כי $\mathbf{F}^{n,n}$ מסמן את המרחב הווקטורי של מטריצות $n \times n$ עם איברים ב־ $\mathbf{F}$ וכי $\dim \mathbf{F}^{n,n} = n^2$ (ראו 3.39 ו־ 3.40).

9.5 $\dim V^{(2)} = (\dim V)^2$

נניח ש־ $e_1, \dots, e_n$ בסיס של V. אזי ההעתקה $\beta \mapsto \mathcal{M}(\beta)$ היא איזומורפיזם של $V^{(2)}$ על $\mathbf{F}^{n,n}$. יתר על כן, $\dim V^{(2)} = (\dim V)^2$.

הוכחה ההעתקה $\beta \mapsto \mathcal{M}(\beta)$ היא בבירור העתקה לינארית של $V^{(2)}$ אל $\mathbf{F}^{n,n}$.
עבור $A \in \mathbf{F}^{n,n}$, נגדיר תבנית בילינארית β_A מעל V על ידי

$$\beta_A(x_1e_1 + \cdots + x_ne_n, y_1e_1 + \cdots + y_ne_n) = \sum_{k=1}^{n} \sum_{j=1}^{n} A_{j,k}x_jy_k$$

עבור $x_1, \dots, x_n, y_1, \dots, y_n \in \mathbf{F}$ (אם $V = \mathbf{F}^n$ ו־ $e_1, \dots, e_n$ הוא הבסיס הסטנדרטי של $\mathbf{F}^n$, אזי β_A זו זהה לתבנית הבילינארית β_A בסעיף השני של דוגמה 9.2).
ההעתקה הלינארית $\beta \mapsto \mathcal{M}(\beta)$ מ־ $V^{(2)}$ אל $\mathbf{F}^{n,n}$ וההעתקה הלינארית $A \mapsto \beta_A$ מ־ $\mathbf{F}^{n,n}$ אל $V^{(2)}$ הופכיות אחת של השנייה מאחר ש־ $\beta_{\mathcal{M}(\beta)} = \beta$ לכל $\beta \in V^{(2)}$ ו־ $\mathcal{M}(\beta_A) = A$ לכל $A \in \mathbf{F}^{n,n}$, כפי שעליכם לוודא.
לפיכך שתי ההעתקות הן איזומורפיזמים ולשני המרחבים שהן מחברות ממד זהה. לכן $\dim V^{(2)} = \dim \mathbf{F}^{n,n} = n^2 = (\dim V)^2$. ■

נזכיר כי C^t מסמן את המטריצה המשוחלפת של C. המטריצה C^t מושגת על ידי החלפת השורות והעמודות של C.

9.6 הרכבה של תבנית בילינארית ואופרטור

נניח ש־ β תבנית בילינארית מעל V ו־ $T \in \mathcal{L}(V)$. נגדיר תבניות בילינאריות α ו־ ρ מעל V על ידי

$$\rho(u, v) = \beta(Tu, v) \quad \text{וגם} \quad \alpha(u, v) = \beta(u, Tv).$$

יהי $e_1, \dots, e_n$ בסיס של V. אזי

$$\mathcal{M}(\rho) = \mathcal{M}(T)^t\mathcal{M}(\beta) \quad \text{וגם} \quad \mathcal{M}(\alpha) = \mathcal{M}(\beta)\mathcal{M}(T).$$

הוכחה אם $j, k \in \{1, \dots, n\}$, אזי

$$\begin{aligned}
\mathcal{M}(\alpha)_{j,k} &= \alpha(e_j, e_k) \\
&= \beta(e_j, Te_k) \\
&= \beta\Big(e_j, \sum_{m=1}^{n} \mathcal{M}(T)_{m,k}\, e_m\Big) \\
&= \sum_{m=1}^{n} \beta(e_j, e_m)\mathcal{M}(T)_{m,k} \\
&= (\mathcal{M}(\beta)\mathcal{M}(T))_{j,k}.
\end{aligned}$$

לפיכך $\mathcal{M}(\alpha) = \mathcal{M}(\beta)\mathcal{M}(T)$. ההוכחה כי $\mathcal{M}(\rho) = \mathcal{M}(T)^t\mathcal{M}(\beta)$ דומה. ■

The result below shows how the matrix of a bilinear form changes if we change the basis. The formula in the result below should be compared to the change-of-basis formula for the matrix of an operator (see 3.84). The two formulas are similar, except that the transpose C^t appears in the formula below and the inverse C^{-1} appears in the change-of-basis formula for the matrix of an operator.

9.7 *change-of-basis formula*

Suppose $\beta \in V^{(2)}$. Suppose $e_1, \dots, e_n$ and $f_1, \dots, f_n$ are bases of V. Let

$$A = \mathcal{M}(\beta, (e_1, \dots, e_n)) \quad \text{and} \quad B = \mathcal{M}(\beta, (f_1, \dots, f_n))$$

and $C = \mathcal{M}(I, (e_1, \dots, e_n), (f_1, \dots, f_n))$. Then

$$A = C^t B C.$$

Proof The linear map lemma (3.4) tells us that there exists an operator $T \in \mathcal{L}(V)$ such that $Tf_k = e_k$ for each $k = 1, \dots, n$. The definition of the matrix of an operator with respect to a basis implies that

$$\mathcal{M}(T, (f_1, \dots, f_n)) = C.$$

Define bilinear forms α, ρ on V by

$$\alpha(u, v) = \beta(u, Tv) \quad \text{and} \quad \rho(u, v) = \alpha(Tu, v) = \beta(Tu, Tv).$$

Then $\beta(e_j, e_k) = \beta(Tf_j, Tf_k) = \rho(f_j, f_k)$ for all $j, k \in \{1, \dots, n\}$. Thus

$$\begin{aligned} A &= \mathcal{M}(\rho, (f_1, \dots, f_n)) \\ &= C^t \mathcal{M}(\alpha, (f_1, \dots, f_n)) \\ &= C^t B C, \end{aligned}$$

where the second and third lines each follow from 9.6. ∎

9.8 example: *the matrix of a bilinear form on* $\mathcal{P}_2(\mathbf{R})$

Define a bilinear form β on $\mathcal{P}_2(\mathbf{R})$ by $\beta(p, q) = p(2) \cdot q'(3)$. Let

$$A = \mathcal{M}(\beta, (1, x - 2, (x - 3)^2)) \quad \text{and} \quad B = \mathcal{M}(\beta, (1, x, x^2))$$

and $C = \mathcal{M}(I, (1, x - 2, (x - 3)^2), (1, x, x^2))$. Then

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 0 & 1 & 6 \\ 0 & 2 & 12 \\ 0 & 4 & 24 \end{pmatrix} \quad \text{and} \quad C = \begin{pmatrix} 1 & -2 & 9 \\ 0 & 1 & -6 \\ 0 & 0 & 1 \end{pmatrix}.$$

Now the change-of-basis formula 9.7 asserts that $A = C^t B C$, which you can verify with matrix multiplication using the matrices above.

המסקנה להלן מראה כיצד המטריצה של תבנית בילינארית משתנה אם אנו משנים את הבסיס. את הנוסחה במסקנה להלן יש להשוות לנוסחת שינוי בסיס עבור המטריצה של אופרטור (ראו 3.84). שתי הנוסחאות דומות, מלבד העובדה כי המשוחלפת C^t מופיעה בנוסחה להלן וההופכית C^{-1} מופיעה בנוסחת שינוי בסיס עבור המטריצה של אופרטור.

9.7 נוסחת שינוי בסיס

נניח ש־ $\beta \in V^{(2)}$. נניח ש־ $e_1, \dots, e_n$ ו־ $f_1, \dots, f_n$ בסיסים של V. תהיינה

$$A = \mathcal{M}(\beta, (e_1, \dots, e_n)) \quad \text{ו־} \quad B = \mathcal{M}(\beta, (f_1, \dots, f_n))$$

ו־ $C = \mathcal{M}(I, (e_1, \dots, e_n), (f_1, \dots, f_n))$. אזי

$$A = C^t B C.$$

הוכחה למת ההעתקה הליניארית (3.4) מורה לנו כי קיים אופרטור $T \in \mathcal{L}(V)$ כך ש־ $Tf_k = e_k$ לכל $k = 1, \dots, n$. הגדרת המטריצה של אופרטור ביחס לבסיס גוררת ש־

$$\mathcal{M}(T, (f_1, \dots, f_n)) = C.$$

נגדיר תבניות בילינאריות α, ρ מעל V על ידי

$$\rho(u, v) = \alpha(Tu, v) = \beta(Tu, Tv) \quad \text{ו־} \quad \alpha(u, v) = \beta(u, Tv).$$

אזי $\beta(e_j, e_k) = \beta(Tf_j, Tf_k) = \rho(f_j, f_k)$ לכל $j, k \in \{1, \dots, n\}$. לפיכך

$$\begin{aligned} A &= \mathcal{M}(\rho, (f_1, \dots, f_n)) \\ &= C^t \mathcal{M}(\alpha, (f_1, \dots, f_n)) \\ &= C^t B C, \end{aligned}$$

כאשר השורות השנייה והשלישית נובעות מ־ 9.6. ■

9.8 דוגמה: המטריצה של תבנית בילינארית מעל $\mathcal{P}_2(\mathbf{R})$

נגדיר תבנית בילינארית β מעל $\mathcal{P}_2(\mathbf{R})$ על ידי $\beta(p, q) = p(2) \cdot q'(3)$. יהיו

$$B = \mathcal{M}(\beta, (1, x, x^2)) \quad \text{ו־} \quad A = \mathcal{M}(\beta, (1, x-2, (x-3)^2))$$

ו־ $C = \mathcal{M}(I, (1, x-2, (x-3)^2), (1, x, x^2))$. אזי

$$C = \begin{pmatrix} 1 & -2 & 9 \\ 0 & 1 & -6 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{וגם} \quad B = \begin{pmatrix} 0 & 1 & 6 \\ 0 & 2 & 12 \\ 0 & 4 & 24 \end{pmatrix} \quad \text{וגם} \quad A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

כעת, נוסחת שינוי הבסיס 9.7 קובעת כי $A = C^t B C$, קביעה שתוכלו לוודא באמצעות כפל מטריצות של לעיל.

Symmetric Bilinear Forms

9.9 definition: *symmetric bilinear form*, $V_{\text{sym}}^{(2)}$

A bilinear form $\rho \in V^{(2)}$ is called *symmetric* if

$$\rho(u, w) = \rho(w, u)$$

for all $u, w \in V$. The set of symmetric bilinear forms on V is denoted by $V_{\text{sym}}^{(2)}$.

9.10 example: *symmetric bilinear forms*

- If V is a real inner product space and $\rho \in V^{(2)}$ is defined by
$$\rho(u, w) = \langle u, w \rangle,$$
then ρ is a symmetric bilinear form on V.
- Suppose V is a real inner product space and $T \in \mathcal{L}(V)$. Define $\rho \in V^{(2)}$ by
$$\rho(u, w) = \langle u, Tw \rangle.$$
Then ρ is a symmetric bilinear form on V if and only if T is a self-adjoint operator (the previous bullet point is the special case $T = I$).
- Suppose $\rho\colon \mathcal{L}(V) \times \mathcal{L}(V) \to \mathbf{F}$ is defined by
$$\rho(S, T) = \operatorname{tr}(ST).$$
Then ρ is a symmetric bilinear form on $\mathcal{L}(V)$ because trace is a linear functional on $\mathcal{L}(V)$ and $\operatorname{tr}(ST) = \operatorname{tr}(TS)$ for all $S, T \in \mathcal{L}(V)$; see 8.56.

9.11 definition: *symmetric matrix*

A square matrix A is called *symmetric* if it equals its transpose.

An operator on V may have a symmetric matrix with respect to some but not all bases of V. In contrast, the next result shows that a bilinear form on V has a symmetric matrix with respect to either all bases of V or with respect to no bases of V.

9.12 *symmetric bilinear forms are diagonalizable*

Suppose $\rho \in V^{(2)}$. Then the following are equivalent.

(a) ρ is a symmetric bilinear form on V.

(b) $\mathcal{M}(\rho, (e_1, \dots, e_n))$ is a symmetric matrix for every basis $e_1, \dots, e_n$ of V.

(c) $\mathcal{M}(\rho, (e_1, \dots, e_n))$ is a symmetric matrix for some basis $e_1, \dots, e_n$ of V.

(d) $\mathcal{M}(\rho, (e_1, \dots, e_n))$ is a diagonal matrix for some basis $e_1, \dots, e_n$ of V.

תבניות ביליניאריות סימטריות

9.9 הגדרה: תבנית ביליניארית סימטרית, $V^{(2)}_{\text{sym}}$

תבנית בילינארית $\rho \in V^{(2)}$ נקראת סימטרית אם
$$\rho(u,w) = \rho(w,u)$$
לכל $u, w \in V$. קבוצת התבניות הביליניאריות הסימטריות מעל V מסומנת על ידי $V^{(2)}_{\text{sym}}$.

9.10 דוגמה: תבניות ביליניאריות סימטריות

- אם V מרחב מכפלה פנימית ממשי ו־ $\rho \in V^{(2)}$ מוגדר על ידי
$$\rho(u,w) = \langle u, w\rangle,$$
אזי ρ תבנית בילינארית סימטרית מעל V.
- נניח ש־ V מרחב מכפלה פנימית ממשי ו־ $T \in \mathcal{L}(V)$. נגדיר $\rho \in V^{(2)}$ על ידי
$$\rho(u,w) = \langle u, Tw\rangle.$$
אזי ρ תבנית בילינארית סימטרית מעל V אם־ורק־אם T אופרטור צמוד לעצמו (הסעיף הקודם בדוגמה הוא המקרה הפרטי של $T = I$).
- נניח ש־ $\mathbf{F}$ $\rho : \mathcal{L}(V) \times \mathcal{L}(V) \to \mathbf{F}$ מוגדרת על ידי
$$\rho(S,T) = \operatorname{tr}(ST).$$
אזי ρ תבנית בילינארית סימטרית מעל $\mathcal{L}(V)$ מאחר שהעקבה היא פונקציונל לינארי מעל $\mathcal{L}(V)$ ו־ $\operatorname{tr}(ST) = \operatorname{tr}(TS)$ לכל $S, T \in \mathcal{L}(V)$; ראו 8.56.

9.11 הגדרה: מטריצה סימטרית

מטריצה ריבועית A נקראת סימטרית אם היא שווה למשוחלפת שלה.

לאופרטור מעל V ייתכן שתהיה מטריצה סימטרית ביחס לבסיס כלשהו אך לא ביחס לכל הבסיסים של V. מנגד, המסקנה הבאה מראה שלתבנית בילינארית מעל V יש מטריצה סימטרית או ביחס לכל בסיסי V או ביחס לאף לא אחד מהבסיסים של V.

9.12 תבניות ביליניאריות סימטריות הן לכסינות

נניח ש־ $\rho \in V^{(2)}$. אזי התנאים הבאים שקולים.

א. ρ תבנית בילינארית סימטרית מעל V.

ב. $\mathcal{M}(\rho, (e_1, \dots, e_n))$ מטריצה סימטרית לכל בסיס $e_1, \dots, e_n$ של V.

ג. $\mathcal{M}(\rho, (e_1, \dots, e_n))$ מטריצה סימטרית עבור בסיס כלשהו $e_1, \dots, e_n$ של V.

ד. $\mathcal{M}(\rho, (e_1, \dots, e_n))$ מטריצה אלכסונית עבור בסיס כלשהו $e_1, \dots, e_n$ של V.

Proof First suppose (a) holds, so ρ is a symmetric bilinear form. Suppose $e_1, \dots, e_n$ is a basis of V and $j, k \in \{1, \dots, n\}$. Then $\rho(e_j, e_k) = \rho(e_k, e_j)$ because ρ is symmetric. Thus $\mathcal{M}(\rho, (e_1, \dots, e_n))$ is a symmetric matrix, showing that (a) implies (b).

Clearly (b) implies (c).

Now suppose (c) holds and $e_1, \dots, e_n$ is a basis of V such that $\mathcal{M}(\rho, (e_1, \dots, e_n))$ is a symmetric matrix. Suppose $u, w \in V$. There exist $a_1, \dots, a_n, b_1, \dots, b_n \in \mathbf{F}$ such that $u = a_1 e_1 + \cdots + a_n e_n$ and $w = b_1 e_1 + \cdots + b_n e_n$. Now

$$\begin{aligned}
\rho(u, w) &= \rho\Big(\sum_{j=1}^{n} a_j e_j, \sum_{k=1}^{n} b_k e_k\Big) \\
&= \sum_{j=1}^{n} \sum_{k=1}^{n} a_j b_k \rho(e_j, e_k) \\
&= \sum_{j=1}^{n} \sum_{k=1}^{n} a_j b_k \rho(e_k, e_j) \\
&= \rho\Big(\sum_{k=1}^{n} b_k e_k, \sum_{j=1}^{n} a_j e_j\Big) \\
&= \rho(w, u),
\end{aligned}$$

where the third line holds because $\mathcal{M}(\rho)$ is a symmetric matrix. The equation above shows that ρ is a symmetric bilinear form, proving that (c) implies (a).

At this point, we have proved that (a), (b), (c) are equivalent. Because every diagonal matrix is symmetric, (d) implies (c). To complete the proof, we will show that (a) implies (d) by induction on $n = \dim V$.

If $n = 1$, then (a) implies (d) because every 1-by-1 matrix is diagonal. Now suppose $n > 1$ and the implication (a) $\implies$ (d) holds for one less dimension. Suppose (a) holds, so ρ is a symmetric bilinear form. If $\rho = 0$, then the matrix of ρ with respect to every basis of V is the zero matrix, which is a diagonal matrix. Hence we can assume that $\rho \neq 0$, which means there exist $u, w \in V$ such that $\rho(u, w) \neq 0$. Now

$$2\rho(u, w) = \rho(u + w, u + w) - \rho(u, u) - \rho(w, w).$$

Because the left side of the equation above is nonzero, the three terms on the right cannot all equal 0. Hence there exists $v \in V$ such that $\rho(v, v) \neq 0$.

Let $U = \{u \in V : \rho(u, v) = 0\}$. Thus U is the null space of the linear functional $u \mapsto \rho(u, v)$ on V. This linear functional is not the zero linear functional because $v \notin U$. Thus $\dim U = n - 1$. By our induction hypothesis, there is a basis $e_1, \dots, e_{n-1}$ of U such that the symmetric bilinear form $\rho|_{U \times U}$ has a diagonal matrix with respect to this basis.

Because $v \notin U$, the list $e_1, \dots, e_{n-1}, v$ is a basis of V. Suppose $k \in \{1, \dots, n-1\}$. Then $\rho(e_k, v) = 0$ by the construction of U. Because ρ is symmetric, we also have $\rho(v, e_k) = 0$. Thus the matrix of ρ with respect to $e_1, \dots, e_{n-1}, v$ is a diagonal matrix, completing the proof that (a) implies (d). ∎

הוכחה תחילה נניח כי (א) מתקיים, ולכן ρ תבנית בילינארית סימטרית. נניח ש־ $e_1, \dots, e_n$ בסיס של V ו־ $j, k \in \{1, \dots, n\}$. אזי $\rho(e_j, e_k) = \rho(e_k, e_j)$ מאחר ש־ ρ סימטרית. לפיכך $\mathcal{M}(\rho, (e_1, \dots, e_n))$ מטריצה סימטרית, ובכך הראינו כי (א) גורר את (ב).

(ב) גורר את (ג) בבירור.

כעת נניח כי (ג) מתקיים ו־ $e_1, \dots, e_n$ בסיס של V כך ש־ $\mathcal{M}(\rho, (e_1, \dots, e_n))$ מטריצה סימטרית. נניח ש־ $u, w \in V$. קיימים $a_1, \dots, a_n, b_1, \dots, b_n \in \mathbf{F}$ כך ש־ $u = a_1 e_1 + \cdots + a_n e_n$ ו־ $w = b_1 e_1 + \cdots + b_n e_n$. כעת

$$\begin{aligned}
\rho(u, w) &= \rho\Big(\sum_{j=1}^{n} a_j e_j, \sum_{k=1}^{n} b_k e_k\Big) \\
&= \sum_{j=1}^{n} \sum_{k=1}^{n} a_j b_k \rho(e_j, e_k) \\
&= \sum_{j=1}^{n} \sum_{k=1}^{n} a_j b_k \rho(e_k, e_j) \\
&= \rho\Big(\sum_{k=1}^{n} b_k e_k, \sum_{j=1}^{n} a_j e_j\Big) \\
&= \rho(w, u),
\end{aligned}$$

כאשר השורה השלישית מתקיימת מאחר ש־ $\mathcal{M}(\rho)$ מטריצה סימטרית. המשוואה לעיל מראה ש־ ρ תבנית בילינארית סימטרית, ובכך הראינו כי (ג) גורר את (א).

בנקודה זו, הוכחנו כי (א), (ב) ו־ (ג) שקולים. מאחר שכל מטריצה אלכסונית היא סימטרית, (ד) גורר את (ג). להשלמת ההוכחה, נראה כי (א) גורר את (ד) על ידי אינדוקציה על $n = \dim V$.

אם $n = 1$, אזי (א) גורר את (ד) מאחר שכל מטריצה 1×1 היא אלכסונית. כעת נניח ש־ $n > 1$ והגרירה (א) $\implies$ (ד) מתקיימת עבור ממד נמוך באחד. נניח ש־ (א) מתקיים, ולכן ρ תבנית בילינארית סימטרית. אם $\rho = 0$, אזי המטריצה של ρ ביחס אל כל בסיס של V היא מטריצת האפס, שהיא מטריצה אלכסונית. לכן אנו יכולים להניח ש־ $\rho \neq 0$, כלומר קיימים $u, w \in V$ כך ש־ $\rho(u, w) \neq 0$. כעת

$$2\rho(u, w) = \rho(u + w, u + w) - \rho(u, u) - \rho(w, w).$$

מאחר שהאגף השמאלי של המשוואה לעיל שונה מאפס, שלושת הביטויים באגף הימני לא יכולים להיות כולם 0. לכן קיים $v \in V$ כך ש־ $\rho(v, v) \neq 0$.

יהי $U = \{u \in V : \rho(u, v) = 0\}$. לפיכך U הוא מרחב האפס של הפונקציונל הליניארי $u \mapsto \rho(u, v)$ מעל V. פונקציונל ליניארי זה אינו פונקציונל האפס מאחר ש־ $v \notin U$. לפיכך $\dim U = n - 1$. לפי הנחת האינדוקציה, קיים בסיס $e_1, \dots, e_{n-1}$ של U כך שיש לתבנית הבילינארית הסימטרית $\rho|_{U \times U}$ מטריצה אלכסונית ביחס לבסיס זה.

מאחר ש־ $v \notin U$, הרשימה $e_1, \dots, e_{n-1}, v$ בסיס של V. נניח ש־ $k \in \{1, \dots, n-1\}$. אזי $\rho(e_k, v) = 0$ מחמת אופן בניית U. מאחר ש־ ρ סימטרית, נקבל גם ש־ $\rho(v, e_k) = 0$. לפיכך המטריצה של ρ ביחס ל־ $e_1, \dots, e_{n-1}, v$ מטריצה אלכסונית ובכך הושלמה ההוכחה כי (א) גורר את (ד). ■

The previous result states that every symmetric bilinear form has a diagonal matrix with respect to some basis. If our vector space happens to be a real inner product space, then the next result shows that every symmetric bilinear form has a diagonal matrix with respect to some *orthonormal* basis. Note that the inner product here is unrelated to the bilinear form.

9.13 *diagonalization of a symmetric bilinear form by an orthonormal basis*

Suppose V is a real inner product space and ρ is a symmetric bilinear form on V. Then ρ has a diagonal matrix with respect to some orthonormal basis of V.

Proof Let $f_1, \dots, f_n$ be an orthonormal basis of V. Let $B = \mathcal{M}(\rho, (f_1, \dots, f_n))$. Then B is a symmetric matrix (by 9.12). Let $T \in \mathcal{L}(V)$ be the operator such that $\mathcal{M}(T, (f_1, \dots, f_n)) = B$. Thus T is self-adjoint.

The real spectral theorem (7.29) states that T has a diagonal matrix with respect to some orthonormal basis $e_1, \dots, e_n$ of V. Let $C = \mathcal{M}(I, (e_1, \dots, e_n), (f_1, \dots, f_n))$. Thus $C^{-1}BC$ is the matrix of T with respect to the basis $e_1, \dots, e_n$ (by 3.84). Hence $C^{-1}BC$ is a diagonal matrix. Now

$$M(\rho, (e_1, \dots, e_n)) = C^t BC = C^{-1}BC,$$

where the first equality holds by 9.7 and the second equality holds because C is a unitary matrix with real entries (which implies that $C^{-1} = C^t$; see 7.57). ∎

Now we turn our attention to alternating bilinear forms. Alternating multilinear forms will play a major role in our approach to determinants later in this chapter.

9.14 definition: *alternating bilinear form*, $V_{\text{alt}}^{(2)}$

A bilinear form $\alpha \in V^{(2)}$ is called *alternating* if

$$\alpha(v, v) = 0$$

for all $v \in V$. The set of alternating bilinear forms on V is denoted by $V_{\text{alt}}^{(2)}$.

9.15 example: *alternating bilinear forms*

- Suppose $n \geq 3$ and $\alpha : \mathbf{F}^n \times \mathbf{F}^n \to \mathbf{F}$ is defined by

$$\alpha((x_1, \dots, x_n), (y_1, \dots, y_n)) = x_1y_2 - x_2y_1 + x_1y_3 - x_3y_1.$$

  Then α is an alternating bilinear form on $\mathbf{F}^n$.

- Suppose $\varphi, \tau \in V'$. Then the bilinear form α on V defined by

$$\alpha(u, w) = \varphi(u)\tau(w) - \varphi(w)\tau(u)$$

  is alternating.

המסקנה הקודמת קובעת כי לכל תבנית בילינארית סימטרית יש מטריצה אלכסונית ביחס לבסיס כלשהו. אם המרחב הווקטורי הוא במקרה גם מרחב מכפלה פנימית ממשי, אזי המסקנה הבאה מראה שלכל תבנית בילינארית סימטרית יש מטריצה אלכסונית ביחס לבסיס אורתונורמלי כלשהו. שימו לב שהמכפלה הפנימית כאן אינה קשורה לתבנית הבילינארית.

9.13 לכסון של תבנית בילינארית סימטרית על ידי בסיס אורתונורמלי

נניח ש־ V מרחב מכפלה פנימית ממשי ו־ ρ תבנית בילינארית סימטרית מעל V. אזי יש ל־ ρ מטריצה אלכסונית ביחס לבסיס אורתונורמלי כלשהו של V.

הוכחה תהי $f_1, \dots, f_n$ בסיס אורתונורמלי של V. תהי $B = \mathcal{M}(\rho, (f_1, \dots, f_n))$. אזי B היא מטריצה סימטרית (לפי 9.12). יהי $T \in \mathcal{L}(V)$ האופרטור כך ש־ $\mathcal{M}(T, (f_1, \dots, f_n)) = B$. לפיכך T צמוד לעצמו.

משפט הפירוק הספקטרלי הממשי (7.29) קובע כי יש ל־ T מטריצה אלכסונית ביחס לבסיס אורתונורמלי $e_1, \dots, e_n$ כלשהו של V. תהי $C = \mathcal{M}(I, (e_1, \dots, e_n), (f_1, \dots, f_n))$. לפיכך $C^{-1}BC$ היא המטריצה של T ביחס לבסיס $e_1, \dots, e_n$ (לפי 3.84). לכן $C^{-1}BC$ מטריצה אלכסונית. כעת

$$\mathcal{M}(\rho, (e_1, \dots, e_n)) = C^t BC = C^{-1}BC,$$

כאשר השוויון הראשון מתקיים לפי 9.7 והשוויון השני מתקיים מאחר ש־ C מטריצה אוניטרית עם איברים ממשיים (מה שגורר ש־ $C^{-1} = C^t$; ראו 7.57). ■

כעת נעבור להתמקד בתבניות בילינאריות מתחלפות. לתבניות מולטי־לינאריות מתחלפות יהיה תפקיד מרכזי בגישתנו לדטרמיננטות בהמשך פרק זה.

9.14 הגדרה: תבנית בילינארית מתחלפת, $V_{\text{alt}}^{(2)}$

תבנית בילינארית $\alpha \in V^{(2)}$ נקראת מתחלפת אם

$$\alpha(v, v) = 0$$

לכל $v \in V$. קבוצת התבניות הבילינאריות המתחלפות מעל V מסומנת על ידי $V_{\text{alt}}^{(2)}$.

9.15 דוגמה: תבניות בילינאריות מתחלפות

- נניח ש־ $n \geq 3$ ו־ $\alpha : \mathbf{F}^n \times \mathbf{F}^n \to \mathbf{F}$ מוגדרת על ידי
$$\alpha((x_1, \dots, x_n), (y_1, \dots, y_n)) = x_1y_2 - x_2y_1 + x_1y_3 - x_3y_1.$$
אזי α תבנית בילינארית מתחלפת מעל $\mathbf{F}^n$.

- נניח ש־ $\varphi, \tau \in V'$. אזי התבנית הבילינארית α מעל V המוגדרת על ידי
$$\alpha(u, w) = \varphi(u)\tau(w) - \varphi(w)\tau(u)$$
היא מתחלפת.

The next result shows that a bilinear form is alternating if and only if switching the order of the two inputs multiplies the output by -1.

9.16 *characterization of alternating bilinear forms*

A bilinear form α on V is alternating if and only if

$$\alpha(u, w) = -\alpha(w, u)$$

for all $u, w \in V$.

Proof First suppose that α is alternating. If $u, w \in V$, then

$$\begin{aligned} 0 &= \alpha(u + w, u + w) \\ &= \alpha(u, u) + \alpha(u, w) + \alpha(w, u) + \alpha(w, w) \\ &= \alpha(u, w) + \alpha(w, u). \end{aligned}$$

Thus $\alpha(u, w) = -\alpha(w, u)$, as desired.

To prove the implication in the other direction, suppose $\alpha(u, w) = -\alpha(w, u)$ for all $u, w \in V$. Then $\alpha(v, v) = -\alpha(v, v)$ for all $v \in V$, which implies that $\alpha(v, v) = 0$ for all $v \in V$. Thus α is alternating. ∎

Now we show that the vector space of bilinear forms on V is the direct sum of the symmetric bilinear forms on V and the alternating bilinear forms on V.

9.17 $V^{(2)} = V^{(2)}_{\text{sym}} \oplus V^{(2)}_{\text{alt}}$

The sets $V^{(2)}_{\text{sym}}$ and $V^{(2)}_{\text{alt}}$ are subspaces of $V^{(2)}$. Furthermore,

$$V^{(2)} = V^{(2)}_{\text{sym}} \oplus V^{(2)}_{\text{alt}}.$$

Proof The definition of symmetric bilinear form implies that the sum of any two symmetric bilinear forms on V is a symmetric bilinear form on V, and every scalar multiple of any symmetric bilinear form on V is a symmetric bilinear form on V. Also, the zero bilinear form is in $V^{(2)}_{\text{sym}}$. Thus $V^{(2)}_{\text{sym}}$ is a subspace of $V^{(2)}$. Similarly, the verification that $V^{(2)}_{\text{alt}}$ is a subspace of $V^{(2)}$ is straightforward.

Next, we want to show that $V^{(2)} = V^{(2)}_{\text{sym}} + V^{(2)}_{\text{alt}}$. To do this, suppose $\beta \in V^{(2)}$. Define $\rho, \alpha \in V^{(2)}$ by

$$\rho(u, w) = \frac{\beta(u, w) + \beta(w, u)}{2} \quad \text{and} \quad \alpha(u, w) = \frac{\beta(u, w) - \beta(w, u)}{2}.$$

Then $\rho \in V^{(2)}_{\text{sym}}$ and $\alpha \in V^{(2)}_{\text{alt}}$, and $\beta = \rho + \alpha$. Thus $V^{(2)} = V^{(2)}_{\text{sym}} + V^{(2)}_{\text{alt}}$.

Finally, to show that the intersection of the two subspaces under consideration equals $\{0\}$, suppose $\beta \in V^{(2)}_{\text{sym}} \cap V^{(2)}_{\text{alt}}$. If $u, w \in V$, then 9.16 implies that

$$\beta(u, w) = -\beta(w, u) = -\beta(u, w)$$

and hence $\beta(u, w) = 0$. Thus $\beta = 0$. Hence $V^{(2)} = V^{(2)}_{\text{sym}} \oplus V^{(2)}_{\text{alt}}$ (by 1.46). ∎

המסקנה הבאה מראה שתבנית בילינארית היא מתחלפת אם־ורק־אם שינוי סדר הרכיבים מכפיל את התוצאה ב־ -1.

9.16 אפיון תבניות בילינאריות מתחלפות

תבנית בילינארית α מעל V היא מתחלפת אם־ורק־אם

$$\alpha(u,w) = -\alpha(w,u)$$

לכל $u, w \in V$.

הוכחה תחילה נניח כי α מתחלפת. אם $u, w \in V$, אזי

$$\begin{aligned} 0 &= \alpha(u+w, u+w) \\ &= \alpha(u,u) + \alpha(u,w) + \alpha(w,u) + \alpha(w,w) \\ &= \alpha(u,w) + \alpha(w,u). \end{aligned}$$

לפיכך $\alpha(u,w) = -\alpha(w,u)$, כפי שרצינו.

להוכחת הגרירה בכיוון השני, נניח ש־ $\alpha(u,w) = -\alpha(w,u)$ לכל $u, w \in V$. אזי $\alpha(v,v) = -\alpha(v,v)$ לכל $v \in V$, מה שגורר ש־ $\alpha(v,v) = 0$ לכל $v \in V$. לפיכך α מתחלפת. ■

כעת נראה כי המרחב הווקטורי של תבניות בילינאריות מעל V הוא סכום ישר של התבניות הבילינאריות הסימטריות מעל V והתבניות הבילינאריות המתחלפות מעל V.

9.17 $V^{(2)} = V^{(2)}_{\text{sym}} \oplus V^{(2)}_{\text{alt}}$

הקבוצות $V^{(2)}_{\text{sym}}$ ו־ $V^{(2)}_{\text{alt}}$ הן תתי־מרחב של $V^{(2)}$. יתר על כן,

$$V^{(2)} = V^{(2)}_{\text{sym}} \oplus V^{(2)}_{\text{alt}}.$$

הוכחה ההגדרה של תבנית בילינארית סימטרית גוררת שהסכום של כל שתי תבניות בילינאריות סימטריות מעל V הוא תבנית בילינארית סימטרית מעל V, וכל כפולת סקלר של תבנית בילינארית סימטרית כלשהי מעל V היא תבנית בילינארית סימטרית מעל V. בנוסף, התבנית הבילינארית האפס נמצאת ב־ $V^{(2)}_{\text{sym}}$. לפיכך $V^{(2)}_{\text{sym}}$ הוא תת־מרחב של $V^{(2)}$. באופן דומה, הווידוא ש־ $V^{(2)}_{\text{alt}}$ הוא תת־מרחב של $V^{(2)}$ הוא ישיר.

כעת, ברצוננו להראות ש־ $V^{(2)} = V^{(2)}_{\text{sym}} + V^{(2)}_{\text{alt}}$. לצורך כך, נניח ש־ $\beta \in V^{(2)}$. נגדיר $\rho, \alpha \in V^{(2)}$ על ידי

$$\alpha(u,w) = \frac{\beta(u,w) - \beta(w,u)}{2} \quad \text{וגם} \quad \rho(u,w) = \frac{\beta(u,w) + \beta(w,u)}{2}.$$

אזי $\rho \in V^{(2)}_{\text{sym}}$ ו־ $\alpha \in V^{(2)}_{\text{alt}}$, וגם $\beta = \rho + \alpha$. לפיכך $V^{(2)} = V^{(2)}_{\text{sym}} + V^{(2)}_{\text{alt}}$.

לבסוף, כדי להראות שהחיתוך של שני תתי־המרחב הנדונים שווה ל־ $\{0\}$, נניח ש־ $\beta \in V^{(2)}_{\text{sym}} \cap V^{(2)}_{\text{alt}}$. אם $u, w \in V$, אזי 9.16 גורר ש־

$$\beta(u,w) = -\beta(w,u) = -\beta(u,w)$$

ולכן $\beta(u,w) = 0$. לפיכך $\beta = 0$. לכן $V^{(2)} = V^{(2)}_{\text{sym}} \oplus V^{(2)}_{\text{alt}}$ (לפי 1.46). ■

Quadratic Forms

9.18 definition: *quadratic form associated with a bilinear form, q_β*

For β a bilinear form on V, define a function $q_\beta : V \to \mathbf{F}$ by $q_\beta(v) = \beta(v, v)$. A function $q : V \to \mathbf{F}$ is called a *quadratic form* on V if there exists a bilinear form β on V such that $q = q_\beta$.

Note that if β is a bilinear form, then $q_\beta = 0$ if and only if β is alternating.

9.19 example: *quadratic form*

Suppose β is the bilinear form on $\mathbf{R}^3$ defined by

$$\beta((x_1, x_2, x_3), (y_1, y_2, y_3)) = x_1y_1 - 4x_1y_2 + 8x_1y_3 - 3x_3y_3.$$

Then q_β is the quadratic form on $\mathbf{R}^3$ given by the formula

$$q_\beta(x_1, x_2, x_3) = x_1^2 - 4x_1x_2 + 8x_1x_3 - 3x_3^2.$$

The quadratic form in the example above is typical of quadratic forms on $\mathbf{F}^n$, as shown in the next result.

9.20 *quadratic forms on $\mathbf{F}^n$*

Suppose n is a positive integer and q is a function from $\mathbf{F}^n$ to $\mathbf{F}$. Then q is a quadratic form on $\mathbf{F}^n$ if and only if there exist numbers $A_{j,k} \in \mathbf{F}$ for $j, k \in \{1, \dots, n\}$ such that

$$q(x_1, \dots, x_n) = \sum_{k=1}^{n} \sum_{j=1}^{n} A_{j,k} x_j x_k$$

for all $(x_1, \dots, x_n) \in \mathbf{F}^n$.

Proof First suppose q is a quadratic form on $\mathbf{F}^n$. Thus there exists a bilinear form β on $\mathbf{F}^n$ such that $q = q_\beta$. Let A be the matrix of β with respect to the standard basis of $\mathbf{F}^n$. Then for all $(x_1, \dots, x_n) \in \mathbf{F}^n$, we have the desired equation

$$q(x_1, \dots, x_n) = \beta((x_1, \dots, x_n), (x_1, \dots, x_n)) = \sum_{k=1}^{n} \sum_{j=1}^{n} A_{j,k} x_j x_k.$$

Conversely, suppose there exist numbers $A_{j,k} \in \mathbf{F}$ for $j, k \in \{1, \dots, n\}$ such that

$$q(x_1, \dots, x_n) = \sum_{k=1}^{n} \sum_{j=1}^{n} A_{j,k} x_j x_k$$

for all $(x_1, \dots, x_n) \in \mathbf{F}^n$. Define a bilinear form β on $\mathbf{F}^n$ by

$$\beta((x_1, \dots, x_n), (y_1, \dots, y_n)) = \sum_{k=1}^{n} \sum_{j=1}^{n} A_{j,k} x_j y_k.$$

Then $q = q_\beta$, as desired. ∎

תבניות ריבועיות

9.18 הגדרה: תבנית ריבועית המתאימה לתבנית בילינארית, q_β

עבור β תבנית בילינארית מעל V, נגדיר פונקציה $q_\beta : V \to \mathbf{F}$ על ידי $q_\beta(v) = \beta(v,v)$. פונקציה $q : V \to \mathbf{F}$ נקראת תבנית ריבועית מעל V אם קיימת תבנית בילינארית β מעל V כך ש־ $q = q_\beta$.

נשים לב כי אם β תבנית בילינארית, אזי $q_\beta = 0$ אם־ורק־אם β מתחלפת.

9.19 דוגמה: תבנית ריבועית

נניח ש־ β היא התבנית הבילינארית מעל $\mathbf{R}^3$ המוגדרת על ידי

$$\beta((x_1,x_2,x_3),(y_1,y_2,y_3)) = x_1y_1 - 4x_1y_2 + 8x_1y_3 - 3x_3y_3.$$

אזי q_β התבנית הריבועית מעל $\mathbf{R}^3$ הניתנת באמצעות הנוסחה

$$q_\beta(x_1,x_2,x_3) = x_1^2 - 4x_1x_2 + 8x_1x_3 - 3x_3^2.$$

התבנית הריבועית בדוגמה לעיל היא תבנית טיפוסית לתבניות ריבועיות מעל $\mathbf{F}^n$, כפי שנראה במסקנה הבאה.

9.20 תבניות ריבועיות מעל $\mathbf{F}^n$

נניח ש־ n מספר שלם חיובי ו־ q פונקציה מ־ $\mathbf{F}^n$ אל $\mathbf{F}$. אזי q תבנית ריבועית מעל $\mathbf{F}^n$ אם־ורק־אם קיימים מספרים $A_{j,k} \in \mathbf{F}$ עבור $j,k \in \{1,\dots,n\}$ כך ש־

$$q(x_1,\dots,x_n) = \sum_{k=1}^{n}\sum_{j=1}^{n} A_{j,k}x_jx_k$$

לכל $(x_1,\dots,x_n) \in \mathbf{F}^n$.

הוכחה תחילה נניח כי q תבנית ריבועית מעל $\mathbf{F}^n$. לפיכך קיימת תבנית בילינארית β מעל $\mathbf{F}^n$ כך ש־ $q = q_\beta$. תהי A המטריצה של β ביחס לבסיס הסטנדרטי של $\mathbf{F}^n$. אזי לכל $(x_1,\dots,x_n) \in \mathbf{F}^n$, מתקיימת המשוואה המבוקשת

$$q(x_1,\dots,x_n) = \beta((x_1,\dots,x_n),(x_1,\dots,x_n)) = \sum_{k=1}^{n}\sum_{j=1}^{n} A_{j,k}x_jx_k.$$

בכיוון ההפוך, נניח כי קיימים מספרים $A_{j,k} \in \mathbf{F}$ עבור $j,k \in \{1,\dots,n\}$ כך ש־

$$q(x_1,\dots,x_n) = \sum_{k=1}^{n}\sum_{j=1}^{n} A_{j,k}x_jx_k$$

לכל $(x_1,\dots,x_n) \in \mathbf{F}^n$. נגדיר תבנית בילינארית β מעל $\mathbf{F}^n$ על ידי

$$\beta((x_1,\dots,x_n),(y_1,\dots,y_n)) = \sum_{k=1}^{n}\sum_{j=1}^{n} A_{j,k}x_jy_k.$$

אזי $q = q_\beta$, כפי שרצינו. ■

Although quadratic forms are defined in terms of an arbitrary bilinear form, the equivalence of (a) and (b) in the result below shows that a *symmetric* bilinear form can always be used.

9.21 *characterizations of quadratic forms*

Suppose $q: V \to \mathbf{F}$ is a function. The following are equivalent.

(a) q is a quadratic form.

(b) There exists a unique symmetric bilinear form ρ on V such that $q = q_\rho$.

(c) $q(\lambda v) = \lambda^2 q(v)$ for all $\lambda \in \mathbf{F}$ and all $v \in V$, and the function

$$(u, w) \mapsto q(u + w) - q(u) - q(w)$$

is a symmetric bilinear form on V.

(d) $q(2v) = 4q(v)$ for all $v \in V$, and the function

$$(u, w) \mapsto q(u + w) - q(u) - q(w)$$

is a symmetric bilinear form on V.

Proof First suppose (a) holds, so q is a quadratic form. Hence there exists a bilinear form β such that $q = q_\beta$. By 9.17, there exist a symmetric bilinear form ρ on V and an alternating bilinear form α on V such that $\beta = \rho + \alpha$. Now

$$q = q_\beta = q_\rho + q_\alpha = q_\rho.$$

If $\rho' \in V_{\text{sym}}^{(2)}$ also satisfies $q_{\rho'} = q$, then $q_{\rho'-\rho} = 0$; thus $\rho' - \rho \in V_{\text{sym}}^{(2)} \cap V_{\text{alt}}^{(2)}$, which implies that $\rho' = \rho$ (by 9.17). This completes the proof that (a) implies (b).

Now suppose (b) holds, so there exists a symmetric bilinear form ρ on V such that $q = q_\rho$. If $\lambda \in \mathbf{F}$ and $v \in V$ then

$$q(\lambda v) = \rho(\lambda v, \lambda v) = \lambda\rho(v, \lambda v) = \lambda^2 \rho(v, v) = \lambda^2 q(v),$$

showing that the first part of (c) holds.

If $u, w \in V$, then

$$q(u + w) - q(u) - q(w) = \rho(u + w, u + w) - \rho(u, u) - \rho(w, w) = 2\rho(u, w).$$

Thus the function $(u, w) \mapsto q(u+w)-q(u)-q(w)$ equals 2ρ, which is a symmetric bilinear form on V, completing the proof that (b) implies (c).

Clearly (c) implies (d).

Now suppose (d) holds. Let ρ be the symmetric bilinear form on V defined by

$$\rho(u, w) = \frac{q(u + w) - q(u) - q(w)}{2}.$$

If $v \in V$, then

$$\rho(v, v) = \frac{q(2v) - q(v) - q(v)}{2} = \frac{4q(v) - 2q(v)}{2} = q(v).$$

Thus $q = q_\rho$, completing the proof that (d) implies (a). ∎

אף שתבניות ריבועיות מוגדרות במונחים של תבנית בילינארית שרירותית, השקילות של (א) ו־ (ב) במסקנה להלן מראה לנו כי תמיד נוכל להשתמש בתבנית בילינארית סימטרית.

9.21 אפיונים של תבניות ריבועיות

נניח ש־ $q: V \to \mathbf{F}$ פונקציה. התנאים הבאים שקולים.

א. q תבנית ריבועית.

ב. קיימת תבנית בילינארית סימטרית יחידה ρ מעל V כך ש־ $q = q_\rho$.

ג. $q(\lambda v) = \lambda^2 q(v)$ לכל $\lambda \in \mathbf{F}$ ולכל $v \in V$, והפונקציה
$(u, w) \mapsto q(u + w) - q(u) - q(w)$
היא תבנית בילינארית סימטרית מעל V.

ד. $q(2v) = 4q(v)$ לכל $v \in V$, והפונקציה
$(u, w) \mapsto q(u + w) - q(u) - q(w)$
תבנית בילינארית סימטרית מעל V.

הוכחה תחילה נניח כי (א) מתקיים, כך ש־ q תבנית ריבועית. לכן קיימת תבנית בילינארית β כך ש־ $q = q_\beta$. לפי 9.17, קיימות תבנית בילינארית סימטרית ρ מעל V ותבנית בילינארית מתחלפת α מעל V כך ש־ $\beta = \rho + \alpha$. כעת

$$q = q_\beta = q_\rho + q_\alpha = q_\rho.$$

אם גם $\rho' \in V^{(2)}_{\text{sym}}$ מקיימת $q_{\rho'} = q$, אזי $q_{\rho'-\rho} = 0$; לפיכך $\rho' - \rho \in V^{(2)}_{\text{sym}} \cap V^{(2)}_{\text{alt}}$, מה שגורר ש־ $\rho = \rho'$ (לפי 9.17). בכך הושלמה ההוכחה כי (א) גורר את (ב).

כעת נניח כי (ב) מתקיים, ולכן קיימת תבנית בילינארית סימטרית ρ מעל V כך ש־ $q = q_\rho$. אם $\lambda \in \mathbf{F}$ ו־ $v \in V$ אזי

$$q(\lambda v) = \rho(\lambda v, \lambda v) = \lambda \rho(v, \lambda v) = \lambda^2 \rho(v, v) = \lambda^2 q(v),$$

ובכך הראינו כי החלק הראשון של (ג) מתקיים.

אם $u, w \in V$, אזי

$$q(u + w) - q(u) - q(w) = \rho(u + w, u + w) - \rho(u, u) - \rho(w, w) = 2\rho(u, w).$$

לפיכך הפונקציה $(u, w) \mapsto q(u + w) - q(u) - q(w)$ שווה ל־ 2ρ, שהיא תבנית בילינארית סימטרית מעל V, ובכך השלמנו את ההוכחה כי (ב) גורר את (ג).

(ג) גורר את (ד) בבירור.

כעת נניח כי (ד) מתקיים. תהי ρ התבנית הבילינארית הסימטרית מעל V המוגדרת על ידי

$$\rho(u, w) = \frac{q(u + w) - q(u) - q(w)}{2}.$$

אם $v \in V$, אזי

$$\rho(v, v) = \frac{q(2v) - q(v) - q(v)}{2} = \frac{4q(v) - 2q(v)}{2} = q(v).$$

לפיכך $q = q_\rho$, ובכך השלמנו את ההוכחה כי (ד) גורר את (א).

9.22 example: *symmetric bilinear form associated with a quadratic form*

Suppose q is the quadratic form on $\mathbf{R}^3$ given by the formula

$$q(x_1, x_2, x_3) = x_1^2 - 4x_1x_2 + 8x_1x_3 - 3x_3^2.$$

A bilinear form β on $\mathbf{R}^3$ such that $q = q_\beta$ is given by Example 9.19, but this bilinear form is not symmetric, as promised by 9.21(b). However, the bilinear form ρ on $\mathbf{R}^3$ defined by

$$\rho((x_1, x_2, x_3), (y_1, y_2, y_3)) = x_1y_1 - 2x_1y_2 - 2x_2y_1 + 4x_1y_3 + 4x_3y_1 - 3x_3y_3$$

is symmetric and satisfies $q = q_\rho$.

The next result states that for each quadratic form we can choose a basis such that the quadratic form looks like a weighted sum of squares of the coordinates, meaning that there are no cross terms of the form x_jx_k with $j \neq k$.

9.23 *diagonalization of quadratic form*

Suppose q is a quadratic form on V.

(a) There exist a basis $e_1, \dots, e_n$ of V and $\lambda_1, \dots, \lambda_n \in \mathbf{F}$ such that

$$q(x_1e_1 + \cdots + x_ne_n) = \lambda_1x_1^2 + \cdots + \lambda_nx_n^2$$

for all $x_1, \dots, x_n \in \mathbf{F}$.

(b) If $\mathbf{F} = \mathbf{R}$ and V is an inner product space, then the basis in (a) can be chosen to be an orthonormal basis of V.

Proof

(a) There exists a symmetric bilinear form ρ on V such that $q = q_\rho$ (by 9.21). Now there exists a basis $e_1, \dots, e_n$ of V such that $\mathcal{M}(\rho, (e_1, \dots, e_n))$ is a diagonal matrix (by 9.12). Let $\lambda_1, \dots, \lambda_n$ denote the entries on the diagonal of this matrix. Thus

$$\rho(e_j, e_k) = \begin{cases} \lambda_j & \text{if } j = k, \\ 0 & \text{if } j \neq k \end{cases}$$

for all $j, k \in \{1, \dots, n\}$. If $x_1, \dots, x_n \in \mathbf{F}$, then

$$\begin{aligned} q(x_1e_1 + \cdots + x_ne_n) &= \rho(x_1e_1 + \cdots + x_ne_n, x_1e_1 + \cdots + x_ne_n) \\ &= \sum_{k=1}^{n} \sum_{j=1}^{n} x_jx_k\rho(e_j, e_k) \\ &= \lambda_1x_1^2 + \cdots + \lambda_nx_n^2, \end{aligned}$$

as desired.

(b) Suppose $\mathbf{F} = \mathbf{R}$ and V is an inner product space. Then 9.13 tells us that the basis in (a) can be chosen to be an orthonormal basis of V. ∎

9.22 דוגמה: תבנית בילינאריות סימטרית המתאימה לתבנית ריבועית

נניח ש־ q התבנית הריבועית מעל $\mathbf{R}^3$ הניתנת באמצעות הנוסחה

$$q(x_1, x_2, x_3) = x_1^2 - 4x_1x_2 + 8x_1x_3 - 3x_3^2.$$

תבנית בילינארית β מעל $\mathbf{R}^3$ כך ש־ $q = q_\beta$ נתונה על ידי דוגמה 9.19, אולם תבנית בילינארית זו אינה סימטרית, כמובטח על ידי 9.21(ב). בכל זאת, התבנית הבילינארית ρ מעל $\mathbf{R}^3$ המוגדרת על ידי

$$\rho((x_1, x_2, x_3), (y_1, y_2, y_3)) = x_1y_1 - 2x_1y_2 - 2x_2y_1 + 4x_1y_3 + 4x_3y_1 - 3x_3y_3$$

היא סימטרית ומקיימת $q = q_\rho$.

המסקנה הבאה קובעת שעבור כל תבנית ריבועית נוכל לבחור בסיס כך שהתבנית הריבועית תיראה כסכום משוקלל של ריבועי הקואורדינטות, כלומר אין איברים מעורבים מהצורה x_jx_k עם $j \neq k$.

9.23 לכסון של תבנית ריבועית

נניח ש־ q תבנית ריבועית מעל V.

א. קיימים בסיס $e_1, \dots, e_n$ של V וסקלרים $\lambda_1, \dots, \lambda_n \in \mathbf{F}$ כך ש־

$$q(x_1e_1 + \cdots + x_ne_n) = \lambda_1x_1^2 + \cdots + \lambda_nx_n^2$$

לכל $x_1, \dots, x_n \in \mathbf{F}$.

ב. אם $\mathbf{F} = \mathbf{R}$ ו־ V מרחב מכפלה פנימית, אזי נוכל לבחור את הבסיס בסעיף (א) כבסיס אורתונורמלי של V.

הוכחה

א. קיימת תבנית בילינארית סימטרית ρ מעל V כך ש־ $q = q_\rho$ (לפי 9.21). כעת קיים בסיס $e_1, \dots, e_n$ של V כך ש־ $\mathcal{M}(\rho, (e_1, \dots, e_n))$ מטריצה אלכסונית (לפי 9.12). נסמן ב־ $\lambda_1, \dots, \lambda_n$ את איברי האלכסון של מטריצה זו. לפיכך

$$\rho(e_j, e_k) = \begin{cases} \lambda_j & \text{if } j = k, \\ 0 & \text{if } j \neq k \end{cases}$$

לכל $j, k \in \{1, \dots, n\}$. אם $x_1, \dots, x_n \in \mathbf{F}$, אזי

$$\begin{aligned} q(x_1e_1 + \cdots + x_ne_n) &= \rho(x_1e_1 + \cdots + x_ne_n, x_1e_1 + \cdots + x_ne_n) \\ &= \sum_{k=1}^{n} \sum_{j=1}^{n} x_jx_k\rho(e_j, e_k) \\ &= \lambda_1x_1^2 + \cdots + \lambda_nx_n^2, \end{aligned}$$

כפי שרצינו.

ב. נניח ש־ $\mathbf{F} = \mathbf{R}$ ו־ V מרחב מכפלה פנימית. אזי 9.13 מורה לנו כי נוכל לבחור את הבסיס ב־ (א) כבסיס אורתונורמלי של V. ■

Exercises 9A

1 Prove that if β is a bilinear form on $\mathbf{F}$, then there exists $c \in \mathbf{F}$ such that

$$\beta(x, y) = cxy$$

for all $x, y \in \mathbf{F}$.

2 Let $n = \dim V$. Suppose β is a bilinear form on V. Prove that there exist $\varphi_1, \dots, \varphi_n, \tau_1, \dots, \tau_n \in V'$ such that

$$\beta(u, v) = \varphi_1(u) \cdot \tau_1(v) + \cdots + \varphi_n(u) \cdot \tau_n(v)$$

for all $u, v \in V$.

This exercise shows that if $n = \dim V$, then every bilinear form on V is of the form given by the last bullet point of Example 9.2.

3 Suppose $\beta : V \times V \to \mathbf{F}$ is a bilinear form on V and also is a linear functional on $V \times V$. Prove that $\beta = 0$.

4 Suppose V is a real inner product space and β is a bilinear form on V. Show that there exists a unique operator $T \in \mathcal{L}(V)$ such that

$$\beta(u, v) = \langle u, Tv \rangle$$

for all $u, v \in V$.

This exercise states that if V is a real inner product space, then every bilinear form on V is of the form given by the third bullet point in 9.2.

5 Suppose β is a bilinear form on a real inner product space V and T is the unique operator on V such that $\beta(u, v) = \langle u, Tv \rangle$ for all $u, v \in V$ (see Exercise 4). Show that β is an inner product on V if and only if T is an invertible positive operator on V.

6 Prove or give a counterexample: If ρ is a symmetric bilinear form on V, then

$$\{v \in V : \rho(v, v) = 0\}$$

is a subspace of V.

7 Explain why the proof of 9.13 (diagonalization of a symmetric bilinear form by an orthonormal basis on a real inner product space) fails if the hypothesis that $\mathbf{F} = \mathbf{R}$ is dropped.

8 Find formulas for $\dim V^{(2)}_{\text{sym}}$ and $\dim V^{(2)}_{\text{alt}}$ in terms of $\dim V$.

9 Suppose that n is a positive integer and $V = \{p \in \mathcal{P}_n(\mathbf{R}) : p(0) = p(1)\}$. Define $\alpha : V \times V \to \mathbf{R}$ by

$$\alpha(p, q) = \int_0^1 pq'.$$

Show that α is an alternating bilinear form on V.

תרגילים טו

1 הוכיחו כי אם β תבנית בילינארית מעל $\mathbf{F}$, אזי קיים $c \in \mathbf{F}$ כך ש־
$$\beta(x,y) = cxy$$
לכל $x, y \in \mathbf{F}$.

2 יהי $n = \dim V$. נניח ש־ β תבנית בילינארית מעל V. הוכיחו כי קיימים $\varphi_1, \dots, \varphi_n, \tau_1, \dots, \tau_n \in V'$ כך ש־
$$\beta(u,v) = \varphi_1(u) \cdot \tau_1(v) + \cdots + \varphi_n(u) \cdot \tau_n(v)$$
לכל $u, v \in V$.

תרגיל זה מראה כי אם $n = \dim V$, אזי כל תבנית בילינארית מעל V היא מהצורה שניתנה בסעיף האחרון של דוגמה 9.2.

3 נניח ש־ $\beta : V \times V \to \mathbf{F}$ תבנית בילינארית מעל V ובנוסף פונקציונל לינארי מעל $V \times V$. הוכיחו כי $\beta = 0$.

4 נניח ש־ V מרחב מכפלה פנימית ממשי ו־ β תבנית בילינארית מעל V. הראו כי קיים אופרטור יחיד $T \in \mathcal{L}(V)$ כך ש־
$$\beta(u,v) = \langle u, Tv \rangle$$
לכל $u, v \in V$.

תרגיל זה קובע כי אם V מרחב מכפלה פנימית ממשי, אזי כל תבנית בילינארית מעל V היא מהצורה שניתנה בסעיף השלישי של דוגמה 9.2.

5 נניח ש־ β תבנית בילינארית מעל מרחב מכפלה פנימית ממשי V ו־ T הוא האופרטור היחיד מעל V כך ש־ $\beta(u,v) = \langle u, Tv \rangle$ לכל $u, v \in V$ (ראו תרגיל 4). הראו כי β היא מכפלה פנימית מעל V אם־ורק־אם T אופרטור חיובי הפיך מעל V.

6 הוכיחו או הביאו דוגמה נגדית: אם ρ תבנית בילינארית סימטרית מעל V, אזי
$$\{v \in V : \rho(v,v) = 0\}$$
תת־מרחב של V.

7 הסבירו מדוע ההוכחה של 9.13 (לכסון של תבנית בילינארית סימטרית על ידי בסיס אורתונורמלי מעל מרחב מכפלה פנימית ממשי) נכשלת אם משמיטים את ההנחה ש־ $\mathbf{F} = \mathbf{R}$.

8 מצאו נוסחאות עבור $\dim V^{(2)}_{\text{sym}}$ ו־ $\dim V^{(2)}_{\text{alt}}$ במונחים של $\dim V$.

9 נניח ש־ n מספר שלם חיובי ו־ $V = \{p \in \mathcal{P}_n(\mathbf{R}) : p(0) = p(1)\}$. נגדיר $\alpha : V \times V \to \mathbf{R}$ על ידי
$$\alpha(p,q) = \int_0^1 pq'.$$
הראו כי α תבנית בילינארית מתחלפת מעל V.

10 Suppose that n is a positive integer and

$$V = \{p \in \mathcal{P}_n(\mathbf{R}) : p(0) = p(1) \text{ and } p'(0) = p'(1)\}.$$

Define $\rho : V \times V \to \mathbf{R}$ by

$$\rho(p, q) = \int_0^1 pq''.$$

Show that ρ is a symmetric bilinear form on V.

10 נניח ש־ n מספר שלם חיובי וגם

$$V = \{p \in \mathcal{P}_n(\mathbf{R}) : p(0) = p(1) \text{ and } p'(0) = p'(1)\}.$$

נגדיר $\rho : V \times V \to \mathbf{R}$ על ידי

$$\rho(p, q) = \int_0^1 pq''.$$

הראו כי ρ תבנית ביליניארית סימטרית מעל V.

אלגברה ליניארית בדרך הישר, מהדורה רביעית, מחבר: שלדון אקסלר
Linear Algebra Done Right, fourth edition, by Sheldon Axler

9B Alternating Multilinear Forms

Multilinear Forms

9.24 definition: V^m

For m a positive integer, define V^m by

$$V^m = \underbrace{V \times \cdots \times V}_{m \text{ times}}.$$

Now we can define m-linear forms as a generalization of the bilinear forms that we discussed in the previous section.

9.25 definition: *m-linear form*, $V^{(m)}$, *multilinear form*

- For m a positive integer, an *m-linear form* on V is a function $\beta : V^m \to \mathbf{F}$ that is linear in each slot when the other slots are held fixed. This means that for each $k \in \{1, \dots, m\}$ and all $u_1, \dots, u_m \in V$, the function

$$v \mapsto \beta(u_1, \dots, u_{k-1}, v, u_{k+1}, \dots, u_m)$$

  is a linear map from V to $\mathbf{F}$.
- The set of m-linear forms on V is denoted by $V^{(m)}$.
- A function β is called a *multilinear form* on V if it is an m-linear form on V for some positive integer m.

In the definition above, the expression $\beta(u_1, \dots, u_{k-1}, v, u_{k+1}, \dots, u_m)$ means $\beta(v, u_2, \dots, u_m)$ if $k = 1$ and means $\beta(u_1, \dots, u_{m-1}, v)$ if $k = m$.

A 1-linear form on V is a linear functional on V. A 2-linear form on V is a bilinear form on V. You can verify that with the usual addition and scalar multiplication of functions, $V^{(m)}$ is a vector space.

9.26 example: *m-linear forms*

- Suppose $\alpha, \rho \in V^{(2)}$. Define a function $\beta : V^4 \to \mathbf{F}$ by

$$\beta(v_1, v_2, v_3, v_4) = \alpha(v_1, v_2)\, \rho(v_3, v_4).$$

  Then $\beta \in V^{(4)}$.
- Define $\beta : (\mathcal{L}(V))^m \to \mathbf{F}$ by

$$\beta(T_1, \dots, T_m) = \operatorname{tr}(T_1 \cdots T_m).$$

  Then β is an m-linear form on $\mathcal{L}(V)$.

טז תבניות מולטי־ליניאריות מתחלפות

תבניות מולטי־ליניאריות

9.24 הגדרה: V^m

עבור m מספר שלם חיובי, נגדיר V^m על ידי

$$V^m = \underbrace{V \times \cdots \times V}_{m \text{ פעמים}}.$$

כעת ביכולתנו להגדיר תבניות m־ליניאריות כהכללה של התבניות הבילינאריות שעליהן דיברנו ביחידה הקודמת.

9.25 הגדרה: תבנית m־ליניארית, $V^{(m)}$, תבנית מולטי־ליניארית

- עבור m מספר שלם חיובי, תבנית m־ליניארית מעל V היא פונקציה $\beta : V^m \to \mathbf{F}$ כך שהיא לינארית בכל רכיב כאשר שאר הרכיבים מקובעים. משמעות הדבר היא כי לכל $k \in \{1, \dots, m\}$ ולכל $u_1, \dots, u_m \in V$, הפונקציה
$$v \mapsto \beta(u_1, \dots, u_{k-1}, v, u_{k+1}, \dots, u_m)$$
היא העתקה לינארית מ־ V אל $\mathbf{F}$.
- קבוצת התבניות ה־ m־ליניאריות מעל V מסומנת על ידי $V^{(m)}$.
- פונקציה β נקראת תבנית מולטי־ליניארית מעל V אם היא תבנית m־ליניארית מעל V עבור m מספר שלם חיובי כלשהו.

בהגדרה המובאת לעיל, משמעות הביטוי $\beta(u_1, \dots, u_{k-1}, v, u_{k+1}, \dots, u_m)$ היא $\beta(v, u_2, \dots, u_m)$ אם $k = 1$ ו־ $\beta(u_1, \dots, u_{m-1}, v)$ אם $k = m$.

תבנית חד־ליניארית מעל V היא פונקציונל ליניארי מעל V. תבנית דו־ליניארית מעל V היא תבנית בילינארית מעל V. באפשרותכם לוודא כי עם פעולות החיבור והכפל בסקלר הרגילות של פונקציות, $V^{(m)}$ מרחב וקטורי.

9.26 דוגמה: תבניות m־ליניאריות

- נניח ש־ $\alpha, \rho \in V^{(2)}$. נגדיר פונקציה $\beta : V^4 \to \mathbf{F}$ על ידי
$$\beta(v_1, v_2, v_3, v_4) = \alpha(v_1, v_2)\, \rho(v_3, v_4).$$
אזי $\beta \in V^{(4)}$.
- נגדיר $\beta : (\mathcal{L}(V))^m \to \mathbf{F}$ על ידי
$$\beta(T_1, \dots, T_m) = \operatorname{tr}(T_1 \cdots T_m).$$
אזי β היא תבנית m־ליניארית מעל $\mathcal{L}(V)$.

Alternating multilinear forms, which we now define, play an important role as we head toward defining determinants.

9.27 definition: *alternating forms*, $V_{\text{alt}}^{(m)}$

Suppose m is a positive integer.

- An m-linear form α on V is called *alternating* if $\alpha(v_1, \dots, v_m) = 0$ whenever $v_1, \dots, v_m$ is a list of vectors in V with $v_j = v_k$ for some two distinct values of j and k in $\{1, \dots, m\}$.
- $V_{\text{alt}}^{(m)} = \{\alpha \in V^{(m)} : \alpha \text{ is an alternating } m\text{-linear form on } V\}$.

You should verify that $V_{\text{alt}}^{(m)}$ is a subspace of $V^{(m)}$. See Example 9.15 for examples of alternating 2-linear forms. See Exercise 2 for an example of an alternating 3-linear form.

The next result tells us that if a linearly dependent list is input to an alternating multilinear form, then the output equals 0.

9.28 *alternating multilinear forms and linear dependence*

Suppose m is a positive integer and α is an alternating m-linear form on V. If $v_1, \dots, v_m$ is a linearly dependent list in V, then

$$\alpha(v_1, \dots, v_m) = 0.$$

Proof Suppose $v_1, \dots, v_m$ is a linearly dependent list in V. By the linear dependence lemma (2.19), some v_k is a linear combination of $v_1, \dots, v_{k-1}$. Thus there exist $b_1, \dots, b_{k-1}$ such that $v_k = b_1 v_1 + \cdots + b_{k-1} v_{k-1}$. Now

$$\begin{aligned}
\alpha(v_1, \dots, v_m) &= \alpha\Big(v_1, \dots, v_{k-1}, \sum_{j=1}^{k-1} b_j v_j, v_{k+1}, \dots, v_m\Big) \\
&= \sum_{j=1}^{k-1} b_j \, \alpha(v_1, \dots, v_{k-1}, v_j, v_{k+1}, \dots, v_m) \\
&= 0.
\end{aligned}$$

The next result states that if $m > \dim V$, then there are no alternating m-linear forms on V other than the function on V^m that is identically 0.

9.29 *no nonzero alternating m-linear forms for $m > \dim V$*

Suppose $m > \dim V$. Then 0 is the only alternating m-linear form on V.

Proof Suppose that α is an alternating m-linear form on V and $v_1, \dots, v_m \in V$. Because $m > \dim V$, this list is not linearly independent (by 2.22). Thus 9.28 implies that $\alpha(v_1, \dots, v_m) = 0$. Hence α is the zero function from V^m to $\mathbf{F}$.

לתבניות מולטי־ליניאריות מתחלפות, שאותן נגדיר כעת, תפקיד מרכזי כאשר ניגש אל הגדרת הדטרמיננטות.

9.27 הגדרה: תבניות מתחלפות, $V_{\text{alt}}^{(m)}$

נניח ש־ m מספר שלם חיובי.

- תבנית m־ליניארית α מעל V נקראת מתחלפת אם $\alpha(v_1,\dots,v_m)=0$ בכל עת ש־ $v_1,\dots,v_m$ היא רשימה של וקטורים ב־ V עם $v_j=v_k$ עבור שני ערכים שונים של j ו־ k ב־ $\{1,\dots,m\}$.
- $V_{\text{alt}}^{(m)}=\{\alpha\in V^{(m)}:\alpha$ תבנית m־ליניארית מתחלפת מעל $V\}$.

עליכם לוודא כי $V_{\text{alt}}^{(m)}$ הוא תת־מרחב של $V^{(m)}$. ראו בדוגמה 9.15 דוגמאות של תבניות דו־ליניאריות מתחלפות. ראו תרגיל 2 עבור דוגמה של תבנית תלת־ליניארית מתחלפת.

המסקנה הבאה אומרת לנו שאם רשימה תלויה ליניארית משמשת כקלט לתבנית מולטי־ליניארית מתחלפת, אזי התוצאה שווה 0.

9.28 תבניות מולטי־ליניאריות מתחלפות ותלות ליניארית

נניח ש־ m מספר שלם חיובי ו־ α תבנית m־ליניארית מתחלפת מעל V. אם $v_1,\dots,v_m$ רשימה תלויה ליניארית ב־ V, אזי

$$\alpha(v_1,\dots,v_m)=0.$$

הוכחה נניח ש־ $v_1,\dots,v_m$ רשימה תלויה ליניארית ב־ V. לפי למת התלות הליניארית (2.19), וקטור v_k כלשהו הוא צירוף ליניארי של $v_1,\dots,v_{k-1}$. לפיכך קיימים $b_1,\dots,b_{k-1}$ כך ש־ $v_k=b_1v_1+\cdots+b_{k-1}v_{k-1}$. כעת

$$\begin{aligned}\alpha(v_1,\dots,v_m)&=\alpha\Big(v_1,\dots,v_{k-1},\sum_{j=1}^{k-1}b_jv_j,v_{k+1},\dots,v_m\Big)\\&=\sum_{j=1}^{k-1}b_j\,\alpha(v_1,\dots,v_{k-1},v_j,v_{k+1},\dots,v_m)\\&=0.\end{aligned}$$

המסקנה הבאה קובעת שאם $m>\dim V$, אזי אין תבניות m־ליניאריות מתחלפות מעל V מלבד הפונקציה מעל V^m שהיא זהותית 0.

9.29 עבור $m>\dim V$ אין תבניות m־ליניאריות מתחלפות שונות מאפס

נניח ש־ $m>\dim V$. אזי 0 היא התבנית ה־ m־ליניארית המתחלפת היחידה מעל V.

הוכחה נניח ש־ α תבנית m־ליניארית מתחלפת מעל V ו־ $v_1,\dots,v_m\in V$. מאחר ש־ $m>\dim V$, רשימה זו אינה בלתי תלויה ליניארית (לפי 2.22). לפיכך 9.28 גוררת ש־ $\alpha(v_1,\dots,v_m)=0$. לכן α היא פונקציית האפס מ־ V^m אל $\mathbf{F}$.

Alternating Multilinear Forms and Permutations

9.30 *swapping input vectors in an alternating multilinear form*

Suppose m is a positive integer, α is an alternating m-linear form on V, and $v_1, \dots, v_m$ is a list of vectors in V. Then swapping the vectors in any two slots of $\alpha(v_1, \dots, v_m)$ changes the value of α by a factor of -1.

Proof Put $v_1 + v_2$ in both the first two slots, getting

$$0 = \alpha(v_1 + v_2, v_1 + v_2, v_3, \dots, v_m).$$

Use the multilinear properties of α to expand the right side of the equation above (as in the proof of 9.16) to get

$$\alpha(v_2, v_1, v_3, \dots, v_m) = -\alpha(v_1, v_2, v_3, \dots, v_m).$$

Similarly, swapping the vectors in any two slots of $\alpha(v_1, \dots, v_m)$ changes the value of α by a factor of -1. ∎

To see what can happen with multiple swaps, suppose α is an alternating 3-linear form on V and $v_1, v_2, v_3 \in V$. To evaluate $\alpha(v_3, v_1, v_2)$ in terms of $\alpha(v_1, v_2, v_3)$, start with $\alpha(v_3, v_1, v_2)$ and swap the entries in the first and third slots, getting $\alpha(v_3, v_1, v_2) = -\alpha(v_2, v_1, v_3)$. Now in the last expression, swap the entries in the first and second slots, getting

$$\alpha(v_3, v_1, v_2) = -\alpha(v_2, v_1, v_3) = \alpha(v_1, v_2, v_3).$$

More generally, we see that if we do an odd number of swaps, then the value of α changes by a factor of -1, and if we do an even number of swaps, then the value of α does not change.

To deal with arbitrary multiple swaps, we need a bit of information about permutations.

9.31 definition: *permutation,* $\operatorname{perm} m$

Suppose m is a positive integer.

- A *permutation* of $(1, \dots, m)$ is a list $(j_1, \dots, j_m)$ that contains each of the numbers $1, \dots, m$ exactly once.
- The set of all permutations of $(1, \dots, m)$ is denoted by $\operatorname{perm} m$.

For example, $(2, 3, 4, 5, 1) \in \operatorname{perm} 5$. You should think of an element of $\operatorname{perm} m$ as a rearrangement of the first m positive integers.

The number of swaps used to change a permutation $(j_1, \dots, j_m)$ to the standard order $(1, \dots, m)$ can depend on the specific swaps selected. The following definition has the advantage of assigning a well-defined sign to every permutation.

תבניות מולטי־ליניאריות מתחלפות ותמורות

9.30 החלפה בין וקטורי קלט בתבנית מולטי־ליניארית מתחלפת

נניח ש־ m מספר שלם חיובי, α תבנית m־ליניארית מתחלפת מעל V, ו־ $v_1, \dots, v_m$ רשימה של וקטורים ב־ V. אזי החלפה בין הווקטורים בכל שני רכיבים של $\alpha(v_1, \dots, v_m)$ משנה את הערך של α בפקטור של -1.

הוכחה נציב $v_1 + v_2$ בשני הרכיבים הראשונים ונקבל

$$0 = \alpha(v_1 + v_2, v_1 + v_2, v_3, \dots, v_m).$$

נשתמש בתכונות המולטי־ליניאריות של α לפתיחת האגף הימני של המשוואה לעיל (כמו בהוכחת 9.16) ונקבל

$$\alpha(v_2, v_1, v_3, \dots, v_m) = -\alpha(v_1, v_2, v_3, \dots, v_m).$$

באופן דומה, החלפה בין הווקטורים בכל שני רכיבים של $\alpha(v_1, \dots, v_m)$ משנה את הערך של α בפקטור של -1. ■

כדי לראות מה יכול לקרות עם מספר החלפות, נניח ש־ α היא תבנית תלת־ליניארית מתחלפת מעל V ו־ $v_1, v_2, v_3 \in V$. כדי לחשב את $\alpha(v_3, v_1, v_2)$ במונחים של $\alpha(v_1, v_2, v_3)$, נתחיל עם $\alpha(v_3, v_1, v_2)$ ונחליף בין הרכיבים הראשון והשלישי ונקבל $\alpha(v_3, v_1, v_2) = -\alpha(v_2, v_1, v_3)$. כעת, בביטוי האחרון, נחליף בין איברי הרכיבים הראשון והשני ונקבל

$$\alpha(v_3, v_1, v_2) = -\alpha(v_2, v_1, v_3) = \alpha(v_1, v_2, v_3).$$

באופן כללי יותר, אנו רואים כי אם אנו מבצעים מספר אי־זוגי של החלפות, אזי הערך של α משתנה בפקטור של -1, ואם אנו מבצעים מספר זוגי של החלפות, אזי הערך של α אינו משתנה.

כדי לעסוק במספר שרירותי של החלפות, נזדקק למעט מידע הנוגע לתמורות.

9.31 הגדרה: תמורה, $\operatorname{perm} m$

נניח ש־ m מספר שלם חיובי.

- תמורה של $(1, \dots, m)$ היא רשימה $(j_1, \dots, j_m)$ המכילה את כל אחד מהמספרים $1, \dots, m$ בדיוק פעם אחת.
- קבוצת כל התמורות של $(1, \dots, m)$ מסומנת על ידי $\operatorname{perm} m$.

לדוגמה, $(2, 3, 4, 5, 1) \in \operatorname{perm} 5$. עליכם לחשוב על איבר של $\operatorname{perm} m$ כסידור מחדש של m השלמים החיוביים הראשונים.

מספר ההחלפות שנבצע כדי לשנות תמורה $(j_1, \dots, j_m)$ לסדר הטבעי $(1, \dots, m)$ יכול להיות תלוי בהחלפות הספציפיות שנבחרו. היתרון בהגדרה הבאה הוא השמת סימן מוגדר היטב לכל תמורה.

9.32 definition: *sign of a permutation*

The *sign* of a permutation $(j_1, \dots, j_m)$ is defined by

$$\text{sign}(j_1, \dots, j_m) = (-1)^N,$$

where N is the number of pairs of integers (k, ℓ) with $1 \le k < \ell \le m$ such that k appears after ℓ in the list $(j_1, \dots, j_m)$.

Hence the sign of a permutation equals 1 if the natural order has been changed an even number of times and equals -1 if the natural order has been changed an odd number of times.

9.33 example: *signs*

- The permutation $(1, \dots, m)$ [no changes in the natural order] has sign 1.
- The only pair of integers (k, ℓ) with $k < \ell$ such that k appears after ℓ in the list $(2, 1, 3, 4)$ is $(1, 2)$. Thus the permutation $(2, 1, 3, 4)$ has sign -1.
- In the permutation $(2, 3, \dots, m, 1)$, the only pairs (k, ℓ) with $k < \ell$ that appear with changed order are $(1, 2), (1, 3), \dots, (1, m)$. Because we have $m - 1$ such pairs, the sign of this permutation equals $(-1)^{m-1}$.

9.34 *swapping two entries in a permutation*

Swapping two entries in a permutation multiplies the sign of the permutation by -1.

Proof Suppose we have two permutations, where the second permutation is obtained from the first by swapping two entries. The two swapped entries were in their natural order in the first permutation if and only if they are not in their natural order in the second permutation. Thus we have a net change (so far) of 1 or -1 (both odd numbers) in the number of pairs not in their natural order.

Consider each entry between the two swapped entries. If an intermediate entry was originally in the natural order with respect to both swapped entries, then it is now in the natural order with respect to neither swapped entry. Similarly, if an intermediate entry was originally in the natural order with respect to neither of the swapped entries, then it is now in the natural order with respect to both swapped entries. If an intermediate entry was originally in the natural order with respect to exactly one of the swapped entries, then that is still true. Thus the net change (for each pair containing an entry between the two swapped entries) in the number of pairs not in their natural order is 2, -2, or 0 (all even numbers).

For all other pairs of entries, there is no change in whether or not they are in their natural order. Thus the total net change in the number of pairs not in their natural order is an odd number. Hence the sign of the second permutation equals -1 times the sign of the first permutation. ∎

9.32 הגדרה: סימן של תמורה

הסימן של תמורה $(j_1, \dots, j_m)$ מוגדר על ידי
$$\text{sign}(j_1, \dots, j_m) = (-1)^N,$$
כאשר N הוא מספר זוגות המספרים השלמים (k, ℓ) עבור $1 \le k < \ell \le m$ כך ש־ k מופיע אחרי ℓ ברשימה $(j_1, \dots, j_m)$.

לכן הסימן של תמורה שווה 1 אם מספר השינויים של הסדר הטבעי הוא זוגי ושווה -1 אם מספר השינויים של הסדר הטבעי הוא אי־זוגי.

9.33 דוגמה: סימנים

- סימן התמורה $(1, \dots, m)$ [ללא שינוי בסדר הטבעי] הוא 1.
- זוג המספרים השלמים היחידי (k, ℓ) עם $k < \ell$ כך ש־ k מופיע אחרי ℓ ברשימה $(2, 1, 3, 4)$ הוא $(1, 2)$. לפיכך סימן התמורה $(2, 1, 3, 4)$ הוא -1.
- בתמורה $(2, 3, \dots, m, 1)$, הזוגות היחידים (k, ℓ) עם $k < \ell$ המופיעים בסדר שונה הם $(1, 2), (1, 3), \dots, (1, m)$. מאחר שיש לנו $m - 1$ זוגות שכאלה, הסימן של תמורה זו שווה $(-1)^{m-1}$.

9.34 החלפת שני איברים בתמורה

החלפת שני איברים בתמורה מכפילה את סימן התמורה ב־ -1.

הוכחה נניח שיש בידינו שתי תמורות, כאשר התמורה השנייה מושגת כתוצאה מהחלפת שני איברים של הראשונה. שני האיברים המוחלפים הופיעו כסדרם הטבעי בתמורה הראשונה אם־ורק־אם הם אינם כסדרם הטבעי בתמורה השנייה. לפיכך מתקבל שינוי נטו (עד כה) של 1 או -1 (שניהם מספרים אי־זוגיים) עבור המספר של הזוגות שאינם כסדרם הטבעי.

נתבונן בכל איבר בין אותם שני איברים מוחלפים. אם איבר המוצב ביניהם היה במקור מוצב בסדר הטבעי ביחס לשני האיברים המוחלפים, אזי איבר זה כעת אינו בסדר הטבעי ביחס לאף אחד משני האיברים המוחלפים. באופן דומה, אם איבר המוצב ביניהם היה במקור שלא בסדר הטבעי ביחס לכל אחד משני האיברים המוחלפים, אזי איבר זה הוא כעת בסדר הטבעי ביחס לכל אחד משני האיברים המוחלפים. אם איבר המוצב ביניהם היה במקור מוצב בסדר הטבעי ביחס לאיבר מוחלף אחד בלבד, אזי מצב זה עדיין מתקיים. לפיכך השינוי נטו (לכל זוג שבו אחד מהאיברים הוא בין אותם שני איברים מוחלפים) במספר הזוגות שאינם בסדר הטבעי הוא -2, 2 או 0 (כולם מספרים זוגיים).

עבור כל שאר זוגות האיברים אין שינוי בשאלה אם הם בסדרם הטבעי אם לא. לפיכך סך השינוי נטו במספר הזוגות שאינם בסדרם הטבעי הוא מספר אי־זוגי. לכן הסימן של התמורה השנייה הוא -1 כפול סימן התמורה הראשונה. ■

9.35 *permutations and alternating multilinear forms*

Suppose m is a positive integer and $\alpha \in V_{\text{alt}}^{(m)}$. Then

$$\alpha(v_{j_1}, \dots, v_{j_m}) = (\text{sign}(j_1, \dots, j_m))\alpha(v_1, \dots, v_m)$$

for every list $v_1, \dots, v_m$ of vectors in V and all $(j_1, \dots, j_m) \in \text{perm}\, m$.

Proof Suppose $v_1, \dots, v_m \in V$ and $(j_1, \dots, j_m) \in \text{perm}\, m$. We can get from $(j_1, \dots, j_m)$ to $(1, \dots, m)$ by a series of swaps of entries in different slots. Each such swap changes the value of α by a factor of -1 (by 9.30) and also changes the sign of the remaining permutation by a factor of -1 (by 9.34). After an appropriate number of swaps, we reach the permutation $1, \dots, m$, which has sign 1. Thus the value of α changed signs an even number of times if $\text{sign}(j_1, \dots, j_m) = 1$ and an odd number of times if $\text{sign}(j_1, \dots, j_m) = -1$, which gives the desired result. ∎

Our use of permutations now leads in a natural way to the following beautiful formula for alternating n-linear forms on an n-dimensional vector space.

9.36 *formula for* $(\dim V)$*-linear alternating forms on* V

Let $n = \dim V$. Suppose $e_1, \dots, e_n$ is a basis of V and $v_1, \dots, v_n \in V$. For each $k \in \{1, \dots, n\}$, let $b_{1,k}, \dots, b_{n,k} \in \mathbf{F}$ be such that

$$v_k = \sum_{j=1}^{n} b_{j,k} e_j.$$

Then

$$\alpha(v_1, \dots, v_n) = \alpha(e_1, \dots, e_n) \sum_{(j_1, \dots, j_n) \in \text{perm}\, n} (\text{sign}(j_1, \dots, j_n)) b_{j_1,1} \cdots b_{j_n,n}$$

for every alternating n-linear form α on V.

Proof Suppose α is an alternating n-linear form on V. Then

$$\begin{aligned}
\alpha(v_1, \dots, v_n) &= \alpha\Big(\sum_{j_1=1}^{n} b_{j_1,1} e_{j_1}, \dots, \sum_{j_n=1}^{n} b_{j_n,n} e_{j_n} \Big) \\
&= \sum_{j_1=1}^{n} \cdots \sum_{j_n=1}^{n} b_{j_1,1} \cdots b_{j_n,n}\, \alpha(e_{j_1}, \dots, e_{j_n}) \\
&= \sum_{(j_1, \dots, j_n) \in \text{perm}\, n} b_{j_1,1} \cdots b_{j_n,n}\, \alpha(e_{j_1}, \dots, e_{j_n}) \\
&= \alpha(e_1, \dots, e_n) \sum_{(j_1, \dots, j_n) \in \text{perm}\, n} (\text{sign}(j_1, \dots, j_n)) b_{j_1,1} \cdots b_{j_n,n},
\end{aligned}$$

where the third line holds because $\alpha(e_{j_1}, \dots, e_{j_n}) = 0$ if $j_1, \dots, j_n$ are not distinct integers, and the last line holds by 9.35. ∎

9.35 תמורות ותבניות מולטי־ליניאריות מתחלפות

נניח ש־ m מספר שלם חיובי ו־ $\alpha \in V_{\text{alt}}^{(m)}$. אזי
$$\alpha(v_{j_1}, \dots, v_{j_m}) = (\text{sign}(j_1, \dots, j_m))\alpha(v_1, \dots, v_m)$$
לכל רשימת וקטורים $v_1, \dots, v_m$ ב־ V ולכל $(j_1, \dots, j_m) \in \text{perm}\, m$.

הוכחה נניח ש־ $v_1, \dots, v_m \in V$ ו־ $(j_1, \dots, j_m) \in \text{perm}\, m$. נוכל להגיע מ־ $(j_1, \dots, j_m)$ אל $(1, \dots, m)$ באמצעות סדרה של החלפות איברים ברכיבים שונים. כל החלפה שכזו משנה את ערכו של α בפקטור של -1 (לפי 9.30) ובנוסף משנה את הסימן של התמורה שנותרה בפקטור של -1 (לפי 9.34). אחרי מספר החלפות מתאים, נגיע לתמורה $1, \dots, m$, שסימנה 1. לפיכך ערך α משנה סימן מספר זוגי של פעמים אם $\text{sign}(j_1, \dots, j_m) = 1$ ומספר אי־זוגי של פעמים אם $\text{sign}(j_1, \dots, j_m) = -1$, מה שמעניק לנו את התוצאה הרצויה. ■

השימוש שלנו בתמורות מוביל כעת באורח טבעי אל הנוסחה היפה המובאת להלן עבור תבניות n־ליניאריות מתחלפות מעל מרחב וקטורי מממד n.

9.36 נוסחה עבור תבניות $(\dim V)$־ליניאריות מתחלפות מעל V

יהי $n = \dim V$. נניח ש־ $e_1, \dots, e_n$ בסיס של V ו־ $v_1, \dots, v_n \in V$. לכל $k \in \{1, \dots, n\}$, יהיו $b_{1,k}, \dots, b_{n,k} \in \mathbf{F}$ כך ש־
$$v_k = \sum_{j=1}^{n} b_{j,k} e_j.$$
אזי
$$\alpha(v_1, \dots, v_n) = \alpha(e_1, \dots, e_n) \sum_{(j_1,\dots,j_n)\in \text{perm}\, n} (\text{sign}(j_1, \dots, j_n)) b_{j_1,1} \cdots b_{j_n,n}$$
לכל תבנית n־ליניארית מתחלפת α מעל V.

הוכחה נניח ש־ α תבנית n־ליניארית מתחלפת מעל V. אזי
$$\begin{aligned}
\alpha(v_1, \dots, v_n) &= \alpha\Big(\sum_{j_1=1}^{n} b_{j_1,1} e_{j_1}, \dots, \sum_{j_n=1}^{n} b_{j_n,n} e_{j_n}\Big) \\
&= \sum_{j_1=1}^{n} \cdots \sum_{j_n=1}^{n} b_{j_1,1} \cdots b_{j_n,n}\, \alpha(e_{j_1}, \dots, e_{j_n}) \\
&= \sum_{(j_1,\dots,j_n)\in \text{perm}\, n} b_{j_1,1} \cdots b_{j_n,n}\, \alpha(e_{j_1}, \dots, e_{j_n}) \\
&= \alpha(e_1, \dots, e_n) \sum_{(j_1,\dots,j_n)\in \text{perm}\, n} (\text{sign}(j_1, \dots, j_n)) b_{j_1,1} \cdots b_{j_n,n},
\end{aligned}$$
כאשר השורה השלישית מתקיימת מאחר ש־ $\alpha(e_{j_1}, \dots, e_{j_n}) = 0$ אם $j_1, \dots, j_n$ אינם מספרים שלמים שונים, והשורה האחרונה מתקיימת לפי 9.35. ■

The following result will be the key to our definition of the determinant in the next section.

9.37 $\dim V_{\text{alt}}^{(\dim V)} = 1$

The vector space $V_{\text{alt}}^{(\dim V)}$ has dimension one.

Proof Let $n = \dim V$. Suppose α and α' are alternating n-linear forms on V with $\alpha \neq 0$. Let $e_1, \dots, e_n$ be such that $\alpha(e_1, \dots, e_n) \neq 0$. There exists $c \in \mathbf{F}$ such that

$$\alpha'(e_1, \dots, e_n) = c\alpha(e_1, \dots, e_n).$$

Furthermore, 9.28 implies that $e_1, \dots, e_n$ is linearly independent and thus is a basis of V.

Suppose $v_1, \dots, v_n \in V$. Let $b_{j,k}$ be as in 9.36 for $j, k = 1, \dots, n$. Then

$$\begin{aligned}
\alpha'(v_1, \dots, v_n) &= \alpha'(e_1, \dots, e_n) \sum_{(j_1,\dots,j_n)\in \operatorname{perm} n} (\operatorname{sign}(j_1, \dots, j_n)) b_{j_1,1} \cdots b_{j_n,n} \\
&= c\alpha(e_1, \dots, e_n) \sum_{(j_1,\dots,j_n)\in \operatorname{perm} n} (\operatorname{sign}(j_1, \dots, j_n)) b_{j_1,1} \cdots b_{j_n,n} \\
&= c\alpha(v_1, \dots, v_n),
\end{aligned}$$

where the first and last lines above come from 9.36. The equation above implies that $\alpha' = c\alpha$. Thus α', α is not a linearly independent list, which implies that $\dim V_{\text{alt}}^{(n)} \leq 1$.

To complete the proof, we only need to show that there exists a nonzero alternating n-linear form α on V (thus eliminating the possibility that $\dim V_{\text{alt}}^{(n)}$ equals 0). To do this, let $e_1, \dots, e_n$ be any basis of V, and let $\varphi_1, \dots, \varphi_n \in V'$ be the linear functionals on V that allow us to express each element of V as a linear combination of $e_1, \dots, e_n$. In other words,

$$v = \sum_{j=1}^{n} \varphi_j(v) e_j$$

for every $v \in V$ (see 3.114). Now for $v_1, \dots, v_n \in V$, define

$$\alpha(v_1, \dots, v_n) = \sum_{(j_1,\dots,j_n)\in \operatorname{perm} n} (\operatorname{sign}(j_1, \dots, j_n)) \varphi_{j_1}(v_1) \cdots \varphi_{j_n}(v_n). \tag{9.38}$$

The verification that α is an n-linear form on V is straightforward.

To see that α is alternating, suppose $v_1, \dots, v_n \in V$ with $v_1 = v_2$. For each $(j_1, \dots, j_n) \in \operatorname{perm} n$, the permutation $(j_2, j_1, j_3, \dots, j_n)$ has the opposite sign. Because $v_1 = v_2$, the contributions from these two permutations to the sum in 9.38 cancel each other. Hence $\alpha(v_1, v_1, v_3, \dots, v_n) = 0$. Similarly, $\alpha(v_1, \dots, v_n) = 0$ if any two vectors in the list $v_1, \dots, v_n$ are equal. Thus α is alternating.

Finally, consider 9.38 with each $v_k = e_k$. Because $\varphi_j(e_k)$ equals 0 if $j \neq k$ and equals 1 if $j = k$, only the permutation $(1, \dots, n)$ makes a nonzero contribution to the right side of 9.38 in this case, giving the equation $\alpha(e_1, \dots, e_n) = 1$. Thus we have produced a nonzero alternating n-linear form α on V, as desired. ∎

המסקנה הבאה תהיה עמוד התווך בהגדרה שלנו לדטרמיננטה ביחידה הבאה.

9.37 $\dim V_{\text{alt}}^{(\dim V)} = 1$

המרחב הווקטורי $V_{\text{alt}}^{(\dim V)}$ הוא חד־ממדי.

הוכחה יהי $n = \dim V$. נניח ש־ α ו־ α' תבניות n־ליניאריות מתחלפות מעל V עם $\alpha \neq 0$. יהיו $e_1, \dots, e_n$ כך ש־ $\alpha(e_1, \dots, e_n) \neq 0$. קיים $c \in \mathbf{F}$ כך ש־

$$\alpha'(e_1, \dots, e_n) = c\alpha(e_1, \dots, e_n).$$

יתר על כן, 9.28 גורר ש־ $e_1, \dots, e_n$ בלתי תלויה ליניארית ולפיכך בסיס של V.

נניח ש־ $v_1, \dots, v_n \in V$. יהי $b_{j,k}$ כמו ב־ 9.36 עבור $j, k = 1, \dots, n$. אזי

$$\begin{aligned}
\alpha'(v_1, \dots, v_n) &= \alpha'(e_1, \dots, e_n) \sum_{(j_1,\dots,j_n)\in \operatorname{perm} n} (\operatorname{sign}(j_1, \dots, j_n)) b_{j_1,1} \cdots b_{j_n,n} \\
&= c\alpha(e_1, \dots, e_n) \sum_{(j_1,\dots,j_n)\in \operatorname{perm} n} (\operatorname{sign}(j_1, \dots, j_n)) b_{j_1,1} \cdots b_{j_n,n} \\
&= c\alpha(v_1, \dots, v_n),
\end{aligned}$$

כאשר השורות הראשונה והאחרונה לעיל נובעות מ־ 9.36. המשוואה לעיל גוררת ש־ $\alpha' = c\alpha$. לפיכך α', α אינה רשימה בלתי תלויה ליניארית, מה שגורר ש־ $\dim V_{\text{alt}}^{(n)} \leq 1$.

להשלמת ההוכחה, כל שנותר לנו להראות הוא כי קיימת תבנית n־ליניארית מתחלפת שונה מאפס α מעל V (ובכך תישלל האפשרות ש־ $\dim V_{\text{alt}}^{(n)}$ שווה 0). לצורך כך, תהי $e_1, \dots, e_n$ בסיס כלשהו של V, ויהיו $\varphi_1, \dots, \varphi_n \in V'$ הפונקציונלים הליניאריים מעל V המאפשרים לנו לבטא כל איבר של V כצירוף ליניארי של $e_1, \dots, e_n$. במילים אחרות,

$$v = \sum_{j=1}^{n} \varphi_j(v) e_j$$

לכל $v \in V$ (ראו 3.114). כעת, עבור $v_1, \dots, v_n \in V$, נגדיר

$$\alpha(v_1, \dots, v_n) = \sum_{(j_1,\dots,j_n)\in \operatorname{perm} n} (\operatorname{sign}(j_1, \dots, j_n)) \varphi_{j_1}(v_1) \cdots \varphi_{j_n}(v_n). \tag{9.38}$$

הווידוא כי α תבנית n־ליניארית מעל V הוא פשוט.

כדי לראות ש־ α מתחלפת, נניח ש־ $v_1, \dots, v_n \in V$ עם $v_1 = v_2$. לכל $(j_1, \dots, j_n) \in \operatorname{perm} n$, לתמורה $(j_2, j_1, j_3, \dots, j_n)$ סימן הפוך. מאחר ש־ $v_1 = v_2$, התרומות של שתי התמורות לסכום 9.38 מבטלות האחת את השנייה. לכן $\alpha(v_1, v_1, v_3, \dots, v_n) = 0$. באופן דומה, $\alpha(v_1, \dots, v_n) = 0$ אם שני וקטורים כלשהם ברשימה $v_1, \dots, v_n$ שווים. לפיכך α מתחלפת.

לסיום, נתבונן ב־ 9.38 כאשר כל $v_k = e_k$. מאחר ש־ $\varphi_j(e_k)$ שווה 0 אם $j \neq k$ ושווה 1 אם $j = k$, רק התמורה $(1, \dots, n)$ תורמת תרומה השונה מאפס לאגף הימני של 9.38 במקרה הזה, ומעניקה למשוואה את הערך $\alpha(e_1, \dots, e_n) = 1$. לפיכך הפקנו תבנית n־ליניארית מתחלפת שונה מאפס α מעל V, כפי שרצינו. ■

The formula 9.38 used in the last proof to construct a nonzero alternating n-linear form came from the formula in 9.36, and that formula arose naturally from the properties of an alternating multilinear form.

Earlier we showed that the value of an alternating multilinear form applied to a linearly dependent list is 0; see 9.28. The next result provides a converse of 9.28 for alternating n-linear forms when $n = \dim V$. In the following result, the statement that α is nonzero means (as usual for a function) that α is not the function on V^n that is identically 0.

9.39 *alternating $(\dim V)$-linear forms and linear independence*

Let $n = \dim V$. Suppose α is a nonzero alternating n-linear form on V and $e_1, \dots, e_n$ is a list of vectors in V. Then

$$\alpha(e_1, \dots, e_n) \neq 0$$

if and only if $e_1, \dots, e_n$ is linearly independent.

Proof First suppose $\alpha(e_1, \dots, e_n) \neq 0$. Then 9.28 implies that $e_1, \dots, e_n$ is linearly independent.

To prove the implication in the other direction, now suppose $e_1, \dots, e_n$ is linearly independent. Because $n = \dim V$, this implies that $e_1, \dots, e_n$ is a basis of V (see 2.38).

Because α is not the zero n-linear form, there exist $v_1, \dots, v_n \in V$ such that $\alpha(v_1, \dots, v_n) \neq 0$. Now 9.36 implies that $\alpha(e_1, \dots, e_n) \neq 0$. ∎

Exercises 9B

1 Suppose m is a positive integer. Show that $\dim V^{(m)} = (\dim V)^m$.

2 Suppose $n \geq 3$ and $\alpha : \mathbf{F}^n \times \mathbf{F}^n \times \mathbf{F}^n \to \mathbf{F}$ is defined by

$$\begin{aligned}&\alpha((x_1, \dots, x_n), (y_1, \dots, y_n), (z_1, \dots, z_n)) \\ &\quad = x_1y_2z_3 - x_2y_1z_3 - x_3y_2z_1 - x_1y_3z_2 + x_3y_1z_2 + x_2y_3z_1.\end{aligned}$$

Show that α is an alternating 3-linear form on $\mathbf{F}^n$.

3 Suppose m is a positive integer and α is an m-linear form on V such that $\alpha(v_1, \dots, v_m) = 0$ whenever $v_1, \dots, v_m$ is a list of vectors in V with $v_j = v_{j+1}$ for some $j \in \{1, \dots, m-1\}$. Prove that α is an alternating m-linear form on V.

4 Prove or give a counterexample: If $\alpha \in V^{(4)}_{\text{alt}}$, then

$$\{(v_1, v_2, v_3, v_4) \in V^4 : \alpha(v_1, v_2, v_3, v_4) = 0\}$$

is a subspace of V^4.

מוקדם יותר הראינו כי הערך של תבנית מולטי־ליניארית מתחלפת המוחלת על רשימה תלויה ליניארית הוא 0; ראו 9.28. המסקנה הבאה מספקת כיוון הפוך ל־ 9.28 עבור תבניות n־ליניאריות מתחלפות כאשר $n = \dim V$. במסקנה הבאה, הקביעה כי α שונה מאפס משמעותה (כמו עבור כל פונקציה) ש־ α אינה הפונקציה מעל V^n שהיא זהותית 0.

הנוסחה 9.38 שבה השתמשנו בהוכחה האחרונה לבניית תבנית n־ליניארית מתחלפת נובעת מהנוסחה ב־ 9.36, ונוסחה זו נובעת באופן טבעי מהתכונות של תבנית מולטי־ליניארית מתחלפת.

9.39 תבניות $(\dim V)$־ליניאריות מתחלפות ואי־תלות ליניארית

יהי $n = \dim V$. נניח ש־ α תבנית n־ליניארית מתחלפת מעל V השונה מאפס ו־ $e_1, \dots, e_n$ רשימה של וקטורים ב־ V. אזי

$$\alpha(e_1, \dots, e_n) \neq 0$$

אם־ורק־אם $e_1, \dots, e_n$ בלתי תלויה ליניארית.

הוכחה תחילה נניח כי $\alpha(e_1, \dots, e_n) \neq 0$. אזי 9.28 גורר ש־ $e_1, \dots, e_n$ בלתי תלויה ליניארית.

להוכחת הגרירה בכיוון השני, כעת נניח ש־ $e_1, \dots, e_n$ בלתי תלויה ליניארית. מאחר ש־ $n = \dim V$, מכאן נובע ש־ $e_1, \dots, e_n$ בסיס של V (ראו 2.38).

מאחר ש־ α אינה תבנית n־ליניארית האפס, קיימים $v_1, \dots, v_n \in V$ כך ש־ $\alpha(v_1, \dots, v_n) \neq 0$. כעת 9.36 גוררת ש־ $\alpha(e_1, \dots, e_n) \neq 0$. ■

תרגילים טז

1 נניח ש־ m מספר שלם חיובי. הראו כי $\dim V^{(m)} = (\dim V)^m$.

2 נניח ש־ $n \geq 3$ ו־ $\alpha : \mathbf{F}^n \times \mathbf{F}^n \times \mathbf{F}^n \to \mathbf{F}$ מוגדרת על ידי

$$\alpha((x_1, \dots, x_n), (y_1, \dots, y_n), (z_1, \dots, z_n)) = x_1y_2z_3 - x_2y_1z_3 - x_3y_2z_1 - x_1y_3z_2 + x_3y_1z_2 + x_2y_3z_1.$$

הראו כי α תבנית 3־ליניארית מתחלפת מעל $\mathbf{F}^n$.

3 נניח ש־ m הוא מספר שלם חיובי ו־ α היא תבנית m־ליניארית מעל V כך ש־ $\alpha(v_1, \dots, v_m) = 0$ בכל עת ש־ $v_1, \dots, v_m$ רשימה של וקטורים ב־ V עם $v_j = v_{j+1}$ עבור $j \in \{1, \dots, m-1\}$ כלשהו. הוכיחו כי α תבנית m־ליניארית מתחלפת מעל V.

4 הוכיחו או הביאו דוגמה נגדית: אם $\alpha \in V^{(4)}_{\text{alt}}$, אזי

$$\{(v_1, v_2, v_3, v_4) \in V^4 : \alpha(v_1, v_2, v_3, v_4) = 0\}$$

תת־מרחב של V^4.

5 Suppose m is a positive integer and β is an m-linear form on V. Define an m-linear form α on V by

$$\alpha(v_1, \dots, v_m) = \sum_{(j_1,\dots,j_m)\in \operatorname{perm} m} (\operatorname{sign}(j_1, \dots, j_m))\beta(v_{j_1}, \dots, v_{j_m})$$

for $v_1, \dots, v_m \in V$. Explain why $\alpha \in V_{\text{alt}}^{(m)}$.

6 Suppose m is a positive integer and β is an m-linear form on V. Define an m-linear form α on V by

$$\alpha(v_1, \dots, v_m) = \sum_{(j_1,\dots,j_m)\in \operatorname{perm} m} \beta(v_{j_1}, \dots, v_{j_m})$$

for $v_1, \dots, v_m \in V$. Explain why

$$\alpha(v_{k_1}, \dots, v_{k_m}) = \alpha(v_1, \dots, v_m)$$

for all $v_1, \dots, v_m \in V$ and all $(k_1, \dots, k_m) \in \operatorname{perm} m$.

7 Give an example of a nonzero alternating 2-linear form α on $\mathbf{R}^3$ and a linearly independent list v_1, v_2 in $\mathbf{R}^3$ such that $\alpha(v_1, v_2) = 0$.

This exercise shows that 9.39 can fail if the hypothesis that $n = \dim V$ is deleted.

5 נניח ש־ m מספר שלם חיובי ו־ β תבנית m־ליניארית מעל V. נגדיר תבנית m־ליניארית α מעל V על ידי

$$\alpha(v_1,\dots,v_m)=\sum_{(j_1,\dots,j_m)\in \operatorname{perm} m}(\operatorname{sign}(j_1,\dots,j_m))\beta(v_{j_1},\dots,v_{j_m})$$

עבור $v_1,\dots,v_m\in V$. הסבירו מדוע $\alpha\in V_{\text{alt}}^{(m)}$.

6 נניח ש־ m מספר שלם חיובי ו־ β תבנית m־ליניארית מעל V. נגדיר תבנית m־ליניארית α מעל V על ידי

$$\alpha(v_1,\dots,v_m)=\sum_{(j_1,\dots,j_m)\in \operatorname{perm} m}\beta(v_{j_1},\dots,v_{j_m})$$

עבור $v_1,\dots,v_m\in V$. הסבירו מדוע

$$\alpha(v_{k_1},\dots,v_{k_m})=\alpha(v_1,\dots,v_m)$$

לכל $v_1,\dots,v_m\in V$ ולכל $(k_1,\dots,k_m)\in \operatorname{perm} m$.

7 הביאו דוגמה של תבנית 2־ליניארית מתחלפת שונה מאפס α מעל $\mathbf{R}^3$ ורשימה בלתי תלויה ליניארית v_1,v_2 ב־ $\mathbf{R}^3$ כך ש־ $\alpha(v_1,v_2)=0$.

תרגיל זה מראה כי 9.39 עלולה להיכשל אם נסיר את ההנחה כי $n=\dim V$.

9C Determinants

Defining the Determinant

The next definition will lead us to a clean, beautiful, basis-free definition of the determinant of an operator.

9.40 definition: α_T

Suppose that m is a positive integer and $T \in \mathcal{L}(V)$. For $\alpha \in V_{\text{alt}}^{(m)}$, define $\alpha_T \in V_{\text{alt}}^{(m)}$ by

$$\alpha_T(v_1, \dots, v_m) = \alpha(Tv_1, \dots, Tv_m)$$

for each list $v_1, \dots, v_m$ of vectors in V.

Suppose $T \in \mathcal{L}(V)$. If $\alpha \in V_{\text{alt}}^{(m)}$ and $v_1, \dots, v_m$ is a list of vectors in V with $v_j = v_k$ for some $j \neq k$, then $Tv_j = Tv_k$, which implies that $\alpha_T(v_1, \dots, v_m) = \alpha(Tv_1, \dots, Tv_m) = 0$. Thus the function $\alpha \mapsto \alpha_T$ is a linear map of $V_{\text{alt}}^{(m)}$ to itself.

We know that $\dim V_{\text{alt}}^{(\dim V)} = 1$ (see 9.37). Every linear map from a one-dimensional vector space to itself is multiplication by some unique scalar. For the linear map $\alpha \mapsto \alpha_T$, we now define $\det T$ to be that scalar.

9.41 definition: *determinant of an operator,* $\det T$

Suppose $T \in \mathcal{L}(V)$. The *determinant* of T, denoted by $\det T$, is defined to be the unique number in $\mathbf{F}$ such that

$$\alpha_T = (\det T)\, \alpha$$

for all $\alpha \in V_{\text{alt}}^{(\dim V)}$.

9.42 example: *determinants of operators*

Let $n = \dim V$.

- If I is the identity operator on V, then $\alpha_I = \alpha$ for all $\alpha \in V_{\text{alt}}^{(n)}$. Thus $\det I = 1$.
- More generally, if $\lambda \in \mathbf{F}$, then $\alpha_{\lambda I} = \lambda^n \alpha$ for all $\alpha \in V_{\text{alt}}^{(n)}$. Thus $\det(\lambda I) = \lambda^n$.
- Still more generally, if $T \in \mathcal{L}(V)$ and $\lambda \in \mathbf{F}$, then $\alpha_{\lambda T} = \lambda^n \alpha_T = \lambda^n (\det T)\alpha$ for all $\alpha \in V_{\text{alt}}^{(n)}$. Thus $\det(\lambda T) = \lambda^n \det T$.
- Suppose $T \in \mathcal{L}(V)$ and there is a basis $e_1, \dots, e_n$ of V consisting of eigenvectors of T, with corresponding eigenvalues $\lambda_1, \dots, \lambda_n$. If $\alpha \in V_{\text{alt}}^{(n)}$, then

$$\alpha_T(e_1, \dots, e_n) = \alpha(\lambda_1 e_1, \dots, \lambda_n e_n) = (\lambda_1 \cdots \lambda_n)\alpha(e_1, \dots, e_n).$$

If $\alpha \neq 0$, then 9.39 implies $\alpha(e_1, \dots, e_n) \neq 0$. Thus the equation above implies

$$\det T = \lambda_1 \cdots \lambda_n.$$

טז דטרמיננטות

הגדרת הדטרמיננטה

הנתיב אל הגדרה נקייה ויפה של הדטרמיננטה של אופרטור, שאינה תלויה בבחירת בסיס, עובר דרך ההגדרה הבאה.

9.40 הגדרה: α_T

נניח ש־m מספר שלם חיובי ו־$T \in \mathcal{L}(V)$. עבור $\alpha \in V_{\text{alt}}^{(m)}$, נגדיר $\alpha_T \in V_{\text{alt}}^{(m)}$ על ידי

$$\alpha_T(v_1, \dots, v_m) = \alpha(Tv_1, \dots, Tv_m)$$

לכל רשימת וקטורים $v_1, \dots, v_m$ ב־V.

נניח ש־$T \in \mathcal{L}(V)$. אם $\alpha \in V_{\text{alt}}^{(m)}$ ו־$v_1, \dots, v_m$ רשימת וקטורים ב־V עם $v_j = v_k$ עבור $j \neq k$ כלשהם, אזי $Tv_j = Tv_k$, מה שגורר ש־$\alpha_T(v_1, \dots, v_m) = \alpha(Tv_1, \dots, Tv_m) = 0$. לפיכך הפונקציה $\alpha \mapsto \alpha_T$ היא העתקה ליניארית של $V_{\text{alt}}^{(m)}$ לעצמו.

אנו יודעים כי $\dim V_{\text{alt}}^{(\dim V)} = 1$ (ראו 9.37). כל העתקה ליניארית ממרחב וקטורי חד־ממדי לעצמו היא כפל בסקלר יחיד כלשהו. עבור ההעתקה הליניארית $\alpha \mapsto \alpha_T$, נגדיר כעת את $\det T$ כסקלר זה.

9.41 הגדרה: דטרמיננטה של אופרטור, $\det T$

נניח ש־$T \in \mathcal{L}(V)$. הדטרמיננטה של T, מסומנת על ידי $\det T$, מוגדרת כמספר היחיד ב־$\mathbf{F}$ כך ש־

$$\alpha_T = (\det T)\, \alpha$$

לכל $\alpha \in V_{\text{alt}}^{(\dim V)}$.

9.42 דוגמה: דטרמיננטות של אופרטורים

יהי $n = \dim V$.

- אם I הוא אופרטור הזהות מעל V, אזי $\alpha_I = \alpha$ לכל $\alpha \in V_{\text{alt}}^{(n)}$. לפיכך $\det I = 1$.
- באופן כללי יותר, אם $\lambda \in \mathbf{F}$, אזי $\alpha_{\lambda I} = \lambda^n \alpha$ לכל $\alpha \in V_{\text{alt}}^{(n)}$. לפיכך $\det(\lambda I) = \lambda^n$.
- באופן כללי עוד יותר, אם $T \in \mathcal{L}(V)$ ו־$\lambda \in \mathbf{F}$, אזי $\alpha_{\lambda T} = \lambda^n \alpha_T = \lambda^n (\det T)\alpha$ לכל $\alpha \in V_{\text{alt}}^{(n)}$. לפיכך $\det(\lambda T) = \lambda^n \det T$.
- נניח ש־$T \in \mathcal{L}(V)$ וקיים בסיס $e_1, \dots, e_n$ של V המורכב מווקטורים עצמיים של T, השייכים לערכים העצמיים $\lambda_1, \dots, \lambda_n$. אם $\alpha \in V_{\text{alt}}^{(n)}$, אזי
$$\alpha_T(e_1, \dots, e_n) = \alpha(\lambda_1 e_1, \dots, \lambda_n e_n) = (\lambda_1 \cdots \lambda_n)\alpha(e_1, \dots, e_n).$$
אם $\alpha \neq 0$, אזי 9.39 גוררת ש־$\alpha(e_1, \dots, e_n) \neq 0$. לפיכך המשוואה לעיל גוררת ש־
$$\det T = \lambda_1 \cdots \lambda_n.$$

Our next task is to define and give a formula for the determinant of a square matrix. To do this, we associate with each square matrix an operator and then define the determinant of the matrix to be the determinant of the associated operator.

9.43 definition: *determinant of a matrix,* $\det A$

Suppose that n is a positive integer and A is an n-by-n square matrix with entries in $\mathbf{F}$. Let $T \in \mathcal{L}(\mathbf{F}^n)$ be the operator whose matrix with respect to the standard basis of $\mathbf{F}^n$ equals A. The *determinant* of A, denoted by $\det A$, is defined by $\det A = \det T$.

9.44 example: *determinants of matrices*

- If I is the n-by-n identity matrix, then the corresponding operator on $\mathbf{F}^n$ is the identity operator I on $\mathbf{F}^n$. Thus the first bullet point of 9.42 implies that the determinant of the identity matrix is 1.
- Suppose A is a diagonal matrix with $\lambda_1, \dots, \lambda_n$ on the diagonal. Then the corresponding operator on $\mathbf{F}^n$ has the standard basis of $\mathbf{F}^n$ as eigenvectors, with eigenvalues $\lambda_1, \dots, \lambda_n$. Thus the last bullet point of 9.42 implies that $\det A = \lambda_1 \cdots \lambda_n$.

For the next result, think of each list $v_1, \dots, v_n$ of n vectors in $\mathbf{F}^n$ as a list of n-by-1 column vectors. The notation $\begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}$ then denotes the n-by-n square matrix whose k^{th} column is v_k for each $k = 1, \dots, n$.

9.45 *determinant is an alternating multilinear form*

Suppose that n is a positive integer. The map that takes a list $v_1, \dots, v_n$ of vectors in $\mathbf{F}^n$ to $\det \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}$ is an alternating n-linear form on $\mathbf{F}^n$.

Proof Let $e_1, \dots, e_n$ be the standard basis of $\mathbf{F}^n$ and suppose $v_1, \dots, v_n$ is a list of vectors in $\mathbf{F}^n$. Let $T \in \mathcal{L}(\mathbf{F}^n)$ be the operator such that $Te_k = v_k$ for $k = 1, \dots, n$. Thus T is the operator whose matrix with respect to $e_1, \dots, e_n$ is $\begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}$. Hence $\det \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix} = \det T$, by definition of the determinant of a matrix.

Let α be an alternating n-linear form on $\mathbf{F}^n$ such that $\alpha(e_1, \dots, e_n) = 1$. Then

$$\begin{aligned} \det \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix} &= \det T \\ &= (\det T)\, \alpha(e_1, \dots, e_n) \\ &= \alpha(Te_1, \dots, Te_n) \\ &= \alpha(v_1, \dots, v_n), \end{aligned}$$

where the third line follows from the definition of the determinant of an operator. The equation above shows that the map that takes a list of vectors $v_1, \dots, v_n$ in $\mathbf{F}^n$ to $\det \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}$ is the alternating n-linear form α on $\mathbf{F}^n$.

משימתנו הבאה היא להגדיר ולתת נוסחה עבור הדטרמיננטה של מטריצה ריבועית. לצורך כך, נשייך לכל מטריצה ריבועית אופרטור ואז נגדיר את הדטרמיננטה של המטריצה כדטרמיננטה של האופרטור המשויך.

9.43 הגדרה: דטרמיננטה של מטריצה, $\det A$

נניח ש־ n מספר שלם חיובי ו־ A היא מטריצה ריבועית $n \times n$ עם איברים ב־ $\mathbf{F}$. יהי $T \in \mathcal{L}(\mathbf{F}^n)$ האופרטור שהמטריצה שלו ביחס לבסיס הסטנדרטי של $\mathbf{F}^n$ שווה ל־ A. הדטרמיננטה של A, מסומנת על ידי $\det A$, מוגדרת על ידי $\det A = \det T$.

9.44 דוגמה: דטרמיננטות של מטריצות

- אם I היא מטריצת הזהות $n \times n$, אזי האופרטור המתאים מעל $\mathbf{F}^n$ הוא אופרטור הזהות I מעל $\mathbf{F}^n$. לפיכך הסעיף הראשון של 9.42 גורר שהדטרמיננטה של מטריצת הזהות היא 1.
- נניח ש־ A מטריצה אלכסונית עם $\lambda_1, \dots, \lambda_n$ על האלכסון. אזי לאופרטור המתאים מעל $\mathbf{F}^n$ יש את הבסיס הסטנדרטי של $\mathbf{F}^n$ כווקטורים עצמיים, עם ערכים עצמיים $\lambda_1, \dots, \lambda_n$. לפיכך הסעיף האחרון של 9.42 גורר ש־ $\det A = \lambda_1 \cdots \lambda_n$.

עבור המסקנה הבאה, חשבו על כל רשימה $v_1, \dots, v_n$ של n וקטורים ב־ $\mathbf{F}^n$ כרשימה של וקטורי עמודה $n \times 1$. הסימון $\begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}$ אם כן מסמן מטריצה ריבועית $n \times n$ שהעמודה ה־ k^{th} שלה היא v_k לכל $k = 1, \dots, n$.

9.45 דטרמיננטה היא תבנית מולטי־ליניארית מתחלפת

נניח ש־ n מספר שלם חיובי. ההעתקה השולחת רשימה של וקטורים $v_1, \dots, v_n$ ב־ $\mathbf{F}^n$ אל $\det \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}$ היא תבנית n־ליניארית מתחלפת מעל $\mathbf{F}^n$.

הוכחה יהי $e_1, \dots, e_n$ הבסיס הסטנדרטי של $\mathbf{F}^n$ ונניח ש־ $v_1, \dots, v_n$ רשימת וקטורים ב־ $\mathbf{F}^n$. יהי $T \in \mathcal{L}(\mathbf{F}^n)$ האופרטור כך ש־ $Te_k = v_k$ עבור $k = 1, \dots, n$. לפיכך T הוא האופרטור שהמטריצה שלו ביחס ל־ $e_1, \dots, e_n$ היא $\begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}$. לכן $\det \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix} = \det T$, לפי ההגדרה של הדטרמיננטה של מטריצה.

תהי α תבנית n־ליניארית מתחלפת מעל $\mathbf{F}^n$ כך ש־ $\alpha(e_1, \dots, e_n) = 1$. אזי

$$\begin{aligned} \det \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix} &= \det T \\ &= (\det T)\, \alpha(e_1, \dots, e_n) \\ &= \alpha(Te_1, \dots, Te_n) \\ &= \alpha(v_1, \dots, v_n), \end{aligned}$$

כאשר השורה השלישית נובעת מהגדרת הדטרמיננטה של אופרטור. המשוואה לעיל מראה שההעתקה השולחת רשימת וקטורים $v_1, \dots, v_n$ ב־ $\mathbf{F}^n$ אל $\det \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}$ היא התבנית ה־ n־ליניארית המתחלפת α מעל $\mathbf{F}^n$. ■

אלגברה ליניארית בדרך הישר, מהדורה רביעית, מחבר: שלדון אקסלר
Linear Algebra Done Right, fourth edition, by Sheldon Axler

The previous result has several important consequences. For example, it immediately implies that a matrix with two identical columns has determinant 0. We will come back to other consequences later, but for now we want to give a formula for the determinant of a square matrix. Recall that if A is a matrix, then $A_{j,k}$ denotes the entry in row j, column k of A.

9.46 *formula for determinant of a matrix*

Suppose that n is a positive integer and A is an n-by-n square matrix. Then

$$\det A = \sum_{(j_1,\dots,j_n)\in \operatorname{perm} n} (\operatorname{sign}(j_1,\dots,j_n)) A_{j_1,1}\cdots A_{j_n,n}.$$

Proof Apply 9.36 with $V = \mathbf{F}^n$ and $e_1,\dots,e_n$ the standard basis of $\mathbf{F}^n$ and α the alternating n-linear form on $\mathbf{F}^n$ that takes $v_1,\dots,v_n$ to $\det\begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}$ [see 9.45]. If each v_k is the kth column of A, then each $b_{j,k}$ in 9.36 equals $A_{j,k}$. Finally,

$$\alpha(e_1,\dots,e_n) = \det\begin{pmatrix} e_1 & \cdots & e_n \end{pmatrix} = \det I = 1.$$

Thus the formula in 9.36 becomes the formula stated in this result. ∎

9.47 example: *explicit formula for determinant*

- If A is a 2-by-2 matrix, then the formula in 9.46 becomes

$$\det A = A_{1,1}A_{2,2} - A_{2,1}A_{1,2}.$$

- If A is a 3-by-3 matrix, then the formula in 9.46 becomes

$$\begin{aligned}\det A = {} & A_{1,1}A_{2,2}A_{3,3} - A_{2,1}A_{1,2}A_{3,3} - A_{3,1}A_{2,2}A_{1,3} \\ & - A_{1,1}A_{3,2}A_{2,3} + A_{3,1}A_{1,2}A_{2,3} + A_{2,1}A_{3,2}A_{1,3}.\end{aligned}$$

The sum in the formula in 9.46 contains $n!$ terms. Because $n!$ grows rapidly as n increases, the formula in 9.46 is not a viable method to evaluate determinants even for moderately sized n. For example, 10! is over three million, and 100! is approximately 10^{158}, leading to a sum that the fastest computer cannot evaluate. We will soon see some results that lead to faster evaluations of determinants than direct use of the sum in 9.46.

9.48 *determinant of upper-triangular matrix*

Suppose that A is an upper-triangular matrix with $\lambda_1,\dots,\lambda_n$ on the diagonal. Then $\det A = \lambda_1\cdots\lambda_n$.

Proof If $(j_1,\dots,j_n) \in \operatorname{perm} n$ with $(j_1,\dots,j_n) \neq (1,\dots,n)$, then $j_k > k$ for some $k \in \{1,\dots,n\}$, which implies that $A_{j_k,k} = 0$. Thus the only permutation that can make a nonzero contribution to the sum in 9.46 is the permutation $(1,\dots,n)$. Because $A_{k,k} = \lambda_k$ for each $k = 1,\dots,n$, this implies that $\det A = \lambda_1\cdots\lambda_n$. ∎

למסקנה הקודמת יש מספר השלכות חשובות. לדוגמה, המסקנה גוררת באופן ישיר שהדטרמיננטה של מטריצה עם שתי עמודות זהות היא 0. מאוחר יותר עוד נשוב אל מספר השלכות נוספות, אך לעת עתה אנו מעוניינים להעניק נוסחה לדטרמיננטה של מטריצה ריבועית. נזכיר כי אם A מטריצה, אזי $A_{j,k}$ מסמן את האיבר בשורה j, עמודה k של A.

9.46 נוסחה עבור דטרמיננטה של מטריצה

נניח ש־ n מספר שלם חיובי ו־ A מטריצה ריבועית $n \times n$. אזי

$$\det A = \sum_{(j_1,\dots,j_n)\in \operatorname{perm} n} (\operatorname{sign}(j_1,\dots,j_n)) A_{j_1,1} \cdots A_{j_n,n}.$$

הוכחה נפעיל את 9.36 עם $V = \mathbf{F}^n$ ו־ $e_1, \dots, e_n$ הבסיס הסטנדרטי של $\mathbf{F}^n$ ו־ α התבנית n־ליניארית המתחלפת מעל $\mathbf{F}^n$ השולחת את $v_1, \dots, v_n$ אל $\det \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}$ [ראו 9.45]. אם כל v_k הוא העמודה ה־ k^{th} של A, אזי כל $b_{j,k}$ ב־ 9.36 שווה ל־ $A_{j,k}$. לסיום,

$$\alpha(e_1, \dots, e_n) = \det \begin{pmatrix} e_1 & \cdots & e_n \end{pmatrix} = \det I = 1.$$

לפיכך הנוסחה ב־ 9.36 הופכת לנוסחה שנקבעה במסקנה זו. ■

9.47 דוגמה: נוסחה מפורשת עבור דטרמיננטה

- אם A מטריצה 2×2, אזי הנוסחה ב־ 9.46 הופכת ל־
$$\det A = A_{1,1}A_{2,2} - A_{2,1}A_{1,2}.$$
- אם A מטריצה 3×3, אזי הנוסחה ב־ 9.46 הופכת ל־
$$\begin{aligned}\det A =& A_{1,1}A_{2,2}A_{3,3} - A_{2,1}A_{1,2}A_{3,3} - A_{3,1}A_{2,2}A_{1,3}\\ &- A_{1,1}A_{3,2}A_{2,3} + A_{3,1}A_{1,2}A_{2,3} + A_{2,1}A_{3,2}A_{1,3}.\end{aligned}$$

הסכום בנוסחה ב־ 9.46 מכיל $n!$ איברים. מאחר ש־ $n!$ גדל במהירות ככל ש־ n גדל, הנוסחה ב־ 9.46 אינה שיטה בת־קיימא לחישוב דטרמיננטות אפילו עבור n בסדר גודל צנוע. לדוגמה, $10!$ הוא מעל לשלושה מיליון, ו־ $100!$ הוא בערך 10^{158}, המוביל לסכום שאפילו המחשב המהיר ביותר אינו יכול לחשב. בקרוב נכיר מסקנות שיובילו אותנו לשיטות חישוב דטרמיננטות מהירות יותר מהחישוב הישיר של הסכום ב־ 9.46.

9.48 דטרמיננטה של מטריצה משולשית עליונה

נניח ש־ A מטריצה משולשית עליונה עם $\lambda_1, \dots, \lambda_n$ על האלכסון. אזי $\det A = \lambda_1 \cdots \lambda_n$.

הוכחה אם $(j_1, \dots, j_n) \in \operatorname{perm} n$ עם $(j_1, \dots, j_n) \neq (1, \dots, n)$, אזי עבור $k \in \{1, \dots, n\}$ כלשהו, מה שגורר ש־ $j_k > k$. לפיכך התמורה היחידה שיכולה להוסיף ערך השונה מאפס לסכום ב־ 9.46 היא התמורה $(1, \dots, n)$. מאחר ש־ $A_{k,k} = \lambda_k$ לכל $k = 1, \dots, n$, מכאן נובע ש־ $\det A = \lambda_1 \cdots \lambda_n$. ■

Properties of Determinants

Our definition of the determinant leads to the following magical proof that the determinant is multiplicative.

9.49 *determinant is multiplicative*

(a) Suppose $S, T \in \mathcal{L}(V)$. Then $\det(ST) = (\det S)(\det T)$.

(b) Suppose A and B are square matrices of the same size. Then

$$\det(AB) = (\det A)(\det B).$$

Proof

(a) Let $n = \dim V$. Suppose $\alpha \in V_{\text{alt}}^{(n)}$ and $v_1, \dots, v_n \in V$. Then

$$\begin{aligned}\alpha_{ST}(v_1, \dots, v_n) &= \alpha(STv_1, \dots, STv_n)\\ &= (\det S)\alpha(Tv_1, \dots, Tv_n)\\ &= (\det S)(\det T)\alpha(v_1, \dots, v_n),\end{aligned}$$

where the first equation follows from the definition of α_{ST}, the second equation follows from the definition of $\det S$, and the third equation follows from the definition of $\det T$. The equation above implies that $\det(ST) = (\det S)(\det T)$.

(b) Let $S, T \in \mathcal{L}(\mathbf{F}^n)$ be such that $\mathcal{M}(S) = A$ and $\mathcal{M}(T) = B$, where all matrices of operators in this proof are with respect to the standard basis of $\mathbf{F}^n$. Then $\mathcal{M}(ST) = \mathcal{M}(S)\mathcal{M}(T) = AB$ (see 3.43). Thus

$$\det(AB) = \det(ST) = (\det S)(\det T) = (\det A)(\det B),$$

where the second equality comes from the result in (a). ∎

The determinant of an operator determines whether the operator is invertible.

9.50 *invertible* $\iff$ *nonzero determinant*

An operator $T \in \mathcal{L}(V)$ is invertible if and only if $\det T \neq 0$. Furthermore, if T is invertible, then $\det(T^{-1}) = \frac{1}{\det T}$.

Proof First suppose T is invertible. Thus $TT^{-1} = I$. Now 9.49 implies that

$$1 = \det I = \det(TT^{-1}) = (\det T)\big(\det(T^{-1})\big).$$

Hence $\det T \neq 0$ and $\det(T^{-1})$ is the multiplicative inverse of $\det T$.

To prove the other direction, now suppose $\det T \neq 0$. Suppose $v \in V$ and $v \neq 0$. Let $v, e_2, \dots, e_n$ be a basis of V and let $\alpha \in V_{\text{alt}}^{(n)}$ be such that $\alpha \neq 0$. Then $\alpha(v, e_2, \dots, e_n) \neq 0$ (by 9.39). Now

$$\alpha(Tv, Te_2, \dots, Te_n) = (\det T)\alpha(v, e_2, \dots, e_n) \neq 0.$$

Thus $Tv \neq 0$. Hence T is invertible. ∎

תכונות של דטרמיננטות

הגדרת הדטרמיננטה שלנו מובילה להוכחה הקסומה הבאה לכך שהדטרמיננטה כפלית.

9.49 דטרמיננטה היא כפלית

א. נניח ש־ $S, T \in \mathcal{L}(V)$. אזי $\det(ST) = (\det S)(\det T)$.

ב. נניח ש־ A ו־ B מטריצות ריבועיות באותו הגודל. אזי $\det(AB) = (\det A)(\det B)$.

הוכחה

א. יהי $n = \dim V$. נניח ש־ $\alpha \in V_{\text{alt}}^{(n)}$ ו־ $v_1, \dots, v_n \in V$. אזי

$$\begin{aligned}\alpha_{ST}(v_1, \dots, v_n) &= \alpha(STv_1, \dots, STv_n)\\ &= (\det S)\alpha(Tv_1, \dots, Tv_n)\\ &= (\det S)(\det T)\alpha(v_1, \dots, v_n),\end{aligned}$$

כאשר המשוואה הראשונה נובעת מהגדרת α_{ST}, המשוואה השנייה נובעת מהגדרת $\det S$, והמשוואה השלישית נובעת מהגדרת $\det T$. המשוואה לעיל גוררת ש־ $\det(ST) = (\det S)(\det T)$.

ב. יהיו $S, T \in \mathcal{L}(\mathbf{F}^n)$ כך ש־ $\mathcal{M}(S) = A$ ו־ $\mathcal{M}(T) = B$, כאשר כל המטריצות של האופרטורים בהוכחה זו הן ביחס לבסיס הסטנדרטי של $\mathbf{F}^n$. אזי $\mathcal{M}(ST) = AB$ (ראו 3.43). לפיכך $\mathcal{M}(S)\mathcal{M}(T)$

$$\det(AB) = \det(ST) = (\det S)(\det T) = (\det A)(\det B),$$

כאשר השוויון השני נובע מהמסקנה בסעיף (א). ∎

הדטרמיננטה של אופרטור קובעת האם האופרטור הפיך.

9.50 הפיך $\iff$ דטרמיננטה שונה מאפס

אופרטור $T \in \mathcal{L}(V)$ הוא הפיך אם־ורק־אם $\det T \neq 0$. יתר על כן, אם T הפיך, אזי $\det(T^{-1}) = \frac{1}{\det T}$.

הוכחה תחילה נניח כי T הפיך. לפיכך $TT^{-1} = I$. כעת 9.49 גורר ש־

$$1 = \det I = \det(TT^{-1}) = (\det T)\big(\det(T^{-1})\big).$$

לכן $\det T \neq 0$ ו־ $\det(T^{-1})$ הוא האיבר ההופכי לכפל של $\det T$.

להוכחת הכיוון האחר, כעת נניח ש־ $\det T \neq 0$. נניח ש־ $v \in V$ ו־ $v \neq 0$. תהי $v, e_2, \dots, e_n$ בסיס של V ותהי $\alpha \in V_{\text{alt}}^{(n)}$ כך ש־ $\alpha \neq 0$. אזי $\alpha(v, e_2, \dots, e_n) \neq 0$ (לפי 9.39). כעת

$$\alpha(Tv, Te_2, \dots, Te_n) = (\det T)\alpha(v, e_2, \dots, e_n) \neq 0.$$

לפיכך $Tv \neq 0$. לכן T הפיך. ∎

An n-by-n matrix A is invertible (see 3.80 for the definition of an invertible matrix) if and only if the operator on $\mathbf{F}^n$ associated with A (via the standard basis of $\mathbf{F}^n$) is invertible. Thus the previous result shows that a square matrix A is invertible if and only if $\det A \neq 0$.

9.51 *eigenvalues and determinants*

Suppose $T \in \mathcal{L}(V)$ and $\lambda \in \mathbf{F}$. Then λ is an eigenvalue of T if and only if $\det(\lambda I - T) = 0$.

Proof The number λ is an eigenvalue of T if and only if $T - \lambda I$ is not invertible (see 5.7), which happens if and only if $\lambda I - T$ is not invertible, which happens if and only if $\det(\lambda I - T) = 0$ (by 9.50). ∎

Suppose $T \in \mathcal{L}(V)$ and $S \colon W \to V$ is an invertible linear map. To prove that $\det(S^{-1}TS) = \det T$, we could try to use 9.49 and 9.50, writing

$$\begin{aligned}\det(S^{-1}TS) &= (\det S^{-1})(\det T)(\det S)\\ &= \det T.\end{aligned}$$

That proof works if $W = V$, but if $W \neq V$ then it makes no sense because the determinant is defined only for linear maps from a vector space to itself, and S maps W to V, making $\det S$ undefined. The proof given below works around this issue and is valid when $W \neq V$.

9.52 *determinant is a similarity invariant*

Suppose $T \in \mathcal{L}(V)$ and $S \colon W \to V$ is an invertible linear map. Then

$$\det(S^{-1}TS) = \det T.$$

Proof Let $n = \dim W = \dim V$. Suppose $\tau \in W_{\text{alt}}^{(n)}$. Define $\alpha \in V_{\text{alt}}^{(n)}$ by

$$\alpha(v_1, \dots, v_n) = \tau(S^{-1}v_1, \dots, S^{-1}v_n)$$

for $v_1, \dots, v_n \in V$. Suppose $w_1, \dots, w_n \in W$. Then

$$\begin{aligned}\tau_{S^{-1}TS}(w_1, \dots, w_n) &= \tau(S^{-1}TSw_1, \dots, S^{-1}TSw_n)\\ &= \alpha(TSw_1, \dots, TSw_n)\\ &= \alpha_T(Sw_1, \dots, Sw_n)\\ &= (\det T)\alpha(Sw_1, \dots, Sw_n)\\ &= (\det T)\tau(w_1, \dots, w_n).\end{aligned}$$

The equation above and the definition of the determinant of the operator $S^{-1}TS$ imply that $\det(S^{-1}TS) = \det T$. ∎

A מטריצה $n \times n$ הפיכה (ראו 3.80 עבור ההגדרה של מטריצה הפיכה) אם־ורק־אם האופרטור מעל $\mathbf{F}^n$ המשויך למטריצה A (באמצעות הבסיס הסטנדרטי של $\mathbf{F}^n$) הפיך. לפיכך המסקנה הקודמת מראה שמטריצה ריבועית A הפיכה אם־ורק־אם $\det A \neq 0$.

9.51 ערכים עצמיים ודטרמיננטות

נניח ש־ $T \in \mathcal{L}(V)$ ו־ $\lambda \in \mathbf{F}$. אזי λ ערך עצמי של T אם־ורק־אם $\det(\lambda I - T) = 0$.

הוכחה המספר λ הוא ערך עצמי של T אם־ורק־אם $T - \lambda I$ אינו הפיך (ראו 5.7), מה שקורה אם־ורק־אם $\det(\lambda I - T) = 0$ אינו הפיך, מה שקורה אם־ורק־אם $\det(\lambda I - T) = 0$ (לפי 9.50). ■

נניח ש־ $T \in \mathcal{L}(V)$ ו־ $S : W \to V$ העתקה ליניארית הפיכה. כדי להוכיח כי $\det(S^{-1}TS) = \det T$, נוכל לנסות להשתמש ב־ 9.49 ו־ 9.50 ולכתוב

$$\begin{aligned}\det(S^{-1}TS) &= (\det S^{-1})(\det T)(\det S)\\ &= \det T.\end{aligned}$$

הוכחה זו עובדת אם $W = V$, אך אם $W \neq V$ היא חסרת משמעות מאחר שהדטרמיננטה מוגדרת רק עבור העתקות ליניאריות ממרחב וקטורי לעצמו, בעוד S מעתיקה את W אל V, ההופכת את $\det S$ לבלתי מוגדרת. ההוכחה הניתנת להלן עוקפת את הבעיה הזו ותקפה גם כאשר $W \neq V$.

9.52 הדטרמיננטה שמורה תחת דִּמְיוֹן

נניח ש־ $T \in \mathcal{L}(V)$ ו־ $S : W \to V$ העתקה ליניארית הפיכה. אזי
$$\det(S^{-1}TS) = \det T.$$

הוכחה יהי $n = \dim W = \dim V$. נניח ש־ $\tau \in W_{\text{alt}}^{(n)}$. נגדיר $\alpha \in V_{\text{alt}}^{(n)}$ על ידי
$$\alpha(v_1, \dots, v_n) = \tau(S^{-1}v_1, \dots, S^{-1}v_n)$$
עבור $v_1, \dots, v_n \in V$. נניח ש־ $w_1, \dots, w_n \in W$. אזי

$$\begin{aligned}\tau_{S^{-1}TS}(w_1, \dots, w_n) &= \tau(S^{-1}TSw_1, \dots, S^{-1}TSw_n)\\ &= \alpha(TSw_1, \dots, TSw_n)\\ &= \alpha_T(Sw_1, \dots, Sw_n)\\ &= (\det T)\alpha(Sw_1, \dots, Sw_n)\\ &= (\det T)\tau(w_1, \dots, w_n).\end{aligned}$$

מהמשוואה לעיל ומהגדרת הדטרמיננטה של האופרטור $S^{-1}TS$ אנו מסיקים כי $\det(S^{-1}TS) = \det T$. ■

For the special case in which $V = \mathbf{F}^n$ and $e_1, \dots, e_n$ is the standard basis of $\mathbf{F}^n$, the next result is true by the definition of the determinant of a matrix. The left side of the equation in the next result does not depend on a choice of basis, which means that the right side is independent of the choice of basis.

9.53 *determinant of operator equals determinant of its matrix*

Suppose $T \in \mathcal{L}(V)$ and $e_1, \dots, e_n$ is a basis of V. Then

$$\det T = \det \mathcal{M}(T, (e_1, \dots, e_n)).$$

Proof Let $f_1, \dots, f_n$ be the standard basis of $\mathbf{F}^n$. Let $S : \mathbf{F}^n \to V$ be the linear map such that $Sf_k = e_k$ for each $k = 1, \dots, n$. Thus $\mathcal{M}(S, (f_1, \dots, f_n), (e_1, \dots, e_n))$ and $\mathcal{M}(S^{-1}, (e_1, \dots, e_n), (f_1, \dots, f_n))$ both equal the n-by-n identity matrix. Hence

$$\mathcal{M}(S^{-1}TS, (f_1, \dots, f_n)) = \mathcal{M}(T, (e_1, \dots, e_n)), \tag{9.54}$$

as follows from two applications of 3.43. Thus

$$\begin{aligned} \det T &= \det(S^{-1}TS) \\ &= \det \mathcal{M}(S^{-1}TS, (f_1, \dots, f_n)) \\ &= \det \mathcal{M}(T, (e_1, \dots, e_n)), \end{aligned}$$

where the first line comes from 9.52, the second line comes from the definition of the determinant of a matrix, and the third line follows from 9.54. ∎

The next result gives a more intuitive way to think about determinants than the definition or the formula in 9.46. We could make the characterization in the result below the definition of the determinant of an operator on a finite-dimensional complex vector space, with the current definition then becoming a consequence of that definition.

9.55 *if $\mathbf{F} = \mathbf{C}$, then determinant equals product of eigenvalues*

Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$. Then $\det T$ equals the product of the eigenvalues of T, with each eigenvalue included as many times as its multiplicity.

Proof There is a basis of V with respect to which T has an upper-triangular matrix with the diagonal entries of the matrix consisting of the eigenvalues of T, with each eigenvalue included as many times as its multiplicity—see 8.37. Thus 9.53 and 9.48 imply that $\det T$ equals the product of the eigenvalues of T, with each eigenvalue included as many times as its multiplicity. ∎

As the next result shows, the determinant interacts nicely with the transpose of a square matrix, with the dual of an operator, and with the adjoint of an operator on an inner product space.

עבור המקרה הפרטי שבו $V = \mathbf{F}^n$ ו־ $e_1, \dots, e_n$ הוא הבסיס הסטנדרטי של $\mathbf{F}^n$, המסקנה הבאה מתקיימת מתוקף הגדרת הדטרמיננטה של מטריצה. האגף השמאלי של המשוואה במסקנה הבאה אינו תלוי בבחירת בסיס, ומשמעות הדבר היא שהאגף הימני אינו תלוי בבחירת בסיס.

9.53 דטרמיננטה של אופרטור שווה לדטרמיננטה של המטריצה שלו

נניח ש־ $T \in \mathcal{L}(V)$ ו־ $e_1, \dots, e_n$ בסיס של V. אזי
$$\det T = \det \mathcal{M}(T, (e_1, \dots, e_n)).$$

הוכחה תהי $f_1, \dots, f_n$ הבסיס הסטנדרטי של $\mathbf{F}^n$. תהי $S \colon \mathbf{F}^n \to V$ ההעתקה הליניארית כך ש־ $Sf_k = e_k$ לכל $k = 1, \dots, n$. לפיכך $\mathcal{M}(S, (f_1, \dots, f_n), (e_1, \dots, e_n))$ שתיהן שוות למטריצת הזהות $n \times n$. לכן $\mathcal{M}(S^{-1}, (e_1, \dots, e_n), (f_1, \dots, f_n))$ ו־

$$\mathcal{M}(S^{-1}TS, (f_1, \dots, f_n)) = \mathcal{M}(T, (e_1, \dots, e_n)), \tag{9.54}$$

כפי שנגרר משני יישומים של 3.43. לפיכך

$$\begin{aligned}\det T &= \det(S^{-1}TS)\\ &= \det \mathcal{M}(S^{-1}TS, (f_1, \dots, f_n))\\ &= \det \mathcal{M}(T, (e_1, \dots, e_n)),\end{aligned}$$

כאשר השורה הראשונה נובעת מ־ 9.52, השורה השנייה נובעת מהגדרת הדטרמיננטה של מטריצה, והשורה השלישית נובעת מ־ 9.54.

המסקנה הבאה מעניקה דרך לחשוב על דטרמיננטות באופן אינטואיטיבי יותר מאשר ההגדרה או הנוסחה ב־ 9.46. יכולנו להציג את האפיון במסקנה להלן כהגדרה של הדטרמיננטה של אופרטור מעל מרחב וקטורי מרוכב מממד סופי, ובמקרה זה ההגדרה הנוכחית הייתה הופכת לתוצאה של הגדרה זו.

9.55 אם $\mathbf{F} = \mathbf{C}$, אזי הדטרמיננטה שווה למכפלת הערכים העצמיים

נניח ש־ $\mathbf{F} = \mathbf{C}$ ו־ $T \in \mathcal{L}(V)$. אזי $\det T$ שווה למכפלת הערכים העצמיים של T, כאשר כל ערך עצמי נכלל במכפלה מספר פעמים השווה לריבויו.

הוכחה קיים בסיס של V אשר ביחס אליו יש ל־ T מטריצה משולשית עליונה כאשר איברי אלכסון המטריצה מורכבים מהערכים העצמיים של T, כאשר מספר המופעים של כל ערך עצמי שווה לריבויו—ראו 8.37. לפיכך 9.53 ו־ 9.48 גוררים ש־ $\det T$ שווה למכפלה של הערכים העצמיים של T, כאשר מספר המופעים של כל ערך עצמי שווה לריבויו.

כפי שהמסקנה הבאה מראה, הדטרמיננטה עובדת יפה עם המשוחלפת של מטריצה ריבועית, עם האופרטור הדואלי ועם האופרטור הצמוד של אופרטור מעל מרחב מכפלה פנימית.

9.56 *determinant of transpose, dual, or adjoint*

(a) Suppose A is a square matrix. Then $\det A^t = \det A$.

(b) Suppose $T \in \mathcal{L}(V)$. Then $\det T' = \det T$.

(c) Suppose V is an inner product space and $T \in \mathcal{L}(V)$. Then

$$\det(T^*) = \overline{\det T}.$$

Proof

(a) Let n be a positive integer. Define $\alpha : (\mathbf{F}^n)^n \to \mathbf{F}$ by

$$\alpha(v_1, \dots, v_n) = \det\Big(\begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}^t\Big)$$

for all $v_1, \dots, v_n \in \mathbf{F}^n$. The formula in 9.46 for the determinant of a matrix shows that α is an n-linear form on $\mathbf{F}^n$.

Suppose $v_1, \dots, v_n \in \mathbf{F}^n$ and $v_j = v_k$ for some $j \neq k$. If B is an n-by-n matrix, then $\begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}^t B$ cannot equal the identity matrix because row j and row k of $\begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}^t B$ are equal. Thus $\begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}^t$ is not invertible, which implies that $\alpha(v_1, \dots, v_n) = 0$. Hence α is an alternating n-linear form on $\mathbf{F}^n$.

Note that α applied to the standard basis of $\mathbf{F}^n$ equals 1. Because the vector space of alternating n-linear forms on $\mathbf{F}^n$ has dimension one (by 9.37), this implies that α is the determinant function. Thus (a) holds.

(b) The equation $\det T' = \det T$ follows from (a) and 9.53 and 3.132.

(c) Pick an orthonormal basis of V. The matrix of T^* with respect to that basis is the conjugate transpose of the matrix of T with respect to that basis (by 7.9). Thus 9.53, 9.46, and (a) imply that $\det(T^*) = \overline{\det T}$. ∎

9.57 *helpful results in evaluating determinants*

(a) If either two columns or two rows of a square matrix are equal, then the determinant of the matrix equals 0.

(b) Suppose A is a square matrix and B is the matrix obtained from A by swapping either two columns or two rows. Then $\det A = -\det B$.

(c) If one column or one row of a square matrix is multiplied by a scalar, then the value of the determinant is multiplied by the same scalar.

(d) If a scalar multiple of one column of a square matrix is added to another column, then the value of the determinant is unchanged.

(e) If a scalar multiple of one row of a square matrix is added to another row, then the value of the determinant is unchanged.

9.56 דטרמיננטה של משוחלפת, דואלי, או אופרטור צמוד

א. נניח ש־ A מטריצה ריבועית. אזי $\det A^{\mathrm{t}} = \det A$.

ב. נניח ש־ $T \in \mathcal{L}(V)$. אזי $\det T' = \det T$.

ג. נניח ש־ V מרחב מכפלה פנימית ו־ $T \in \mathcal{L}(V)$. אזי $\det(T^*) = \overline{\det T}$.

הוכחה

א. יהי n מספר שלם חיובי. נגדיר $\alpha : (\mathbf{F}^n)^n \to \mathbf{F}$ על ידי

$$\alpha(v_1, \dots, v_n) = \det\Big(\begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}^{\mathrm{t}}\Big)$$

לכל $v_1, \dots, v_n \in \mathbf{F}^n$. הנוסחה ב־ 9.46 עבור הדטרמיננטה של מטריצה מראה כי α היא תבנית n־ליניארית מעל $\mathbf{F}^n$.

נניח ש־ $v_1, \dots, v_n \in \mathbf{F}^n$ ו־ $v_j = v_k$ עבור $j \neq k$ כלשהם. אם B מטריצה $n \times n$, אזי $\begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}^{\mathrm{t}} B$ אינה יכולה להיות שווה למטריצת הזהות מאחר ששורה j ושורה k של B $\begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}^{\mathrm{t}}$ שוות. לפיכך $\begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}$ אינה הפיכה, מה שגורר ש־ $\alpha(v_1, \dots, v_n) = 0$. לכן α היא תבנית n־ליניארית מתחלפת מעל $\mathbf{F}^n$.

שימו לב ש־ α מופעלת על הבסיס הסטנדרטי של $\mathbf{F}^n$ שווה 1. מאחר שהמרחב הווקטורי של תבניות n־ליניאריות מתחלפות מעל $\mathbf{F}^n$ הוא חד־ממדי (לפי 9.37), מכאן נובע ש־ α היא פונקציית הדטרמיננטה. לפיכך (א) מתקיים.

ב. המשוואה $\det T' = \det T$ נובעת מ־ (א) ו־ 9.53 ו־ 3.132.

ג. נבחר בסיס אורתונורמלי של V. המטריצה של T^* ביחס לבסיס זה היא הצמודה־משוחלפת של המטריצה של T ביחס לבסיס זה (לפי 7.9). לפיכך $\det(T^*) = \overline{\det T}$, ו־ (א) גוררים ש־ 9.53, 9.46. ■

9.57 תוצאות שימושיות בחישוב דטרמיננטות

א. אם שתי עמודות או שתי שורות של מטריצה ריבועית שוות, אזי הדטרמיננטה של המטריצה שווה 0.

ב. נניח ש־ A מטריצה ריבועית ו־ B המטריצה שהושגה מ־ A באמצעות החלפת שתי עמודות או שתי שורות. אזי $\det A = -\det B$.

ג. אם עמודה אחת או שורה אחת של מטריצה ריבועית מוכפלת בסקלר, אזי ערך הדטרמיננטה מוכפל על ידי אותו סקלר.

ד. אם נחבר כפולת סקלר של עמודה אחת של מטריצה ריבועית לעמודה אחרת, אזי ערך הדטרמיננטה אינו משתנה.

ה. אם נחבר כפולת סקלר של שורה אחת של מטריצה ריבועית לשורה אחרת, אזי ערך הדטרמיננטה אינו משתנה.

Proof All the assertions in this result follow from the result that the maps $v_1, \dots, v_n \mapsto \det \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}$ and $v_1, \dots, v_n \mapsto \det \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}^t$ are both alternating n-linear forms on $\mathbf{F}^n$ [see 9.45 and 9.56(a)].

For example, to prove (d) suppose $v_1, \dots, v_n \in \mathbf{F}^n$ and $c \in \mathbf{F}$. Then

$$\begin{aligned}
&\det \begin{pmatrix} v_1 + cv_2 & v_2 & \cdots & v_n \end{pmatrix} \\
&\quad = \det \begin{pmatrix} v_1 & v_2 & \cdots & v_n \end{pmatrix} + c \det \begin{pmatrix} v_2 & v_2 & v_3 & \cdots & v_n \end{pmatrix} \\
&\quad = \det \begin{pmatrix} v_1 & v_2 & \cdots & v_n \end{pmatrix},
\end{aligned}$$

where the first equation follows from the multilinearity property and the second equation follows from the alternating property. The equation above shows that adding a multiple of the second column to the first column does not change the value of the determinant. The same conclusion holds for any two columns. Thus (d) holds.

The proof of (e) follows from (d) and from 9.56(a). The proofs of (a), (b), and (c) use similar tools and are left to the reader. ∎

For matrices whose entries are concrete numbers, the result above leads to a much faster way to evaluate the determinant than direct application of the formula in 9.46. Specifically, apply the Gaussian elimination procedure of swapping rows [by 9.57(b), this changes the determinant by a factor of -1], multiplying a row by a nonzero constant [by 9.57(c), this changes the determinant by the same constant], and adding a multiple of one row to another row [by 9.57(e), this does not change the determinant] to produce an upper-triangular matrix, whose determinant is the product of the diagonal entries (by 9.48). If your software keeps track of the number of row swaps and of the constants used when multiplying a row by a constant, then the determinant of the original matrix can be computed.

Because a number $\lambda \in \mathbf{F}$ is an eigenvalue of an operator $T \in \mathcal{L}(V)$ if and only if $\det(\lambda I - T) = 0$ (by 9.51), you may be tempted to think that one way to find eigenvalues quickly is to choose a basis of V, let $A = \mathcal{M}(T)$, evaluate $\det(\lambda I - A)$, and then solve the equation $\det(\lambda I - A) = 0$ for λ. However, that procedure is rarely efficient, except when $\dim V = 2$ (or when $\dim V$ equals 3 or 4 if you are willing to use the cubic or quartic formulas). One problem is that the procedure described in the paragraph above for evaluating a determinant does not work when the matrix includes a symbol (such as the λ in $\lambda I - A$). This problem arises because decisions need to be made in the Gaussian elimination procedure about whether certain quantities equal 0, and those decisions become complicated in expressions involving a symbol λ.

Recall that an operator on a finite-dimensional inner product space is unitary if it preserves norms (see 7.51 and the paragraph following it). Every eigenvalue of a unitary operator has absolute value 1 (by 7.54). Thus the product of the eigenvalues of a unitary operator has absolute value 1. Hence (at least in the case $\mathbf{F} = \mathbf{C}$) the determinant of a unitary operator has absolute value 1 (by 9.55). The next result gives a proof that works without the assumption that $\mathbf{F} = \mathbf{C}$.

הוכחה הטענות במסקנה זו נובעות מהתוצאה כי $v_1, \dots, v_n \mapsto \det \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}$ וכי $v_1, \dots, v_n \mapsto \det \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}^t$ שתיהן תבניות n-ליניאריות מתחלפות מעל $\mathbf{F}^n$ [ראו 9.45 ו־9.56(א)].

לדוגמה, להוכחת (ד) נניח ש־ $v_1, \dots, v_n \in \mathbf{F}^n$ ו־ $c \in \mathbf{F}$. אזי

$$\begin{aligned} &\det \begin{pmatrix} v_1 + cv_2 & v_2 & \cdots & v_n \end{pmatrix} \\ &= \det \begin{pmatrix} v_1 & v_2 & \cdots & v_n \end{pmatrix} + c \det \begin{pmatrix} v_2 & v_2 & v_3 & \cdots & v_n \end{pmatrix} \\ &= \det \begin{pmatrix} v_1 & v_2 & \cdots & v_n \end{pmatrix}, \end{aligned}$$

כאשר המשוואה הראשונה נובעת מהתכונה המולטי־ליניארית והמשוואה השנייה נובעת מהתכונה המתחלפת. המשוואה לעיל מראה שחיבור כפולה של העמודה השנייה לעמודה הראשונה אינו משנה את ערך הדטרמיננטה. אותה מסקנה מתקיימת עבור כל שתי עמודות. לפיכך (ד) מתקיים.

ההוכחה של (ה) נובעת מ־ (ד) ומ־ 9.56(א). ההוכחות של (א), (ב) ו־ (ג) משתמשות בכלים דומים ומושארות לקוראים. ∎

עבור מטריצות שאיבריהן מספרים קונקרטיים, המסקנה לעיל מובילה לשיטה לחישוב הדטרמיננטה שמהירה באופן משמעותי מן היישום הישיר של הנוסחה ב־ 9.46. בפרט, הפעילו את שיטת החילוץ של גאוס: החלפת שורות [לפי 9.57(ב), הדבר משנה את הדטרמיננטה בפקטור של -1], הכפלת שורה בקבוע השונה מאפס [לפי 9.57(ג), הדטרמיננטה מוכפלת באותו הקבוע], והוספת כפולת שורה אחת לשורה אחרת [לפי 9.57(ה), אינה משנה את הדטרמיננטה] — להפקת מטריצה משולשית עליונה, שהדטרמיננטה שלה היא מכפלת איברי האלכסון (לפי 9.48). אם התוכנה שבה אתם משתמשים מבצעת מעקב אחרי מספר השורות שהוחלפו והקבועים שבהם נעשה שימוש להכפלת שורות, אזי חישוב הדטרמיננטה של המטריצה המקורית בר השגה.

מאחר שמספר $\lambda \in \mathbf{F}$ הוא ערך עצמי של אופרטור $T \in \mathcal{L}(V)$ אם־ורק־אם $\det(\lambda I - T) = 0$ (לפי 9.51), אתם עלולים להתפתות לחשוב שדרך מהירה למצוא ערכים עצמיים היא לבחור בסיס של V, לסמן $A = \mathcal{M}(T)$, לחשב את $\det(\lambda I - A)$, ואז לפתור את המשוואה $\det(\lambda I - A) = 0$ עבור λ. אולם הליך זה יעיל רק לעיתים נדירות, מלבד כאשר $\dim V = 2$ (או כאשר $\dim V$ שווה 3 או 4 אם ברצונכם להשתמש בנוסחאות למציאת שורשים למשוואות ממעלה שלישית או רביעית). בעיה אחת היא שההליך לחישוב דטרמיננטה שתואר בפסקה שלעיל אינו עובד כאשר המטריצה מכילה סמל (כמו ה־ λ ב־ $\lambda I - A$). בעיה זו נובעת מכך שבשיטת החילוץ של גאוס יש להכריע אם גדלים מסוימים שווים 0, והכרעות אלה נעשות מסובכות בביטויים המערבים סמל λ.

נזכיר כי אופרטור מעל מרחב מכפלה פנימית מממד סופי הוא אוניטרי אם הוא משמר נורמות (ראו 7.51 והפסקה שאחריה). לכל ערך עצמי של אופרטור אוניטרי יש ערך מוחלט 1 (לפי 7.54). לפיכך למכפלת הערכים העצמיים של אופרטור אוניטרי יש ערך מוחלט 1. לכן (לכל הפחות במקרה של $\mathbf{F} = \mathbf{C}$) לדטרמיננטה של אופרטור אוניטרי יש ערך מוחלט 1 (לפי 9.55). המסקנה הבאה מספקת הוכחה התקפה ללא ההנחה ש־ $\mathbf{F} = \mathbf{C}$.

9.58 *every unitary operator has determinant with absolute value 1*

Suppose V is an inner product space and $S \in \mathcal{L}(V)$ is a unitary operator. Then $|\det S| = 1$.

Proof Because S is unitary, $I = S^*S$ (see 7.53). Thus

$$1 = \det(S^*S) = (\det S^*)(\det S) = \overline{(\det S)}(\det S) = |\det S|^2,$$

where the second equality comes from 9.49(a) and the third equality comes from 9.56(c). The equation above implies that $|\det S| = 1$. ∎

The determinant of a positive operator on an inner product space meshes well with the analogy that such operators correspond to the nonnegative real numbers.

9.59 *every positive operator has nonnegative determinant*

Suppose V is an inner product space and $T \in \mathcal{L}(V)$ is a positive operator. Then $\det T \geq 0$.

Proof By the spectral theorem (7.29 or 7.31), V has an orthonormal basis consisting of eigenvectors of T. Thus by the last bullet point of 9.42, $\det T$ equals a product of the eigenvalues of T, possibly with repetitions. Each eigenvalue of T is a nonnegative number (by 7.38). Thus we conclude that $\det T \geq 0$. ∎

Suppose V is an inner product space and $T \in \mathcal{L}(V)$. Recall that the list of nonnegative square roots of the eigenvalues of T^*T (each included as many times as its multiplicity) is called the list of singular values of T (see Section 7E).

9.60 $|\det T|$ = *product of singular values of* T

Suppose V is an inner product space and $T \in \mathcal{L}(V)$. Then

$$|\det T| = \sqrt{\det(T^*T)} = \text{product of singular values of } T.$$

Proof We have

$$|\det T|^2 = \overline{(\det T)}(\det T) = \big(\det(T^*)\big)(\det T) = \det(T^*T),$$

where the middle equality comes from 9.56(c) and the last equality comes from 9.49(a). Taking square roots of both sides of the equation above shows that $|\det T| = \sqrt{\det(T^*T)}$.

Let $s_1, \ldots, s_n$ denote the list of singular values of T. Thus $s_1^2, \ldots, s_n^2$ is the list of eigenvalues of T^*T (with appropriate repetitions), corresponding to an orthonormal basis of V consisting of eigenvectors of T^*T. Hence the last bullet point of 9.42 implies that

$$\det(T^*T) = s_1^2 \cdots s_n^2.$$

Thus $|\det T| = s_1 \cdots s_n$, as desired. ∎

9.58 לכל אופרטור אוניטרי יש דטרמיננטה עם ערך מוחלט 1

נניח ש־ V מרחב מכפלה פנימית ו־ $S \in \mathcal{L}(V)$ אופרטור אוניטרי. אזי $|\det S| = 1$.

הוכחה מאחר ש־ S אוניטרי, $I = S^*S$ (ראו 7.53). לפיכך

$$1 = \det(S^*S) = (\det S^*)(\det S) = \overline{(\det S)}(\det S) = |\det S|^2,$$

כאשר השוויון השני נובע מ־ 9.49(א) והשוויון השלישי נובע מ־ 9.56(ג). המשוואה לעיל גוררת ש־ $|\det S| = 1$.

הדטרמיננטה של אופרטור חיובי מעל מרחב מכפלה פנימית משתלבת היטב באנלוגיה שלפיה אופרטורים כאלה מקבילים למספרים הממשיים האי־שליליים.

9.59 לכל אופרטור חיובי יש דטרמיננטה אי־שלילית

נניח ש־ V מרחב מכפלה פנימית ו־ $T \in \mathcal{L}(V)$ אופרטור חיובי. אזי $\det T \geq 0$.

הוכחה לפי משפט הפירוק הספקטרלי (7.29 או 7.31), יש ל־ V בסיס אורתונורמלי המורכב מווקטורים עצמיים של T. לפיכך לפי הסעיף האחרון של 9.42, $\det T$ שווה למכפלת הערכים העצמיים של T, עם אפשרות לחזרות. כל ערך עצמי של T הוא מספר אי־שלילי (לפי 7.38). לפיכך אנו מסיקים ש־ $\det T \geq 0$.

נניח ש־ V מרחב מכפלה פנימית ו־ $T \in \mathcal{L}(V)$. ניזכר כי רשימת השורשים הריבועיים האי־שליליים של הערכים העצמיים של T^*T (כאשר מספר המופעים של כל ערך עצמי שווה לריבויו) נקראת רשימת הערכים הסינגולריים של T (ראו יחידה 3ז).

9.60 $|\det T| =$ מכפלת ערכים סינגולריים של T

נניח ש־ V מרחב מכפלה פנימית ו־ $T \in \mathcal{L}(V)$. אזי

$|\det T| = \sqrt{\det(T^*T)} =$ מכפלת ערכים סינגולריים של T.

הוכחה מתקיים

$$|\det T|^2 = \overline{(\det T)}(\det T) = \big(\det(T^*)\big)(\det T) = \det(T^*T),$$

כאשר השוויון האמצעי נובע מ־ 9.56(ג) והשוויון האחרון נובע מ־ 9.49(א). הוצאת שורשים לשני אגפי המשוואה לעיל מראה ש־ $|\det T| = \sqrt{\det(T^*T)}$.

נסמן ב־ $s_1, \dots, s_n$ את רשימת הערכים הסינגולריים של T. לפיכך $s_1^2, \dots, s_n^2$ היא רשימת הערכים העצמיים של T^*T (עם חזרות בהתאם), המשויכת לבסיס אורתונורמלי של V המורכב מווקטורים עצמיים של T^*T. לכן הסעיף האחרון של 9.42 גורר ש־ $\det(T^*T) = s_1^2 \cdots s_n^2$.

לפיכך $|\det T| = s_1 \cdots s_n$, כפי שרצינו.

An operator T on a real inner product space changes volume by a factor of the product of the singular values (by 7.111). Thus the next result follows immediately from 7.111 and 9.60. This result explains why the absolute value of a determinant appears in the change of variables formula in multivariable calculus.

9.61 *T changes volume by factor of $|\det T|$*

Suppose $T \in \mathcal{L}(\mathbf{R}^n)$ and $\Omega \subseteq \mathbf{R}^n$. Then

$$\text{volume } T(\Omega) = |\det T|(\text{volume } \Omega).$$

For operators on finite-dimensional complex vector spaces, we now connect the determinant to a polynomial that we have previously seen.

9.62 *if $\mathbf{F} = \mathbf{C}$, then characteristic polynomial of T equals $\det(zI - T)$*

Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$. Let $\lambda_1, \dots, \lambda_m$ denote the distinct eigenvalues of T, and let $d_1, \dots, d_m$ denote their multiplicities. Then

$$\det(zI - T) = (z - \lambda_1)^{d_1} \cdots (z - \lambda_m)^{d_m}.$$

Proof There exists a basis of V with respect to which T has an upper-triangular matrix with each λ_k appearing on the diagonal exactly d_k times (by 8.37). With respect to this basis, $zI - T$ has an upper-triangular matrix with $z - \lambda_k$ appearing on the diagonal exactly d_k times for each k. Thus 9.48 gives the desired equation. ∎

Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$. The characteristic polynomial of T was defined in 8.26 as the polynomial on the right side of the equation in 9.62. We did not previously define the characteristic polynomial of an operator on a finite-dimensional real vector space because such operators may have no eigenvalues, making a definition using the right side of the equation in 9.62 inappropriate.

We now present a new definition of the characteristic polynomial, motivated by 9.62. This new definition is valid for both real and complex vector spaces. The equation in 9.62 shows that this new definition is equivalent to our previous definition when $\mathbf{F} = \mathbf{C}$ (8.26).

9.63 definition: *characteristic polynomial*

Suppose $T \in \mathcal{L}(V)$. The polynomial defined by

$$z \mapsto \det(zI - T)$$

is called the *characteristic polynomial* of T.

The formula in 9.46 shows that the characteristic polynomial of an operator $T \in \mathcal{L}(V)$ is a monic polynomial of degree $\dim V$. The zeros in $\mathbf{F}$ of the characteristic polynomial of T are exactly the eigenvalues of T (by 9.51).

אופרטור T מעל מרחב מכפלה פנימית ממשי משנה נפח בפקטור של מכפלת הערכים הסינגולריים (לפי 7.111). לפיכך המסקנה הבאה נובעת ישירות מ־ 7.111 ו־ 9.60. מסקנה זו מסבירה מדוע הערך המוחלט של דטרמיננטה מופיע במשפט החלפת המשתנים הכללי בחדו"א.

9.61 T משנה נפח בפקטור של $|\det T|$

נניח ש־ $T \in \mathcal{L}(\mathbf{R}^n)$ ו־ $\Omega \subseteq \mathbf{R}^n$. אזי

$$\text{volume } T(\Omega) = |\det T|(\text{volume } \Omega).$$

עבור אופרטורים מעל מרחבים וקטוריים מרוכבים מממד סופי, אנו נחבר כעת את הדטרמיננטה לפולינום שאותו כבר פגשנו.

9.62 אם $\mathbf{F} = \mathbf{C}$, אזי הפולינום האופייני של T שווה ל־ $\det(zI - T)$

נניח ש־ $\mathbf{F} = \mathbf{C}$ ו־ $T \in \mathcal{L}(V)$. נסמן ב־ $\lambda_1, \dots, \lambda_m$ את הערכים העצמיים השונים של T, וב־ $d_1, \dots, d_m$ את הריבויים שלהם. אזי

$$\det(zI - T) = (z - \lambda_1)^{d_1} \cdots (z - \lambda_m)^{d_m}.$$

הוכחה קיים בסיס של V אשר ביחס אליו יש ל־ T מטריצה משולשית עליונה כאשר כל λ_k מופיע על האלכסון בדיוק d_k פעמים (לפי 8.37). ביחס לבסיס זה, יש ל־ $zI - T$ מטריצה משולשית עליונה כאשר $z - \lambda_k$ מופיע על האלכסון בדיוק d_k פעמים לכל k. לפיכך מ־ 9.48 נקבל את המשוואה הרצויה. ∎

נניח ש־ $\mathbf{F} = \mathbf{C}$ ו־ $T \in \mathcal{L}(V)$. הפולינום האופייני של T הוגדר ב־ 8.26 כפולינום באגף הימני של המשוואה ב־ 9.62. לא הגדרנו עד כה פולינום אופייני של אופרטור מעל מרחב וקטורי ממשי מממד סופי מאחר שלאופרטורים אלו ייתכן שלא יהיו ערכים עצמיים, מה שהופך הגדרה תוך שימוש באגף הימני של המשוואה ב־ 9.62 לבלתי מתאימה.

כעת נציג הגדרה חדשה של הפולינום האופייני, כשהמוטיבציה שלה נובעת מ־ 9.62. הגדרה חדשה זו תקפה במרחבים וקטוריים ממשיים ומרוכבים כאחד. המשוואה ב־ 9.62 מראה לנו כי הגדרה חדשה זו שקולה להגדרה הקודמת כאשר 8.26 ($\mathbf{F} = \mathbf{C}$).

9.63 הגדרה: פולינום אופייני

נניח ש־ $T \in \mathcal{L}(V)$. הפולינום המוגדר על ידי

$$z \mapsto \det(zI - T)$$

נקרא הפולינום האופייני של T.

הנוסחה ב־ 9.46 מראה שהפולינום האופייני של אופרטור $T \in \mathcal{L}(V)$ הוא פולינום מתוקן ממעלה $\dim V$. השורשים ב־ $\mathbf{F}$ של הפולינום האופייני של T הם בדיוק הערכים העצמיים של T (לפי 9.51).

Previously we proved the Cayley–Hamilton theorem (8.29) in the complex case. Now we can extend that result to operators on real vector spaces.

9.64 *Cayley–Hamilton theorem*

Suppose $T \in \mathcal{L}(V)$ and q is the characteristic polynomial of T. Then $q(T) = 0$.

Proof If $\mathbf{F} = \mathbf{C}$, then the equation $q(T) = 0$ follows from 9.62 and 8.29.

Now suppose $\mathbf{F} = \mathbf{R}$. Fix a basis of V, and let A be the matrix of T with respect to this basis. Let S be the operator on $\mathbf{C}^{\dim V}$ such that the matrix of S (with respect to the standard basis of $\mathbf{C}^{\dim V}$) is A. For all $z \in \mathbf{R}$ we have

$$q(z) = \det(zI - T) = \det(zI - A) = \det(zI - S).$$

Thus q is the characteristic polynomial of S. The case $\mathbf{F} = \mathbf{C}$ (first sentence of this proof) now implies that $0 = q(S) = q(A) = q(T)$. ∎

The Cayley–Hamilton theorem (9.64) implies that the characteristic polynomial of an operator $T \in \mathcal{L}(V)$ is a polynomial multiple of the minimal polynomial of T (by 5.29). Thus if the degree of the minimal polynomial of T equals $\dim V$, then the characteristic polynomial of T equals the minimal polynomial of T. This happens for a very large percentage of operators, including over 99.999% of 4-by-4 matrices with integer entries in $[-100, 100]$ (see the paragraph following 5.25).

The last sentence in our next result was previously proved in the complex case (see 8.54). Now we can give a proof that works on both real and complex vector spaces.

9.65 *characteristic polynomial, trace, and determinant*

Suppose $T \in \mathcal{L}(V)$. Let $n = \dim V$. Then the characteristic polynomial of T can be written as

$$z^n - (\operatorname{tr} T)z^{n-1} + \cdots + (-1)^n(\det T).$$

Proof The constant term of a polynomial function of z is the value of the polynomial when $z = 0$. Thus the constant term of the characteristic polynomial of T equals $\det(-T)$, which equals $(-1)^n \det T$ (by the third bullet point of 9.42).

Fix a basis of V, and let A be the matrix of T with respect to this basis. The matrix of $zI - T$ with respect to this basis is $zI - A$. The term coming from the identity permutation $(1, \dots, n)$ in the formula 9.46 for $\det(zI - A)$ is

$$(z - A_{1,1}) \cdots (z - A_{n,n}).$$

The coefficient of z^{n-1} in the expression above is $-(A_{1,1} + \cdots + A_{n,n})$, which equals $-\operatorname{tr} T$. The terms in the formula for $\det(zI - A)$ coming from other elements of perm n contain at most $n-2$ factors of the form $z - A_{k,k}$ and thus do not contribute to the coefficient of z^{n-1} in the characteristic polynomial of T. ∎

מוקדם יותר הוכחנו את משפט קיילי־המילטון (8.29) במקרה המרוכב. כעת אנו יכולים להרחיב מסקנה זו לאופרטורים מעל מרחבים וקטוריים ממשיים.

9.64 משפט קיילי־המילטון

נניח ש־ $T \in \mathcal{L}(V)$ ו־ q הוא הפולינום האופייני של T. אזי $q(T) = 0$.

הוכחה אם $\mathbf{F} = \mathbf{C}$, אזי המשוואה $q(T) = 0$ נובעת מ־ 9.62 ו־ 8.29.

כעת נניח ש־ $\mathbf{F} = \mathbf{R}$. נְקַבֵּעַ בסיס של V, ותהי A המטריצה של T ביחס לבסיס זה. יהי S האופרטור מעל $\mathbf{C}^{\dim V}$ כך שהמטריצה של S (ביחס לבסיס הסטנדרטי של $\mathbf{C}^{\dim V}$) היא A. לכל $z \in \mathbf{R}$ מתקיים

$$q(z) = \det(zI - T) = \det(zI - A) = \det(zI - S).$$

לפיכך q הוא הפולינום האופייני של S. המקרה של $\mathbf{F} = \mathbf{C}$ (המשפט הראשון בהוכחה זו) כעת גורר ש־ $0 = q(S) = q(A) = q(T)$. ■

משפט קיילי־המילטון (9.64) גורר שהפולינום האופייני של אופרטור $T \in \mathcal{L}(V)$ הוא כפולת פולינום של הפולינום המינימלי של T (לפי 5.29). לפיכך אם מעלת הפולינום המינימלי של T שווה ל־ $\dim V$, אזי הפולינום האופייני של T שווה לפולינום המינימלי של T. הדבר קורה לאחוז גבוה מאוד של אופרטורים, כולל מעל ל־ 99.999% של מטריצות 4×4 שאיבריהן מספרים שלמים בקטע $[-100, 100]$ (ראו הפסקה שאחרי 5.25).

המשפט האחרון במסקנה הבאה הוכח מקודם עבור המקרה המרוכב (ראו 8.54). כעת נספק הוכחה העובדת במרחבים וקטוריים ממשיים ומרוכבים כאחד.

9.65 פולינום אופייני, עקבה ודטרמיננטה

נניח ש־ $T \in \mathcal{L}(V)$. יהי $n = \dim V$. אזי הפולינום האופייני של T ניתן לכתיבה כ־

$$z^n - (\operatorname{tr} T)z^{n-1} + \cdots + (-1)^n(\det T).$$

הוכחה האיבר הקבוע של פונקציית פולינום של z הוא ערך הפולינום כאשר $z = 0$. לפיכך האיבר הקבוע של הפולינום האופייני של T שווה ל־ $\det(-T)$, השווה ל־ $(-1)^n \det T$ (לפי הסעיף השלישי ב־ 9.42).

נְקַבֵּעַ בסיס של V, ותהי A המטריצה של T ביחס לבסיס זה. המטריצה של $zI - T$ ביחס לבסיס זה היא $zI - A$. האיבר הנובע מתמורת הזהות $(1, \ldots, n)$ בנוסחה 9.46 עבור $\det(zI - A)$ הוא

$$(z - A_{1,1}) \cdots (z - A_{n,n}).$$

המקדם של z^{n-1} בביטוי לעיל הוא $-(A_{1,1} + \cdots + A_{n,n})$, השווה ל־ $-\operatorname{tr} T$. האיברים בנוסחה עבור $\det(zI - A)$ הנובעים מאיברים אחרים של $\operatorname{perm} n$ מכילים לכל היותר $n - 2$ גורמים מהצורה $z - A_{k,k}$ ולפיכך אינם תורמים למקדם של z^{n-1} בפולינום האופייני של T. ■

The next result was proved by Jacques Hadamard (1865–1963) in 1893.

In the result below, think of the columns of the n-by-n matrix A as elements of $\mathbf{F}^n$. The norms appearing below then arise from the standard inner product on $\mathbf{F}^n$. Recall that the notation $R_{\cdot,k}$ in the proof below means the kth column of the matrix R (as was defined in 3.44).

9.66 *Hadamard's inequality*

Suppose A is an n-by-n matrix. Let $v_1, \dots, v_n$ denote the columns of A. Then

$$|\det A| \le \prod_{k=1}^{n} \|v_k\|.$$

Proof If A is not invertible, then $\det A = 0$ and hence the desired inequality holds in this case.

Thus assume that A is invertible. The QR factorization (7.58) tells us that there exist a unitary matrix Q and an upper-triangular matrix R whose diagonal contains only positive numbers such that $A = QR$. We have

$$\begin{aligned}
|\det A| &= |\det Q| \, |\det R| \\
&= |\det R| \\
&= \prod_{k=1}^{n} R_{k,k} \\
&\le \prod_{k=1}^{n} \|R_{\cdot,k}\| \\
&= \prod_{k=1}^{n} \|QR_{\cdot,k}\| \\
&= \prod_{k=1}^{n} \|v_k\|,
\end{aligned}$$

where the first line comes from 9.49(b), the second line comes from 9.58, the third line comes from 9.48, and the fifth line holds because Q is an isometry. ∎

To give a geometric interpretation to Hadamard's inequality, suppose $\mathbf{F} = \mathbf{R}$. Let $T \in \mathcal{L}(\mathbf{R}^n)$ be the operator such that $Te_k = v_k$ for each $k = 1, \dots, n$, where $e_1, \dots, e_n$ is the standard basis of $\mathbf{R}^n$. Then T maps the box $P(e_1, \dots, e_n)$ onto the parallelepiped $P(v_1, \dots, v_n)$ [see 7.102 and 7.105 for a review of this notation and terminology]. Because the box $P(e_1, \dots, e_n)$ has volume 1, this implies (by 9.61) that the parallelepiped $P(v_1, \dots, v_n)$ has volume $|\det T|$, which equals $|\det A|$. Thus Hadamard's inequality above can be interpreted to say that among all parallelepipeds whose edges have lengths $\|v_1\|, \dots, \|v_n\|$, the ones with largest volume have orthogonal edges $\left(\text{and thus have volume } \prod_{k=1}^{n} \|v_k\|\right)$.

For a necessary and sufficient condition for Hadamard's inequality to be an equality, see Exercise 18.

המסקנה הבאה הוכחה על ידי ז'אק סלומון אָדָמָר (1865–1963) בשנת 1893.

במסקנה להלן, חשבו על עמודות A מטריצה $n \times n$ כאיברים של $\mathbf{F}^n$. הנורמות המופיעות להלן הן תוצאת המכפלה הפנימית הסטנדרטית מעל $\mathbf{F}^n$. ניזכר כי הסימון $R_{\cdot,k}$ בהוכחה להלן משמעותו היא העמודה ה־ k^{th} של המטריצה R (כפי שהוגדר ב־ 3.44).

9.66 אי־שוויון אָדָמָר

נניח ש־ A מטריצה $n \times n$. נסמן ב־ $v_1, \dots, v_n$ את עמודות A. אזי

$$|\det A| \le \prod_{k=1}^{n} \|v_k\|.$$

הוכחה אם A אינה הפיכה, אזי $\det A = 0$ ולכן האי־שוויון הרצוי מתקיים במקרה זה.

לפיכך נניח ש־ A הפיכה. הפירוק QR (ראו 7.58) אומר לנו כי קיימת מטריצה אוניטרית Q ומטריצה משולשית עליונה R שהאלכסון שלה מכיל רק איברים חיוביים כך ש־ $A = QR$. מתקיים

$$\begin{aligned}
|\det A| &= |\det Q| \, |\det R| \\
&= |\det R| \\
&= \prod_{k=1}^{n} R_{k,k} \\
&\le \prod_{k=1}^{n} \|R_{\cdot,k}\| \\
&= \prod_{k=1}^{n} \|QR_{\cdot,k}\| \\
&= \prod_{k=1}^{n} \|v_k\|,
\end{aligned}$$

כאשר השורה הראשונה נובעת מ־ 9.49(ב), השורה השנייה נובעת מ־ 9.58, השורה השלישית נובעת מ־ 9.48 והשורה החמישית מתקיימת מאחר ש־ Q איזומטריה. ■

להענקת פרשנות גאומטרית לאי־שוויון אדמר, נניח ש־ $\mathbf{F} = \mathbf{R}$. יהי $T \in \mathcal{L}(\mathbf{R}^n)$ האופרטור כך ש־ $Te_k = v_k$ לכל $k = 1, \dots, n$, כאשר $e_1, \dots, e_n$ הוא הבסיס הסטנדרטי של $\mathbf{R}^n$. אזי T מעתיק את התיבה $P(e_1, \dots, e_n)$ על המקבילון $P(v_1, \dots, v_n)$ [ראו 7.102 ו־ 7.105 לסקירה של סימון וטרמינולוגיה אלו]. מאחר שנפח התיבה $P(e_1, \dots, e_n)$ הוא 1, מכאן נובע (לפי 9.61) שנפח המקבילון $P(v_1, \dots, v_n)$ הוא $|\det T|$, השווה ל־ $|\det A|$. לפיכך נוכל להעניק פרשנות לאי־שוויון אדמר לעיל ככזה הקובע כי מבין כל המקבילונים שאורכי צלעותיהם $\|v_1\|, \dots, \|v_n\|$, בעלי הנפח הגדול ביותר הם אלו שצלעותיהם ניצבות זו לזו (ולפיכך נפחם הוא $\prod_{k=1}^{n} \|v_k\|$).

עבור תנאי הכרחי ומספיק שאי־שוויון אָדָמָר יהיה שוויון, ראו תרגיל 18.

The matrix in the next result is called the *Vandermonde matrix*. Vandermonde matrices have important applications in polynomial interpolation, the discrete Fourier transform, and other areas of mathematics. The proof of the next result is a nice illustration of the power of switching between matrices and linear maps.

9.67 *determinant of Vandermonde matrix*

Suppose $n > 1$ and $\beta_1, \dots, \beta_n \in \mathbf{F}$. Then

$$\det \begin{pmatrix} 1 & \beta_1 & \beta_1^2 & \cdots & \beta_1^{n-1} \\ 1 & \beta_2 & \beta_2^2 & \cdots & \beta_2^{n-1} \\ & & \ddots & & \\ 1 & \beta_n & \beta_n^2 & \cdots & \beta_n^{n-1} \end{pmatrix} = \prod_{1 \le j < k \le n} (\beta_k - \beta_j).$$

Proof Let $1, z, \dots, z^{n-1}$ be the standard basis of $\mathcal{P}_{n-1}(\mathbf{F})$ and let $e_1, \dots, e_n$ denote the standard basis of $\mathbf{F}^n$. Define a linear map $S \colon \mathcal{P}_{n-1}(\mathbf{F}) \to \mathbf{F}^n$ by

$$Sp = (p(\beta_1), \dots, p(\beta_n)).$$

Let A denote the Vandermonde matrix shown in the statement of this result. Note that

$$A = \mathcal{M}(S, (1, z, \dots, z^{n-1}), (e_1, \dots, e_n)).$$

Let $T \colon \mathcal{P}_{n-1}(\mathbf{F}) \to \mathcal{P}_{n-1}(\mathbf{F})$ be the operator on $\mathcal{P}_{n-1}(\mathbf{F})$ such that $T1 = 1$ and

$$Tz^k = (z - \beta_1)(z - \beta_2) \cdots (z - \beta_k)$$

for $k = 1, \dots, n-1$. Let $B = \mathcal{M}(T, (1, z, \dots, z^{n-1}), (1, z, \dots, z^{n-1}))$. Then B is an upper-triangular matrix all of whose diagonal entries equal 1. Thus $\det B = 1$ (by 9.48).

Let $C = \mathcal{M}(ST, (1, z, \dots, z^{n-1}), (e_1, \dots, e_n))$. Thus $C = AB$ (by 3.81), which implies that

$$\det A = (\det A)(\det B) = \det C.$$

The definitions of C, S, and T show that C equals

$$\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & \beta_2 - \beta_1 & 0 & \cdots & 0 \\ 1 & \beta_3 - \beta_1 & (\beta_3 - \beta_1)(\beta_3 - \beta_2) & \cdots & 0 \\ & & & \ddots & \\ 1 & \beta_n - \beta_1 & (\beta_n - \beta_1)(\beta_n - \beta_2) & \cdots & (\beta_n - \beta_1)(\beta_n - \beta_2) \cdots (\beta_n - \beta_{n-1}) \end{pmatrix}.$$

Now $\det A = \det C = \prod_{1 \le j < k \le n} (\beta_k - \beta_j)$, where we have used 9.56(a) and 9.48. ∎

המטריצה במסקנה הבאה נקראת מטריצת וַנְדֶרְמוֹנְד. למטריצות וַנְדֶרְמוֹנְד יש יישומים חשובים באינטרפולציה פולינומית, התמרת פורייה הבדידה ותחומים נוספים במתמטיקה. הוכחת המסקנה הבאה היא המחשה יפה של העוצמה שבמעבר הלוך ושוב בין מטריצות להעתקות ליניאריות.

9.67 דטרמיננטה של מטריצת וַנְדֶרְמוֹנְד

נניח ש־ $n > 1$ ו־ $\beta_1, \dots, \beta_n \in \mathbf{F}$. אזי

$$\det \begin{pmatrix} 1 & \beta_1 & \beta_1^2 & \cdots & \beta_1^{n-1} \\ 1 & \beta_2 & \beta_2^2 & \cdots & \beta_2^{n-1} \\ & & \ddots & & \\ 1 & \beta_n & \beta_n^2 & \cdots & \beta_n^{n-1} \end{pmatrix} = \prod_{1 \le j < k \le n} (\beta_k - \beta_j).$$

הוכחה תהי $1, z, \dots, z^{n-1}$ הבסיס הסטנדרטי של $\mathcal{P}_{n-1}(\mathbf{F})$ ונסמן ב־ $e_1, \dots, e_n$ את הבסיס הסטנדרטי של $\mathbf{F}^n$. נגדיר העתקה ליניארית $S : \mathcal{P}_{n-1}(\mathbf{F}) \to \mathbf{F}^n$ על ידי

$$Sp = (p(\beta_1), \dots, p(\beta_n)).$$

נסמן ב־ A את מטריצת ונדרמונד המוצגת בניסוח מסקנה זו. נשים לב כי

$$A = \mathcal{M}(S, (1, z, \dots, z^{n-1}), (e_1, \dots, e_n)).$$

יהי $T : \mathcal{P}_{n-1}(\mathbf{F}) \to \mathcal{P}_{n-1}(\mathbf{F})$ האופרטור מעל $\mathcal{P}_{n-1}(\mathbf{F})$ כך ש־ $T1 = 1$ ו־

$$Tz^k = (z - \beta_1)(z - \beta_2) \cdots (z - \beta_k)$$

עבור $k = 1, \dots, n-1$. יהי $B = \mathcal{M}(T, (1, z, \dots, z^{n-1}), (1, z, \dots, z^{n-1}))$. אזי B היא מטריצה משולשית עליונה שכל איברי האלכסון שלה שווים 1. לפיכך $\det B = 1$ (לפי 9.48).

תהי $C = \mathcal{M}(ST, (1, z, \dots, z^{n-1}), (e_1, \dots, e_n))$. לפיכך $C = AB$ (לפי 3.81), מה שגורר ש־

$$\det A = (\det A)(\det B) = \det C.$$

הגדרותיהן של C, S ו־ T מראות כי C שווה ל־

$$\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & \beta_2 - \beta_1 & 0 & \cdots & 0 \\ 1 & \beta_3 - \beta_1 & (\beta_3 - \beta_1)(\beta_3 - \beta_2) & \cdots & 0 \\ & & & \ddots & \\ 1 & \beta_n - \beta_1 & (\beta_n - \beta_1)(\beta_n - \beta_2) & \cdots & (\beta_n - \beta_1)(\beta_n - \beta_2) \cdots (\beta_n - \beta_{n-1}) \end{pmatrix}.$$

כעת $\det A = \det C = \prod_{1 \le j < k \le n} (\beta_k - \beta_j)$, כאשר השתמשנו ב־ 9.56(א) ו־ 9.48. ■

Exercises 9C

1 Prove or give a counterexample: $S, T \in \mathcal{L}(V) \implies \det(S+T) = \det S + \det T$.

2 Suppose the first column of a square matrix A consists of all zeros except possibly the first entry $A_{1,1}$. Let B be the matrix obtained from A by deleting the first row and the first column of A. Show that $\det A = A_{1,1} \det B$.

3 Suppose $T \in \mathcal{L}(V)$ is nilpotent. Prove that $\det(I + T) = 1$.

4 Suppose V is an inner product space and $S \in \mathcal{L}(V)$. Prove that S is unitary if and only if $|\det S| = \|S\| = 1$.

5 Suppose A is a block upper-triangular matrix

$$A = \begin{pmatrix} A_1 & & * \\ & \ddots & \\ 0 & & A_m \end{pmatrix},$$

where each A_k along the diagonal is a square matrix. Prove that

$$\det A = (\det A_1) \cdots (\det A_m).$$

6 Suppose $A = \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}$ is an n-by-n matrix, with v_k denoting the kth column of A. Show that if $(m_1, \dots, m_n) \in \operatorname{perm} n$, then

$$\det \begin{pmatrix} v_{m_1} & \cdots & v_{m_n} \end{pmatrix} = (\operatorname{sign}(m_1, \dots, m_n)) \det A.$$

7 Suppose $T \in \mathcal{L}(V)$ is invertible. Let p denote the characteristic polynomial of T and let q denote the characteristic polynomial of T^{-1}. Prove that

$$q(z) = \frac{1}{p(0)} z^{\dim V} p\left(\frac{1}{z}\right)$$

for all nonzero $z \in \mathbf{F}$.

8 Suppose $T \in \mathcal{L}(V)$ is an operator with no eigenvalues (which implies that $\mathbf{F} = \mathbf{R}$). Prove that $\det T > 0$.

9 Suppose that V is a real vector space of even dimension, $T \in \mathcal{L}(V)$, and $\det T < 0$. Prove that T has at least two distinct eigenvalues.

10 Suppose V is a real vector space of odd dimension and $T \in \mathcal{L}(V)$. Without using the minimal polynomial, prove that T has an eigenvalue.

This result was previously proved without using determinants or the characteristic polynomial—see 5.34.

11 Prove or give a counterexample: If $\mathbf{F} = \mathbf{R}$, $T \in \mathcal{L}(V)$, and $\det T > 0$, then T has a square root.

If $\mathbf{F} = \mathbf{C}$, $T \in \mathcal{L}(V)$, and $\det T \neq 0$, then T has a square root (see 8.41).

תרגילים ט3

1 הוכיחו או הביאו דוגמה נגדית: $\det(S+T) = \det S + \det T \implies S, T \in \mathcal{L}(V)$.

2 נניח שהעמודה הראשונה של מטריצה ריבועית A מורכבת כולה מאפסים מלבד האפשרות שהאיבר הראשון $A_{1,1}$ אינו אפס. תהי B המטריצה המושגת מ־ A על ידי מחיקת השורה הראשונה והעמודה הראשונה של A. הראו כי $\det A = A_{1,1} \det B$.

3 נניח ש־ $T \in \mathcal{L}(V)$ נילפוטנטי. הוכיחו כי $\det(I+T) = 1$.

4 נניח ש־ V מרחב מכפלה פנימית ו־ $S \in \mathcal{L}(V)$. הוכיחו כי S אוניטרי אם־ורק־אם $|\det S| = \|S\| = 1$.

5 נניח ש־ A מטריצת בלוקים משולשית עליונה

$$A = \begin{pmatrix} A_1 & & * \\ & \ddots & \\ 0 & & A_m \end{pmatrix},$$

כאשר כל A_k לאורך האלכסון היא מטריצה ריבועית. הוכיחו כי

$$\det A = (\det A_1) \cdots (\det A_m).$$

6 נניח ש־ $A = \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}$ היא מטריצה $n \times n$, כאשר v_k מסמן את העמודה ה־ k^{th} של A. הראו כי אם $(m_1, \dots, m_n) \in \operatorname{perm} n$, אזי

$$\det \begin{pmatrix} v_{m_1} & \cdots & v_{m_n} \end{pmatrix} = (\operatorname{sign}(m_1, \dots, m_n)) \det A.$$

7 נניח ש־ $T \in \mathcal{L}(V)$ הפיך. נסמן ב־ p את הפולינום האופייני של T וב־ q את הפולינום האופייני של T^{-1}. הוכיחו כי

$$q(z) = \frac{1}{p(0)} z^{\dim V} p\left(\frac{1}{z}\right)$$

לכל $z \in \mathbf{F}$ השונה מאפס.

8 נניח ש־ $T \in \mathcal{L}(V)$ אופרטור ללא ערכים עצמיים (מה שגורר ש־ $\mathbf{F} = \mathbf{R}$). הוכיחו כי $\det T > 0$.

9 נניח ש־ V מרחב וקטורי ממשי מממד זוגי, $T \in \mathcal{L}(V)$ ו־ $\det T < 0$. הוכיחו כי יש ל־ T לכל הפחות שני ערכים עצמיים שונים.

10 נניח ש־ V מרחב וקטורי ממשי מממד אי־זוגי ו־ $T \in \mathcal{L}(V)$. מבלי שתשתמשו בפולינום המינימלי, הוכיחו כי יש ל־ T ערך עצמי.

תוצאה זו הוכחה מוקדם יותר ללא שימוש בדטרמיננטה או בפולינום האופייני—ראו 5.34.

11 הוכיחו או הביאו דוגמה נגדית: אם $T \in \mathcal{L}(V)$, $\mathbf{F} = \mathbf{R}$ ו־ $\det T > 0$, אזי יש ל־ T שורש ריבועי.

אם $T \in \mathcal{L}(V)$, $\mathbf{F} = \mathbf{C}$ ו־ $\det T \neq 0$, אזי יש ל־ T שורש ריבועי (ראו 8.41).

12 Suppose $S, T \in \mathcal{L}(V)$ and S is invertible. Define $p\colon \mathbf{F} \to \mathbf{F}$ by

$$p(z) = \det(zS - T).$$

Prove that p is a polynomial of degree $\dim V$ and that the coefficient of $z^{\dim V}$ in this polynomial is $\det S$.

13 Suppose $\mathbf{F} = \mathbf{C}$, $T \in \mathcal{L}(V)$, and $n = \dim V > 2$. Let $\lambda_1, \ldots, \lambda_n$ denote the eigenvalues of T, with each eigenvalue included as many times as its multiplicity.

(a) Find a formula for the coefficient of z^{n-2} in the characteristic polynomial of T in terms of $\lambda_1, \ldots, \lambda_n$.

(b) Find a formula for the coefficient of z in the characteristic polynomial of T in terms of $\lambda_1, \ldots, \lambda_n$.

14 Suppose V is an inner product space and T is a positive operator on V. Prove that

$$\det \sqrt{T} = \sqrt{\det T}.$$

15 Suppose V is an inner product space and $T \in \mathcal{L}(V)$. Use the polar decomposition to give a proof that

$$|\det T| = \sqrt{\det(T^*T)}$$

that is different from the proof given earlier (see 9.60).

16 Suppose $T \in \mathcal{L}(V)$. Define $g\colon \mathbf{F} \to \mathbf{F}$ by $g(x) = \det(I + xT)$. Show that $g'(0) = \operatorname{tr} T$.

Look for a clean solution to this exercise, without using the explicit but complicated formula for the determinant of a matrix.

17 Suppose a, b, c are positive numbers. Find the volume of the ellipsoid

$$\left\{(x, y, z) \in \mathbf{R}^3 : \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} < 1\right\}$$

by finding a set $\Omega \subseteq \mathbf{R}^3$ whose volume you know and an operator T on $\mathbf{R}^3$ such that $T(\Omega)$ equals the ellipsoid above.

18 Suppose that A is an invertible square matrix. Prove that Hadamard's inequality (9.66) is an equality if and only if each column of A is orthogonal to the other columns.

19 Suppose V is an inner product space, $e_1, \ldots, e_n$ is an orthonormal basis of V, and $T \in \mathcal{L}(V)$ is a positive operator.

(a) Prove that $\det T \le \prod_{k=1}^{n} \langle Te_k, e_k \rangle$.

(b) Prove that if T is invertible, then the inequality in (a) is an equality if and only if e_k is an eigenvector of T for each $k = 1, \ldots, n$.

12 נניח ש־ $S, T \in \mathcal{L}(V)$ ו־ S הפיך. נגדיר $p : \mathbf{F} \to \mathbf{F}$ על ידי

$$p(z) = \det(zS - T).$$

הוכיחו כי p פולינום ממעלה $\dim V$ וכי המקדם של $z^{\dim V}$ בפולינום זה הוא $\det S$.

13 נניח ש־ $\mathbf{F} = \mathbf{C}$, $T \in \mathcal{L}(V)$ ו־ $n = \dim V > 2$. נסמן ב־ $\lambda_1, \dots, \lambda_n$ את הערכים העצמיים של T, כאשר כל ערך עצמי נכלל מספר פעמים השווה לריבויו.

א. מצאו נוסחה עבור המקדם של z^{n-2} בפולינום האופייני של T במונחים של $\lambda_1, \dots, \lambda_n$.

ב. מצאו נוסחה עבור המקדם של z בפולינום האופייני של T במונחים של $\lambda_1, \dots, \lambda_n$.

14 נניח ש־ V מרחב מכפלה פנימית ו־ T אופרטור חיובי מעל V. הוכיחו כי

$$\det \sqrt{T} = \sqrt{\det T}.$$

15 נניח ש־ V מרחב מכפלה פנימית ו־ $T \in \mathcal{L}(V)$. השתמשו בפירוק הפולרי להוכיח כי

$$|\det T| = \sqrt{\det(T^*T)}$$

באופן שונה מההוכחה שניתנה לעיל (ראו 9.60).

16 נניח ש־ $T \in \mathcal{L}(V)$. נגדיר $g : \mathbf{F} \to \mathbf{F}$ על ידי $g(x) = \det(I + xT)$. הראו כי $g'(0) = \operatorname{tr} T$.

חפשו פתרון נקי לתרגיל זה, מבלי שתשתמשו בנוסחה המפורשת אך המסובכת של הדטרמיננטה של מטריצה.

17 נניח ש־ a, b, c מספרים חיוביים. מצאו את נפח האליפסואיד

$$\left\{(x, y, z) \in \mathbf{R}^3 : \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} < 1\right\}$$

על ידי מציאת קבוצה $\Omega \subseteq \mathbf{R}^3$ שנפחה ידוע ואופרטור T מעל $\mathbf{R}^3$ כך ש־ $T(\Omega)$ שווה לאליפסואיד לעיל.

18 נניח ש־ A מטריצה ריבועית הפיכה. הוכיחו כי אי־שוויון אדמר (9.66) הוא שוויון אם־ורק־אם כל עמודה של A אורתוגונלית לעמודות האחרות.

19 נניח ש־ V מרחב מכפלה פנימית, $e_1, \dots, e_n$ בסיס אורתונורמלי של V ו־ $T \in \mathcal{L}(V)$ אופרטור חיובי.

א. הוכיחו כי $\det T \le \prod_{k=1}^{n} \langle Te_k, e_k \rangle$.

ב. הוכיחו כי אם T הפיך, אזי האי־שוויון ב־ (א) הוא שוויון אם־ורק־אם e_k הוא וקטור עצמי של T לכל $k = 1, \dots, n$.

20 Suppose A is an n-by-n matrix, and suppose c is such that $|A_{j,k}| \le c$ for all $j, k \in \{1, \dots, n\}$. Prove that

$$|\det A| \le c^n n^{n/2}.$$

The formula for the determinant of a matrix (9.46) shows that $|\det A| \le c^n n!$. *However, the estimate given by this exercise is much better. For example, if* $c = 1$ *and* $n = 100$, *then* $c^n n! \approx 10^{158}$, *but the estimate given by this exercise is the much smaller number* 10^{100}. *If* n *is an integer power of* 2, *then the inequality above is sharp and cannot be improved.*

21 Suppose n is a positive integer and $\delta : \mathbf{C}^{n,n} \to \mathbf{C}$ is a function such that

$$\delta(AB) = \delta(A) \cdot \delta(B)$$

for all $A, B \in \mathbf{C}^{n,n}$ and $\delta(A)$ equals the product of the diagonal entries of A for each diagonal matrix $A \in \mathbf{C}^{n,n}$. Prove that

$$\delta(A) = \det A$$

for all $A \in \mathbf{C}^{n,n}$.

Recall that $\mathbf{C}^{n,n}$ *denotes the set of* n*-by-*n *matrices with entries in* $\mathbf{C}$. *This exercise shows that the determinant is the unique function defined on square matrices that is multiplicative and has the desired behavior on diagonal matrices. This result is analogous to Exercise 10 in Section 8D, which shows that the trace is uniquely determined by its algebraic properties.*

I find that in my own elementary lectures, I have, for pedagogical reasons, pushed determinants more and more into the background. Too often I have had the experience that, while the students acquired facility with the formulas, which are so useful in abbreviating long expressions, they often failed to gain familiarity with their *meaning*, and skill in manipulation prevented the student from going into all the details of the subject and so gaining a mastery.

—*Elementary Mathematics from an Advanced Standpoint*: *Geometry*, Felix Klein

20 נניח ש־ A מטריצה $n \times n$, ונניח ש־ c כך ש־ $|A_{j,k}| \le c$ לכל $j,k \in \{1,\dots,n\}$. הוכיחו כי

$$|\det A| \le c^n n^{n/2}.$$

הנוסחה עבור הדטרמיננטה של מטריצה (9.46) מראה כי $|\det A| \le c^n n!$. בכל זאת, ההערכה שניתנת בתרגיל זה טובה לאין שיעור. לדוגמה, אם $c = 1$ ו־ $n = 100$, אזי $c^n n! \approx 10^{158}$, אך ההערכה שניתנת בתרגיל זה היא המספר הקטן בהרבה 10^{100}. אם n חזקה שלמה של 2, אזי האי־שוויון לעיל הדוק ואין בו מקום לשיפור.

21 נניח ש־ n מספר שלם חיובי ו־ $\delta : \mathbf{C}^{n,n} \to \mathbf{C}$ פונקציה כך ש־

$$\delta(AB) = \delta(A) \cdot \delta(B)$$

לכל $A, B \in \mathbf{C}^{n,n}$ ו־ $\delta(A)$ שווה למכפלת איברי האלכסון של A לכל מטריצה אלכסונית $A \in \mathbf{C}^{n,n}$. הוכיחו כי

$$\delta(A) = \det A$$

לכל $A \in \mathbf{C}^{n,n}$.

נזכיר כי $\mathbf{C}^{n,n}$ מסמן את קבוצת המטריצות $n \times n$ עם איברים ב־ $\mathbf{C}$. תרגיל זה מראה כי הדטרמיננטה היא הפונקציה היחידה המוגדרת מעל מטריצה ריבועית שהיא כפלית ובעלת התנהגות רצויה מעל מטריצות אלכסוניות. מסקנה זו אנלוגית לתרגיל 10 ביחידה ח4, מה שמראה שהעקבה נקבעת באופן יחיד על ידי תכונותיה האלגבריות.

אני נוכח לראות כי בהרצאות היסודיות שלי, אני נוהג, מטעמים פדגוגיים, להדחיק את הדטרמיננטות אל השוליים האחוריים. לעיתים קרובות מדי חוויתי כי, בעוד הסטודנטים רוכשים מיומנות עם הנוסחאות, שהן כה שימושיות לקיצור ביטויים ארוכים, לעיתים קרובות הם כושלים בהיכרות עם משמעויותיהן, וכי המיומנות בביצוע מניפולציות אלגבריות מנעה מהסטודנטים לרדת לכל פרטי הנושא וכך מנעה מהם את השגת השליטה בו.

—מתמטיקה יסודית מנקודת מבט מתקדמת: גאומטריה, פליקס קליין

9D Tensor Products

Tensor Product of Two Vector Spaces

The motivation for our next topic comes from wanting to form the product of a vector $v \in V$ and a vector $w \in W$. This product will be denoted by $v \otimes w$, pronounced "v tensor w", and will be an element of some new vector space called $V \otimes W$ (also pronounced "V tensor W").

We already have a vector space $V \times W$ (see Section 3E), called the product of V and W. However, $V \times W$ will not serve our purposes here because it does not provide a natural way to multiply an element of V by an element of W. We would like our tensor product to satisfy some of the usual properties of multiplication. For example, we would like the distributive property to be satisfied, meaning that if $v_1, v_2, v \in V$ and $w_1, w_2, w \in W$, then

$$(v_1 + v_2) \otimes w = v_1 \otimes w + v_2 \otimes w \quad \text{and} \quad v \otimes (w_1 + w_2) = v \otimes w_1 + v \otimes w_2.$$

We would also like scalar multiplication to interact well with this new multiplication, meaning that

To produce $\otimes$ in TEX, type `\otimes`.

$$\lambda(v \otimes w) = (\lambda v) \otimes w = v \otimes (\lambda w)$$

for all $\lambda \in \mathbf{F}$, $v \in V$, and $w \in W$.

Furthermore, it would be nice if each basis of V when combined with each basis of W produced a basis of $V \otimes W$. Specifically, if $e_1, \dots, e_m$ is a basis of V and $f_1, \dots, f_n$ is a basis of W, then we would like a list (in any order) consisting of $e_j \otimes f_k$, as j ranges from 1 to m and k ranges from 1 to n, to be a basis of $V \otimes W$. This implies that $\dim(V \otimes W)$ should equal $(\dim V)(\dim W)$. Recall that $\dim(V \times W) = \dim V + \dim W$ (see 3.92), which shows that the product $V \times W$ will not serve our purposes here.

To produce a vector space whose dimension is $(\dim V)(\dim W)$ in a natural fashion from V and W, we look at the vector space of bilinear functionals, as defined below.

9.68 definition: *bilinear functional on $V \times W$, the vector space $\mathcal{B}(V, W)$*

- A *bilinear functional* on $V \times W$ is a function $\beta : V \times W \to \mathbf{F}$ such that $v \mapsto \beta(v, w)$ is a linear functional on V for each $w \in W$ and $w \mapsto \beta(v, w)$ is a linear functional on W for each $v \in V$.
- The vector space of bilinear functionals on $V \times W$ is denoted by $\mathcal{B}(V, W)$.

If $W = V$, then a bilinear functional on $V \times W$ is a bilinear form; see 9.1.

The operations of addition and scalar multiplication on $\mathcal{B}(V, W)$ are defined to be the usual operations of addition and scalar multiplication of functions. As you can verify, these operations make $\mathcal{B}(V, W)$ into a vector space whose additive identity is the zero function from $V \times W$ to $\mathbf{F}$.

ט§ מכפלות טנזוריות

מכפלה טנזורית של שני מרחבים וקטוריים

המוטיבציה לנושא הבא שבו נעסוק נובעת מהרצון להגדיר את המכפלה של וקטור $v \in V$ עם וקטור $w \in W$. מכפלה זו תסומן על ידי $v \otimes w$, מבוטאת "v טנזור w", ותהיה איבר במרחב וקטורי חדש כלשהו הנקרא $V \otimes W$ (גם הוא מבוטא כ־ "V טנזור W").

את המרחב הווקטורי $V \times W$ כבר הכרנו (ראו יחידה ג5), הנקרא המכפלה של V ו־ W. עם זאת, $V \times W$ לא יישרת את צרכינו מאחר שאינו מספק דרך טבעית להכפלת איבר של V עם איבר של W. אנו רוצים שהמכפלה הטנזורית תקיים חלק מהתכונות הרגילות של כפל. לדוגמה, אנו מעוניינים בקיומה של תכונת הפילוג, כלומר, אם $v_1, v_2, v \in V$ ו־ $w_1, w_2, w \in W$, אזי

$$v \otimes (w_1 + w_2) = v \otimes w_1 + v \otimes w_2 \quad \text{וגם} \quad (v_1 + v_2) \otimes w = v_1 \otimes w + v_2 \otimes w.$$

להפקת $\otimes$ ב־ TEX, הקלידו \otimes.

נרצה בנוסף כי כפל בסקלר יעבוד בצורה חלקה עם פעולת כפל חדשה זו, כלומר

$$\lambda(v \otimes w) = (\lambda v) \otimes w = v \otimes (\lambda w)$$

לכל $\lambda \in \mathbf{F}$, $v \in V$ ו־ $w \in W$.

יתר על כן, יהיה נחמד אם כל בסיס של V בשילוב עם כל בסיס של W יפיקו בסיס של $V \otimes W$. בפרט, אם $e_1, \dots, e_m$ בסיס של V ו־ $f_1, \dots, f_n$ בסיס של W, אזי נרצה שהרשימה (בכל סדר שהוא) המורכבת מ־ $e_j \otimes f_k$, כאשר j נע מ־ 1 עד m ו־ k נע מ־ 1 עד n, תהיה בסיס של $V \otimes W$. מכאן נובע כי $\dim(V \otimes W)$ צריך להיות שווה ל־ $(\dim V)(\dim W)$. ניזכר כי $\dim(V \times W) = \dim V + \dim W$ (ראו 3.92), מה שמסביר מדוע המכפלה $V \times W$ לא תשרת את צרכינו כאן.

להפקת מרחב וקטורי שממדו הוא $(\dim V)(\dim W)$ באופן טבעי מ־ V ו־ W, נסתכל על המרחב הווקטורי של פונקציונלים בילינאריים, כפי שהוא מוגדר להלן.

9.68 הגדרה: פונקציונל בילינארי מעל $V \times W$, המרחב הווקטורי $\mathcal{B}(V, W)$

- פונקציונל בילינארי מעל $V \times W$ הוא פונקציה $\beta : V \times W \to \mathbf{F}$ כך ש־ $v \mapsto \beta(v, w)$ הוא פונקציונל לינארי מעל V לכל $w \in W$ ו־ $w \mapsto \beta(v, w)$ פונקציונל לינארי מעל W לכל $v \in V$.
- המרחב הווקטורי של פונקציונלים בילינאריים מעל $V \times W$ מסומן על ידי $\mathcal{B}(V, W)$.

אם $W = V$, אזי פונקציונל בילינארי מעל $V \times W$ הוא תבנית בילינארית; ראו 9.1.

פעולות החיבור וכפל בסקלר מעל $\mathcal{B}(V, W)$ מוגדרות כפעולות החיבור והכפל בסקלר הרגילות של פונקציות. כפי שבאפשרותכם לוודא, פעולות אלו הופכות את $\mathcal{B}(V, W)$ למרחב וקטורי שאיבר הזהות שלו לחיבור הוא פונקציית האפס מ־ $V \times W$ אל $\mathbf{F}$.

9.69 example: *bilinear functionals*

- Suppose $\varphi \in V'$ and $\tau \in W'$. Define $\beta : V \times W \to \mathbf{F}$ by $\beta(v, w) = \varphi(v)\tau(w)$. Then β is a bilinear functional on $V \times W$.
- Suppose $v \in V$ and $w \in W$. Define $\beta : V' \times W' \to \mathbf{F}$ by $\beta(\varphi, \tau) = \varphi(v)\tau(w)$. Then β is a bilinear functional on $V' \times W'$.
- Define $\beta : V \times V' \to \mathbf{F}$ by $\beta(v, \varphi) = \varphi(v)$. Then β is a bilinear functional on $V \times V'$.
- Suppose $\varphi \in V'$. Define $\beta : V \times \mathcal{L}(V) \to \mathbf{F}$ by $\beta(v, T) = \varphi(Tv)$. Then β is a bilinear functional on $V \times \mathcal{L}(V)$.
- Suppose m and n are positive integers. Define $\beta : \mathbf{F}^{m,n} \times \mathbf{F}^{n,m} \to \mathbf{F}$ by $\beta(A, B) = \operatorname{tr}(AB)$. Then β is a bilinear functional on $\mathbf{F}^{m,n} \times \mathbf{F}^{n,m}$.

9.70 *dimension of the vector space of bilinear functionals*

$\dim \mathcal{B}(V, W) = (\dim V)(\dim W)$.

Proof Let $e_1, \dots, e_m$ be a basis of V and $f_1, \dots, f_n$ be a basis of W. For a bilinear functional $\beta \in \mathcal{B}(V, W)$, let $\mathcal{M}(\beta)$ be the m-by-n matrix whose entry in row j, column k is $\beta(e_j, f_k)$. The map $\beta \mapsto \mathcal{M}(\beta)$ is a linear map of $\mathcal{B}(V, W)$ into $\mathbf{F}^{m,n}$.

For a matrix $C \in \mathbf{F}^{m,n}$, define a bilinear functional β_C on $V \times W$ by

$$\beta_C(a_1 e_1 + \cdots + a_m e_m, b_1 f_1 + \cdots + b_n f_n) = \sum_{k=1}^{n} \sum_{j=1}^{m} C_{j,k} a_j b_k$$

for $a_1, \dots, a_m, b_1, \dots, b_n \in \mathbf{F}$.

The linear map $\beta \mapsto \mathcal{M}(\beta)$ from $\mathcal{B}(V, W)$ to $\mathbf{F}^{m,n}$ and the linear map $C \mapsto \beta_C$ from $\mathbf{F}^{m,n}$ to $\mathcal{B}(V, W)$ are inverses of each other because $\beta_{\mathcal{M}(\beta)} = \beta$ for all $\beta \in \mathcal{B}(V, W)$ and $\mathcal{M}(\beta_C) = C$ for all $C \in \mathbf{F}^{m,n}$, as you should verify.

Thus both maps are isomorphisms and the two spaces that they connect have the same dimension. Hence $\dim \mathcal{B}(V, W) = \dim \mathbf{F}^{m,n} = mn = (\dim V)(\dim W)$. ∎

Several different definitions of $V \otimes W$ appear in the mathematical literature. These definitions are equivalent to each other, at least in the finite-dimensional context, because any two vector spaces of the same dimension are isomorphic.

The result above states that $\mathcal{B}(V, W)$ has the dimension that we seek, as do $\mathcal{L}(V, W)$ and $\mathbf{F}^{\dim V, \dim W}$. Thus it may be tempting to define $V \otimes W$ to be $\mathcal{B}(V, W)$ or $\mathcal{L}(V, W)$ or $\mathbf{F}^{\dim V, \dim W}$. However, none of those definitions would lead to a basis-free definition of $v \otimes w$ for $v \in V$ and $w \in W$.

The following definition, while it may seem a bit strange and abstract at first, has the huge advantage that it defines $v \otimes w$ in a basis-free fashion. We define $V \otimes W$ to be the vector space of bilinear functionals on $V' \times W'$ instead of the more tempting choice of the vector space of bilinear functionals on $V \times W$.

9.69 דוגמה: פונקציונלים ביליניאריים

- נניח ש־ $\varphi \in V'$ ו־ $\tau \in W'$. נגדיר $\beta : V \times W \to \mathbf{F}$ על ידי $\beta(v,w) = \varphi(v)\tau(w)$. אזי β פונקציונל ביליניארי מעל $V \times W$.
- נניח ש־ $v \in V$ ו־ $w \in W$. נגדיר $\beta : V' \times W' \to \mathbf{F}$ על ידי $\beta(\varphi,\tau) = \varphi(v)\tau(w)$. אזי β פונקציונל ביליניארי מעל $V' \times W'$.
- נגדיר $\beta : V \times V' \to \mathbf{F}$ על ידי $\beta(v,\varphi) = \varphi(v)$. אזי β פונקציונל ביליניארי מעל $V \times V'$.
- נניח ש־ $\varphi \in V'$. נגדיר $\beta : V \times \mathcal{L}(V) \to \mathbf{F}$ על ידי $\beta(v,T) = \varphi(Tv)$. אזי β פונקציונל ביליניארי מעל $V \times \mathcal{L}(V)$.
- נניח ש־ m ו־ n מספרים שלמים חיוביים. נגדיר $\beta : \mathbf{F}^{m,n} \times \mathbf{F}^{n,m} \to \mathbf{F}$ על ידי $\beta(A,B) = \operatorname{tr}(AB)$. אזי β פונקציונל ביליניארי מעל $\mathbf{F}^{m,n} \times \mathbf{F}^{n,m}$.

9.70 ממד המרחב הווקטורי של פונקציונלים ביליניאריים

$$\dim \mathcal{B}(V,W) = (\dim V)(\dim W).$$

הוכחה תהי $e_1, \dots, e_m$ בסיס של V ו־ $f_1, \dots, f_n$ בסיס של W. עבור פונקציונל ביליניארי $\beta \in \mathcal{B}(V,W)$, תהי $\mathcal{M}(\beta)$ המטריצה $m \times n$ שהאיבר בשורה j, עמודה k הוא $\beta(e_j, f_k)$. ההעתקה $\beta \mapsto \mathcal{M}(\beta)$ היא העתקה ליניארית של $\mathcal{B}(V,W)$ אל $\mathbf{F}^{m,n}$.

עבור מטריצה $C \in \mathbf{F}^{m,n}$, נגדיר פונקציונל ביליניארי β_C מעל $V \times W$ על ידי

$$\beta_C(a_1e_1 + \cdots + a_me_m, b_1f_1 + \cdots + b_nf_n) = \sum_{k=1}^{n}\sum_{j=1}^{m} C_{j,k}a_jb_k$$

עבור $a_1, \dots, a_m, b_1, \dots, b_n \in \mathbf{F}$.

ההעתקה הליניארית $\beta \mapsto \mathcal{M}(\beta)$ מ־ $\mathcal{B}(V,W)$ אל $\mathbf{F}^{m,n}$ וההעתקה הליניארית $C \mapsto \beta_C$ מ־ $\mathbf{F}^{m,n}$ אל $\mathcal{B}(V,W)$ הן ההופכיות האחת של השנייה שכן $\beta_{\mathcal{M}(\beta)} = \beta$ לכל $\beta \in \mathcal{B}(V,W)$ ו־ $\mathcal{M}(\beta_C) = C$ לכל $C \in \mathbf{F}^{m,n}$, כפי שעליכם לוודא.

לפיכך שתי ההעתקות הן איזומורפיזמים ולשני המרחבים שהן מחברות יש ממד זהה. לכן $\dim \mathcal{B}(V,W) = \dim \mathbf{F}^{m,n} = mn = (\dim V)(\dim W)$. ■

מספר הגדרות שונות של $V \otimes W$ מופיעות בספרות המתמטית. הגדרות אלו שקולות אחת לשנייה, לכל הפחות בהקשר של ממד סופי, מאחר שכל שני מרחבים וקטוריים מממד זהה הם איזומורפיים.

המסקנה לעיל קובעת כי יש ל־ $\mathcal{B}(V,W)$ את הממד שביקשנו, כפי שיש ל־ $\mathcal{L}(V,W)$ ו־ $\mathbf{F}^{\dim V, \dim W}$. לפיכך אנו עלולים להתפתות ולהגדיר את $V \otimes W$ כ־ $\mathcal{B}(V,W)$ או $\mathcal{L}(V,W)$ או $\mathbf{F}^{\dim V, \dim W}$. עם זאת, אף אחת מהגדרות אלו לא תוביל להגדרה שאינה תלויה בבחירת בסיס של $v \otimes w$ עבור $v \in V$ ו־ $w \in W$.

להגדרה הבאה — אף שבהתרשמות ראשונה אולי תיראה מוזרה ומופשטת — יש יתרון עצום בכך שהיא מגדירה $v \otimes w$ באופן שאינו תלוי בבחירת בסיס. אנו מגדירים את $V \otimes W$ כמרחב הווקטורי של פונקציונלים ביליניאריים מעל $V' \times W'$ במקום האפשרות המפתה יותר של מרחב וקטורי של פונקציונלים ביליניאריים מעל $V \times W$.

9.71 definition: *tensor product, $V \otimes W$, $v \otimes w$*

- The *tensor product* $V \otimes W$ is defined to be $\mathcal{B}(V', W')$.
- For $v \in V$ and $w \in W$, the *tensor product* $v \otimes w$ is the element of $V \otimes W$ defined by
$$(v \otimes w)(\varphi, \tau) = \varphi(v)\tau(w)$$
for all $(\varphi, \tau) \in V' \times W'$.

We can quickly prove that the definition of $V \otimes W$ gives it the desired dimension.

9.72 *dimension of the tensor product of two vector spaces*

$$\dim(V \otimes W) = (\dim V)(\dim W).$$

Proof Because a vector space and its dual have the same dimension (by 3.111), we have $\dim V' = \dim V$ and $\dim W' = \dim W$. Thus 9.70 tells us that the dimension of $\mathcal{B}(V', W')$ equals $(\dim V)(\dim W)$. ∎

To understand the definition of the tensor product $v \otimes w$ of two vectors $v \in V$ and $w \in W$, focus on the kind of object it is. An element of $V \otimes W$ is a bilinear functional on $V' \times W'$, and in particular it is a function from $V' \times W'$ to $\mathbf{F}$. Thus for each element of $V' \times W'$, it should produce an element of $\mathbf{F}$. The definition above has this behavior, because $v \otimes w$ applied to a typical element (φ, τ) of $V' \times W'$ produces the number $\varphi(v)\tau(w)$.

The somewhat abstract nature of $v \otimes w$ should not matter. The important point is the behavior of these objects. The next result shows that tensor products of vectors have the desired bilinearity properties.

9.73 *bilinearity of tensor product*

Suppose $v, v_1, v_2 \in V$ and $w, w_1, w_2 \in W$ and $\lambda \in \mathbf{F}$. Then

$$(v_1 + v_2) \otimes w = v_1 \otimes w + v_2 \otimes w \quad \text{and} \quad v \otimes (w_1 + w_2) = v \otimes w_1 + v \otimes w_2$$

and

$$\lambda(v \otimes w) = (\lambda v) \otimes w = v \otimes (\lambda w).$$

Proof Suppose $(\varphi, \tau) \in V' \times W'$. Then

$$\begin{aligned}((v_1 + v_2) \otimes w)(\varphi, \tau) &= \varphi(v_1 + v_2)\tau(w) \\ &= \varphi(v_1)\tau(w) + \varphi(v_2)\tau(w) \\ &= (v_1 \otimes w)(\varphi, \tau) + (v_2 \otimes w)(\varphi, \tau) \\ &= (v_1 \otimes w + v_2 \otimes w)(\varphi, \tau).\end{aligned}$$

Thus $(v_1 + v_2) \otimes w = v_1 \otimes w + v_2 \otimes w$.

The other two equalities are proved similarly. ∎

9.71 הגדרה: מכפלה טנזורית, $V \otimes W$, $v \otimes w$

- המכפלה הטנזורית $V \otimes W$ מוגדרת כ־ $\mathcal{B}(V', W')$.
- עבור $v \in V$ ו־ $w \in W$, המכפלה הטנזורית $v \otimes w$ היא איבר של $V \otimes W$ המוגדר על ידי
$$(v \otimes w)(\varphi, \tau) = \varphi(v)\tau(w)$$
לכל $(\varphi, \tau) \in V' \times W'$.

ביכולתנו להוכיח בזריזות כי הגדרת $V \otimes W$ מעניקה לו את הממד הרצוי.

9.72 ממד המכפלה הטנזורית של שני מרחבים וקטוריים

$$\dim(V \otimes W) = (\dim V)(\dim W).$$

הוכחה מאחר שלמרחב וקטורי ולדואלי שלו ממד זהה (לפי 3.111), מתקיים $\dim V' = \dim V$ ו־ $\dim W' = \dim W$. לפיכך 9.70 אומרת לנו כי הממד של $\mathcal{B}(V', W')$ שווה $(\dim V)(\dim W)$. ■

להבנת הגדרת המכפלה הטנזורית $v \otimes w$ של שני וקטורים $v \in V$ ו־ $w \in W$, התמקדו בסוג האובייקט הזה. איבר של $V \otimes W$ הוא פונקציונל ביליניארי מעל $V' \times W'$, ובפרט הוא פונקציה מ־ $V' \times W'$ אל $\mathbf{F}$. לפיכך עבור כל איבר של $V' \times W'$, היא אמורה להפיק איבר של $\mathbf{F}$. ההגדרה לעיל מוודאת שכך תתנהג הפונקציה, מאחר ש־ $v \otimes w$ המופעלת על איבר טיפוסי (φ, τ) של $V' \times W'$ מפיקה את המספר $\varphi(v)\tau(w)$.

האופי המופשט במידת מה של $v \otimes w$ לא אמור להטריד את מנוחתכם. הנקודה החשובה היא התנהגות אובייקטים אלו. המסקנה הבאה מראה שלמכפלות טנזוריות של וקטורים תכונות הבילינאריות הרצויות.

9.73 בילינאריות של מכפלה טנזורית

נניח ש־ $v, v_1, v_2 \in V$ ו־ $w, w_1, w_2 \in W$ ו־ $\lambda \in \mathbf{F}$. אזי
$v \otimes (w_1 + w_2) = v \otimes w_1 + v \otimes w_2$ וגם $(v_1 + v_2) \otimes w = v_1 \otimes w + v_2 \otimes w$
וגם
$$\lambda(v \otimes w) = (\lambda v) \otimes w = v \otimes (\lambda w).$$

הוכחה נניח ש־ $(\varphi, \tau) \in V' \times W'$. אזי
$$\begin{aligned}((v_1 + v_2) \otimes w)(\varphi, \tau) &= \varphi(v_1 + v_2)\tau(w) \\ &= \varphi(v_1)\tau(w) + \varphi(v_2)\tau(w) \\ &= (v_1 \otimes w)(\varphi, \tau) + (v_2 \otimes w)(\varphi, \tau) \\ &= (v_1 \otimes w + v_2 \otimes w)(\varphi, \tau).\end{aligned}$$
לפיכך $(v_1 + v_2) \otimes w = v_1 \otimes w + v_2 \otimes w$.
את שני השוויונות הנוספים ניתן להוכיח באופן דומה. ■

Lists are, by definition, ordered. The order matters when, for example, we form the matrix of an operator with respect to a basis. For lists in this section with two indices, such as $\{e_j \otimes f_k\}_{j=1,\dots,m;k=1,\dots,n}$ in the next result, the ordering does not matter and we do not specify it—just choose any convenient ordering.

The linear independence of elements of $V \otimes W$ in (a) of the result below captures the idea that there are no relationships among vectors in $V \otimes W$ other than the relationships that come from bilinearity of the tensor product (see 9.73) and the relationships that may be present due to linear dependence of a list of vectors in V or a list of vectors in W.

9.74 *basis of $V \otimes W$*

Suppose $e_1, \dots, e_m$ is a list of vectors in V and $f_1, \dots, f_n$ is a list of vectors in W.

(a) If $e_1, \dots, e_m$ and $f_1, \dots, f_n$ are both linearly independent lists, then

$$\{e_j \otimes f_k\}_{j=1,\dots,m;k=1,\dots,n}$$

is a linearly independent list in $V \otimes W$.

(b) If $e_1, \dots, e_m$ is a basis of V and $f_1, \dots, f_n$ is a basis of W, then the list $\{e_j \otimes f_k\}_{j=1,\dots,m;k=1,\dots,n}$ is a basis of $V \otimes W$.

Proof To prove (a), suppose $e_1, \dots, e_m$ and $f_1, \dots, f_n$ are both linearly independent lists. This linear independence and the linear map lemma (3.4) imply that there exist $\varphi_1, \dots, \varphi_m \in V'$ and $\tau_1, \dots, \tau_n \in W'$ such that

$$\varphi_j(e_k) = \begin{cases} 1 & \text{if } j = k, \\ 0 & \text{if } j \neq k \end{cases} \quad \text{and} \quad \tau_j(f_k) = \begin{cases} 1 & \text{if } j = k, \\ 0 & \text{if } j \neq k, \end{cases}$$

where $j, k \in \{1, \dots, m\}$ in the first equation and $j, k \in \{1, \dots, n\}$ in the second equation.

Suppose $\{a_{j,k}\}_{j=1,\dots,m;k=1,\dots,n}$ is a list of scalars such that

$$\sum_{k=1}^{n} \sum_{j=1}^{m} a_{j,k}(e_j \otimes f_k) = 0. \tag{9.75}$$

Note that $(e_j \otimes f_k)(\varphi_M, \tau_N)$ equals 1 if $j = M$ and $k = N$, and equals 0 otherwise. Thus applying both sides of 9.75 to (φ_M, τ_N) shows that $a_{M,N} = 0$, proving that $\{e_j \otimes f_k\}_{j=1,\dots,m;k=1,\dots,n}$ is linearly independent.

Now (b) follows from (a), the equation $\dim V \otimes W = (\dim V)(\dim W)$ [see 9.72], and the result that a linearly independent list of the right length is a basis (see 2.38). ∎

Every element of $V \otimes W$ is a finite sum of elements of the form $v \otimes w$, where $v \in V$ and $w \in W$, as implied by (b) in the result above. However, if $\dim V > 1$ and $\dim W > 1$, then Exercise 4 shows that

$$\{v \otimes w : (v, w) \in V \times W\} \neq V \otimes W.$$

רשימות הן, לפי הגדרה, סדורות. לסדר יש חשיבות כאשר, לדוגמה, אנו יוצרים את המטריצה של אופרטור ביחס לבסיס. עבור רשימות ביחידה זו עם שני אינדקסים, דוגמת $\{e_j \otimes f_k\}_{j=1,\dots,m;k=1,\dots,n}$ שבמסקנה הבאה, הסדר לא משפיע ואיננו מציינים אותו—ניתן לבחור כל סדר לפי נוחות.

אי התלות הליניארית של איברי $V \otimes W$ ב־ (א) של המסקנה להלן לוכדת את הרעיון כי אין קשרים בין וקטורים ב־ $V \otimes W$ מלבד הקשרים הנובעים מהבילינאריות של המכפלה הטנזורית (ראו 9.73) והקשרים שאולי יופיעו כתוצאה מתלות ליניארית של רשימת וקטורים ב־ V או רשימת וקטורים ב־ W.

9.74 בסיס של $V \otimes W$

נניח ש־ $e_1, \dots, e_m$ רשימת וקטורים ב־ V ו־ $f_1, \dots, f_n$ רשימת וקטורים ב־ W.

א. אם $e_1, \dots, e_m$ ו־ $f_1, \dots, f_n$ שתיהן רשימות בלתי תלויות ליניארית, אזי
$$\{e_j \otimes f_k\}_{j=1,\dots,m;k=1,\dots,n}$$
היא רשימה בלתי תלויה ליניארית ב־ $V \otimes W$.

ב. אם $e_1, \dots, e_m$ בסיס של V ו־ $f_1, \dots, f_n$ בסיס של W, אזי הרשימה $\{e_j \otimes f_k\}_{j=1,\dots,m;k=1,\dots,n}$ היא בסיס של $V \otimes W$.

הוכחה להוכחת (א), נניח ש־ $e_1, \dots, e_m$ ו־ $f_1, \dots, f_n$ שתיהן רשימות בלתי תלויות ליניארית. אי־תלות ליניארית זו ולמת ההעתקה הליניארית (3.4) גוררים שקיימים $\varphi_1, \dots, \varphi_m \in V'$ ו־ $\tau_1, \dots, \tau_n \in W'$ כך ש־
$$\varphi_j(e_k) = \begin{cases} 1 & \text{if } j = k, \\ 0 & \text{if } j \neq k \end{cases} \quad \text{וגם} \quad \tau_j(f_k) = \begin{cases} 1 & \text{if } j = k, \\ 0 & \text{if } j \neq k, \end{cases}$$
כאשר $j, k \in \{1, \dots, m\}$ במשוואה הראשונה ו־ $j, k \in \{1, \dots, n\}$ במשוואה השנייה.

נניח ש־ $\{a_{j,k}\}_{j=1,\dots,m;k=1,\dots,n}$ רשימת סקלרים כך ש־
$$\sum_{k=1}^{n} \sum_{j=1}^{m} a_{j,k}(e_j \otimes f_k) = 0. \tag{9.75}$$
שימו לב ש־ $(e_j \otimes f_k)(\varphi_M, \tau_N)$ שווה 1 אם $j = M$ ו־ $k = N$, ושווה 0 בכל מקרה אחר. לפיכך הפעלת שני אגפי 9.75 על (φ_M, τ_N) מראה ש־ $a_{M,N} = 0$, ובכך הוכחנו כי $\{e_j \otimes f_k\}_{j=1,\dots,m;k=1,\dots,n}$ בלתי תלויה ליניארית.

כעת (ב) נובע מ־ (א), המשוואה $\dim V \otimes W = (\dim V)(\dim W)$ [ראו 9.72], ומהמסקנה כי רשימה בלתי תלויה ליניארית באורך הנכון היא בסיס (ראו 2.38). ■

כל איבר של $V \otimes W$ הוא סכום סופי של איברים מהצורה $v \otimes w$, כאשר $v \in V$ ו־ $w \in W$, כפי שנובע מסעיף (ב) במסקנה לעיל. עם זאת, אם $\dim V > 1$ ו־ $\dim W > 1$, אזי תרגיל 4 מראה כי
$$\{v \otimes w : (v, w) \in V \times W\} \neq V \otimes W.$$

9.76 example: *tensor product of element of* $\mathbf{F}^m$ *with element of* $\mathbf{F}^n$

Suppose m and n are positive integers. Let $e_1, \dots, e_m$ denote the standard basis of $\mathbf{F}^m$ and let $f_1, \dots, f_n$ denote the standard basis of $\mathbf{F}^n$. Suppose

$$v = (v_1, \dots, v_m) \in \mathbf{F}^m \quad \text{and} \quad w = (w_1, \dots, w_n) \in \mathbf{F}^n.$$

Then

$$\begin{aligned} v \otimes w &= \Big(\sum_{j=1}^{m} v_j e_j\Big) \otimes \Big(\sum_{k=1}^{n} w_k f_k\Big) \\ &= \sum_{k=1}^{n} \sum_{j=1}^{m} (v_j w_k)(e_j \otimes f_k). \end{aligned}$$

Thus with respect to the basis $\{e_j \otimes f_k\}_{j=1,\dots,m;k=1,\dots,n}$ of $\mathbf{F}^m \otimes \mathbf{F}^n$ provided by 9.74(b), the coefficients of $v \otimes w$ are the numbers $\{v_j w_k\}_{j=1,\dots,m;k=1,\dots,n}$. If instead of writing these numbers in a list, we write them in an m-by-n matrix with $v_j w_k$ in row j, column k, then we can identify $v \otimes w$ with the m-by-n matrix

$$\begin{pmatrix} v_1 w_1 & \cdots & v_1 w_n \\ & \ddots & \\ v_m w_1 & \cdots & v_m w_n \end{pmatrix}.$$

See Exercises 5 and 6 for practice in using the identification from the example above.

We now define bilinear maps, which differ from bilinear functionals in that the target space can be an arbitrary vector space rather than just the scalar field.

9.77 definition: *bilinear map*

A *bilinear map* from $V \times W$ to a vector space U is a function $\Gamma : V \times W \to U$ such that $v \mapsto \Gamma(v, w)$ is a linear map from V to U for each $w \in W$ and $w \mapsto \Gamma(v, w)$ is a linear map from W to U for each $v \in V$.

9.78 example: *bilinear maps*

- Every bilinear functional on $V \times W$ is a bilinear map from $V \times W$ to $\mathbf{F}$.
- The function $\Gamma : V \times W \to V \otimes W$ defined by $\Gamma(v, w) = v \otimes w$ is a bilinear map from $V \times W$ to $V \otimes W$ (by 9.73).
- The function $\Gamma : \mathcal{L}(V) \times \mathcal{L}(V) \to \mathcal{L}(V)$ defined by $\Gamma(S, T) = ST$ is a bilinear map from $\mathcal{L}(V) \times \mathcal{L}(V)$ to $\mathcal{L}(V)$.
- The function $\Gamma : V \times \mathcal{L}(V, W) \to W$ defined by $\Gamma(v, T) = Tv$ is a bilinear map from $V \times \mathcal{L}(V, W)$ to W.

9.76 דוגמה: מכפלה טנזורית של איבר של $\mathbf{F}^m$ עם איבר של $\mathbf{F}^n$

נניח ש־ m ו־ n מספרים שלמים חיוביים. נסמן ב־ $e_1, \dots, e_m$ את הבסיס הסטנדרטי של $\mathbf{F}^m$ וב־ $f_1, \dots, f_n$ את הבסיס הסטנדרטי של $\mathbf{F}^n$. נניח ש־

$$v = (v_1, \dots, v_m) \in \mathbf{F}^m \quad \text{וגם} \quad w = (w_1, \dots, w_n) \in \mathbf{F}^n.$$

אזי

$$\begin{aligned} v \otimes w &= \Big(\sum_{j=1}^{m} v_j e_j\Big) \otimes \Big(\sum_{k=1}^{n} w_k f_k\Big) \\ &= \sum_{k=1}^{n} \sum_{j=1}^{m} (v_j w_k)(e_j \otimes f_k). \end{aligned}$$

לפיכך ביחס אל הבסיס $\{e_j \otimes f_k\}_{j=1,\dots,m;k=1,\dots,n}$ של $\mathbf{F}^m \otimes \mathbf{F}^n$ שניתן על ידי 9.74(ב), המקדמים של $v \otimes w$ הם המספרים $\{v_j w_k\}_{j=1,\dots,m;k=1,\dots,n}$. אם במקום לכתוב מספרים אלו כרשימה, נכתוב אותם כמטריצה $m \times n$ עם $v_j w_k$ בשורה j, עמודה k, אזי נוכל לזהות את $v \otimes w$ עם המטריצה $m \times n$

$$\begin{pmatrix} v_1 w_1 & \cdots & v_1 w_n \\ & \ddots & \\ v_m w_1 & \cdots & v_m w_n \end{pmatrix}.$$

ראו תרגילים 5 ו־ 6 לתרגול השימוש בזיהוי מהדוגמה לעיל.

כעת אנו מגדירים העתקות בליניאריות, השונות מפונקציונלים בליניאריים בכך שמרחב היעד יכול להיות מרחב וקטורי שרירותי במקום רק שדה הסקלרים.

9.77 הגדרה: העתקה בילינארית

העתקה בילינארית מ־ $V \times W$ אל מרחב וקטורי U היא פונקציה $\Gamma : V \times W \to U$ כך ש־ $v \mapsto \Gamma(v, w)$ היא העתקה לינארית מ־ V אל U לכל $w \in W$ ו־ $w \mapsto \Gamma(v, w)$ העתקה לינארית מ־ W אל U לכל $v \in V$.

9.78 דוגמה: העתקות בילינאריות

- כל פונקציונל בילינארי מעל $V \times W$ הוא העתקה בילינארית מ־ $V \times W$ אל $\mathbf{F}$.
- הפונקציה $\Gamma : V \times W \to V \otimes W$ המוגדרת על ידי $\Gamma(v, w) = v \otimes w$ היא העתקה בילינארית מ־ $V \times W$ אל $V \otimes W$ (לפי 9.73).
- הפונקציה $\Gamma : \mathcal{L}(V) \times \mathcal{L}(V) \to \mathcal{L}(V)$ המוגדרת על ידי $\Gamma(S, T) = ST$ היא העתקה בילינארית מ־ $\mathcal{L}(V) \times \mathcal{L}(V)$ אל $\mathcal{L}(V)$.
- הפונקציה $\Gamma : V \times \mathcal{L}(V, W) \to W$ המוגדרת על ידי $\Gamma(v, T) = Tv$ היא העתקה בילינארית מ־ $V \times \mathcal{L}(V, W)$ אל W.

Tensor products allow us to convert bilinear maps on $V \times W$ into linear maps on $V \otimes W$ (and vice versa), as shown by the next result. In the mathematical literature, (a) of the result below is called the "universal property" of tensor products.

9.79 *converting bilinear maps to linear maps*

Suppose U is a vector space.

(a) Suppose $\Gamma : V \times W \to U$ is a bilinear map. Then there exists a unique linear map $\hat{\Gamma} : V \otimes W \to U$ such that

$$\hat{\Gamma}(v \otimes w) = \Gamma(v, w)$$

for all $(v, w) \in V \times W$.

(b) Conversely, suppose $T : V \otimes W \to U$ is a linear map. Then there exists a unique bilinear map $T^{\#} : V \times W \to U$ such that

$$T^{\#}(v, w) = T(v \otimes w)$$

for all $(v, w) \in V \times W$.

Proof Let $e_1, \dots, e_m$ be a basis of V and let $f_1, \dots, f_n$ be a basis of W. By the linear map lemma (3.4) and 9.74(b), there exists a unique linear map $\hat{\Gamma} : V \otimes W \to U$ such that

$$\hat{\Gamma}(e_j \otimes f_k) = \Gamma(e_j, f_k)$$

for all $j \in \{1, \dots, m\}$ and $k \in \{1, \dots, n\}$.

Now suppose $(v, w) \in V \times W$. There exist $a_1, \dots, a_m, b_1, \dots, b_n \in \mathbf{F}$ such that $v = a_1 e_1 + \cdots + a_m e_m$ and $w = b_1 f_1 + \cdots + b_n f_n$. Thus

$$\begin{aligned}
\hat{\Gamma}(v \otimes w) &= \hat{\Gamma}\Big(\sum_{k=1}^{n} \sum_{j=1}^{m} (a_j b_k)(e_j \otimes f_k)\Big) \\
&= \sum_{k=1}^{n} \sum_{j=1}^{m} a_j b_k \hat{\Gamma}(e_j \otimes f_k) \\
&= \sum_{k=1}^{n} \sum_{j=1}^{m} a_j b_k \Gamma(e_j, f_k) \\
&= \Gamma(v, w),
\end{aligned}$$

as desired, where the second line holds because $\hat{\Gamma}$ is linear, the third line holds by the definition of $\hat{\Gamma}$, and the fourth line holds because Γ is bilinear.

The uniqueness of the linear map $\hat{\Gamma}$ satisfying $\hat{\Gamma}(v \otimes w) = \Gamma(v, w)$ follows from 9.74(b), completing the proof of (a).

To prove (b), define a function $T^{\#} : V \times W \to U$ by $T^{\#}(v, w) = T(v \otimes w)$ for all $(v, w) \in V \times W$. The bilinearity of the tensor product (see 9.73) and the linearity of T imply that $T^{\#}$ is bilinear.

Clearly the choice of $T^{\#}$ that satisfies the conditions is unique. ∎

מכפלות טנזוריות מאפשרות לנו להמיר העתקות בילינאריות מעל $V\times W$ להעתקות לינאריות מעל $V\otimes W$ (ולהיפך), כפי שמראה לנו המסקנה הבאה. בספרות המתמטית, סעיף (א) של המסקנה להלן נקרא "התכונה האוניברסלית" של מכפלות טנזוריות.

9.79 המרת העתקות בילינאריות להעתקות לינאריות

נניח ש־ U מרחב וקטורי.

א. נניח ש־ $\Gamma: V\times W\to U$ העתקה בילינארית. אזי קיימת העתקה לינארית יחידה $\hat{\Gamma}: V\otimes W\to U$ כך ש־

$$\hat{\Gamma}(v\otimes w)=\Gamma(v,w)$$

לכל $(v,w)\in V\times W$.

ב. בכיוון ההפוך, נניח ש־ $T: V\otimes W\to U$ העתקה לינארית. אזי קיימת העתקה בילינארית יחידה $T^{\#}: V\times W\to U$ כך ש־

$$T^{\#}(v,w)=T(v\otimes w)$$

לכל $(v,w)\in V\times W$.

הוכחה יהי $e_1,\dots,e_m$ בסיס של V ויהי $f_1,\dots,f_n$ בסיס של W. באמצעות למת ההעתקה הלינארית (3.4) ו־9.74(ב), קיימת העתקה לינארית יחידה $\hat{\Gamma}: V\otimes W\to U$ כך ש־

$$\hat{\Gamma}(e_j\otimes f_k)=\Gamma(e_j,f_k)$$

לכל $j\in\{1,\dots,m\}$ ו־ $k\in\{1,\dots,n\}$.

כעת נניח ש־ $(v,w)\in V\times W$. קיימים $a_1,\dots,a_m,b_1,\dots,b_n\in\mathbf{F}$ כך ש־ $v=a_1e_1+\cdots+a_me_m$ ו־ $w=b_1f_1+\cdots+b_nf_n$. לפיכך

$$\begin{aligned}\hat{\Gamma}(v\otimes w)&=\hat{\Gamma}\Big(\sum_{k=1}^{n}\sum_{j=1}^{m}(a_jb_k)(e_j\otimes f_k)\Big)\\&=\sum_{k=1}^{n}\sum_{j=1}^{m}a_jb_k\hat{\Gamma}(e_j\otimes f_k)\\&=\sum_{k=1}^{n}\sum_{j=1}^{m}a_jb_k\Gamma(e_j,f_k)\\&=\Gamma(v,w),\end{aligned}$$

כפי שרצינו, כאשר השורה השנייה מתקיימת מאחר ש־ $\hat{\Gamma}$ לינארית, השורה השלישית מתקיימת עקב ההגדרה של $\hat{\Gamma}$, והשורה הרביעית מתקיימת מאחר ש־ Γ בילינארית.

היחידות של ההעתקה לינארית $\hat{\Gamma}$ המקיימת $\hat{\Gamma}(v\otimes w)=\Gamma(v,w)$ נובעת מ־9.74(ב), ובכך השלמנו את הוכחת (א).

להוכחת (ב), נגדיר פונקציה $T^{\#}: V\times W\to U$ על ידי $T^{\#}(v,w)=T(v\otimes w)$ לכל $(v,w)\in V\times W$. הבילינאריות של המכפלה הטנזורית (ראו 9.73) והלינאריות של T גוררים ש־ $T^{\#}$ בילינארית.

בחירה של $T^{\#}$ המקיימת את התנאים היא בוודאות יחידה.

To prove 9.79(a), we could not just define $\hat{\Gamma}(v \otimes w) = \Gamma(v, w)$ for all $v \in V$ and $w \in W$ (and then extend $\hat{\Gamma}$ linearly to all of $V \otimes W$) because elements of $V \otimes W$ do not have unique representations as finite sums of elements of the form $v \otimes w$. Our proof used a basis of V and a basis of W to get around this problem.

Although our construction of $\hat{\Gamma}$ in the proof of 9.79(a) depended on a basis of V and a basis of W, the equation $\hat{\Gamma}(v \otimes w) = \Gamma(v, w)$ that holds for all $v \in V$ and $w \in W$ shows that $\hat{\Gamma}$ does not depend on the choice of bases for V and W.

Tensor Product of Inner Product Spaces

The result below features three inner products—one on $V \otimes W$, one on V, and one on W, although we use the same symbol $\langle \cdot, \cdot \rangle$ for all three inner products.

9.80 *inner product on tensor product of two inner product spaces*

Suppose V and W are inner product spaces. Then there is a unique inner product on $V \otimes W$ such that

$$\langle v \otimes w, u \otimes x \rangle = \langle v, u \rangle \langle w, x \rangle$$

for all $v, u \in V$ and $w, x \in W$.

Proof Suppose $e_1, \dots, e_m$ is an orthonormal basis of V and $f_1, \dots, f_n$ is an orthonormal basis of W. Define an inner product on $V \otimes W$ by

$$\text{9.81} \qquad \left\langle \sum_{k=1}^{n} \sum_{j=1}^{m} b_{j,k} e_j \otimes f_k, \sum_{k=1}^{n} \sum_{j=1}^{m} c_{j,k} e_j \otimes f_k \right\rangle = \sum_{k=1}^{n} \sum_{j=1}^{m} b_{j,k} \overline{c_{j,k}}.$$

The straightforward verification that 9.81 defines an inner product on $V \otimes W$ is left to the reader [use 9.74(b)].

Suppose that $v, u \in V$ and $w, x \in W$. Let $v_1, \dots, v_m \in \mathbf{F}$ be such that $v = v_1 e_1 + \cdots + v_m e_m$, with similar expressions for u, w, and x. Then

$$\begin{aligned}
\langle v \otimes w, u \otimes x \rangle &= \left\langle \sum_{j=1}^{m} v_j e_j \otimes \sum_{k=1}^{n} w_k f_k, \sum_{j=1}^{m} u_j e_j \otimes \sum_{k=1}^{n} x_k f_k \right\rangle \\
&= \left\langle \sum_{k=1}^{n} \sum_{j=1}^{m} v_j w_k e_j \otimes f_k, \sum_{k=1}^{n} \sum_{j=1}^{m} u_j x_k e_j \otimes f_k \right\rangle \\
&= \sum_{k=1}^{n} \sum_{j=1}^{m} v_j \overline{u_j} w_k \overline{x_k} \\
&= \left(\sum_{j=1}^{m} v_j \overline{u_j} \right) \left(\sum_{k=1}^{n} w_k \overline{x_k} \right) \\
&= \langle v, u \rangle \langle w, x \rangle.
\end{aligned}$$

There is only one inner product on $V \otimes W$ such that $\langle v \otimes w, u \otimes x \rangle = \langle v, u \rangle \langle w, x \rangle$ for all $v, u \in V$ and $w, x \in W$ because every element of $V \otimes W$ can be written as a linear combination of elements of the form $v \otimes w$ [by 9.74(b)]. ∎

להוכחת 9.79(א), לא היינו יכולים פשוט להגדיר $\hat{\Gamma}(v\otimes w)=\Gamma(v,w)$ לכל $v\in V$ ו־ $w\in W$ (ואז להרחיב את $\hat{\Gamma}$ באופן ליניארי לכל איברי $V\otimes W$) מאחר שאין לאיברים של $V\otimes W$ ייצוג יחיד כסכום סופי מהצורה $v\otimes w$. ההוכחה שלנו משתמשת בבסיס של V ובסיס של W לעקוף בעיה זו.

אף שבניית $\hat{\Gamma}$ בהוכחת 9.79(א) תלויה בבסיס של V ובסיס של W, המשוואה $\hat{\Gamma}(v\otimes w)=\Gamma(v,w)$ המתקיימת לכל $v\in V$ ו־ $w\in W$ מראה כי $\hat{\Gamma}$ אינה תלויה בבחירת בסיסים עבור V ו־ W.

מכפלה טנזורית של מרחבי מכפלה פנימית

במסקנה להלן מופיעות שלוש מכפלות פנימיות שונות—אחת מעל $V\otimes W$, אחת מעל V ואחת מעל W, אף שאנו משתמשים באותו הסמל $\langle\cdot,\cdot\rangle$ לכל שלושת המכפלות הפנימיות.

9.80 מכפלה פנימית מעל מכפלה טנזורית של שני מרחבי מכפלה פנימית

נניח ש־ V ו־ W מרחבי מכפלה פנימית. אזי קיימת מכפלה פנימית יחידה מעל $V\otimes W$ כך ש־

$$\langle v\otimes w,u\otimes x\rangle=\langle v,u\rangle\langle w,x\rangle$$

לכל $v,u\in V$ ו־ $w,x\in W$.

הוכחה נניח ש־ $e_1,\dots,e_m$ בסיס אורתונורמלי של V ו־ $f_1,\dots,f_n$ בסיס אורתונורמלי של W. נגדיר מכפלה פנימית מעל $V\otimes W$ על ידי

$$\left\langle\sum_{k=1}^{n}\sum_{j=1}^{m}b_{j,k}\,e_j\otimes f_k,\sum_{k=1}^{n}\sum_{j=1}^{m}c_{j,k}\,e_j\otimes f_k\right\rangle=\sum_{k=1}^{n}\sum_{j=1}^{m}b_{j,k}\overline{c_{j,k}}. \tag{9.81}$$

הווידוא המיידי ש־ 9.81 מגדירה מכפלה פנימית מעל $V\otimes W$ מושאר לקוראים [השתמשו ב־ 9.74(ב)].

נניח ש־ $v,u\in V$ ו־ $w,x\in W$. יהי $v_1,\dots,v_m\in\mathbf{F}$ כך ש־ $v=v_1e_1+\cdots+v_me_m$, וביטויים דומים עבור u, w ו־ x. אזי

$$\begin{aligned}\langle v\otimes w,u\otimes x\rangle&=\left\langle\sum_{j=1}^{m}v_je_j\otimes\sum_{k=1}^{n}w_kf_k,\sum_{j=1}^{m}u_je_j\otimes\sum_{k=1}^{n}x_kf_k\right\rangle\\&=\left\langle\sum_{k=1}^{n}\sum_{j=1}^{m}v_jw_k\,e_j\otimes f_k,\sum_{k=1}^{n}\sum_{j=1}^{m}u_jx_k\,e_j\otimes f_k\right\rangle\\&=\sum_{k=1}^{n}\sum_{j=1}^{m}v_j\overline{u_j}w_k\overline{x_k}\\&=\left(\sum_{j=1}^{m}v_j\overline{u_j}\right)\left(\sum_{k=1}^{n}w_k\overline{x_k}\right)\\&=\langle v,u\rangle\langle w,x\rangle.\end{aligned}$$

קיימת מכפלה פנימית יחידה מעל $V\otimes W$ כך ש־ $\langle v\otimes w,u\otimes x\rangle=\langle v,u\rangle\langle w,x\rangle$ לכל $v,u\in V$ ו־ $w,x\in W$ מאחר שכל איבר של $V\otimes W$ ניתן לכתיבה כצירוף ליניארי של איברים מהצורה $v\otimes w$ [לפי 9.74(ב)].

The definition below of a natural inner product on $V \otimes W$ is now justified by 9.80. We could not have simply defined $\langle v \otimes w, u \otimes x\rangle$ to be $\langle v, u\rangle\langle w, x\rangle$ (and then used additivity in each slot separately to extend the definition to $V \otimes W$) without some proof because elements of $V \otimes W$ do not have unique representations as finite sums of elements of the form $v \otimes w$.

9.82 definition: *inner product on tensor product of two inner product spaces*

Suppose V and W are inner product spaces. The inner product on $V \otimes W$ is the unique function $\langle \cdot, \cdot \rangle$ from $(V \otimes W) \times (V \otimes W)$ to $\mathbf{F}$ such that

$$\langle v \otimes w, u \otimes x\rangle = \langle v, u\rangle\langle w, x\rangle$$

for all $v, u \in V$ and $w, x \in W$.

Take $u = v$ and $x = w$ in the equation above and then take square roots to show that

$$\|v \otimes w\| = \|v\| \, \|w\|$$

for all $v \in V$ and all $w \in W$.

The construction of the inner product in the proof of 9.80 depended on an orthonormal basis $e_1, \dots, e_m$ of V and an orthonormal basis $f_1, \dots, f_n$ of W. Formula 9.81 implies that $\{e_j \otimes f_k\}_{j=1,\dots,m;k=1,\dots,n}$ is a doubly indexed orthonormal list in $V \otimes W$ and hence is an orthonormal basis of $V \otimes W$ [by 9.74(b)]. The importance of the next result arises because the orthonormal bases used there can be different from the orthonormal bases used to define the inner product in 9.80. Although the notation for the bases is the same in the proof of 9.80 and in the result below, think of them as two different sets of orthonormal bases.

9.83 *orthonormal basis of $V \otimes W$*

Suppose V and W are inner product spaces, and $e_1, \dots, e_m$ is an orthonormal basis of V and $f_1, \dots, f_n$ is an orthonormal basis of W. Then

$$\{e_j \otimes f_k\}_{j=1,\dots,m;k=1,\dots,n}$$

is an orthonormal basis of $V \otimes W$.

Proof We know that $\{e_j \otimes f_k\}_{j=1,\dots,m;k=1,\dots,n}$ is a basis of $V \otimes W$ [by 9.74(b)]. Thus we only need to verify orthonormality. To do this, suppose $j, M \in \{1, \dots, m\}$ and $k, N \in \{1, \dots, n\}$. Then

$$\langle e_j \otimes f_k, e_M \otimes f_N\rangle = \langle e_j, e_M\rangle\langle f_k, f_N\rangle = \begin{cases} 1 & \text{if } j = M \text{ and } k = N, \\ 0 & \text{otherwise.} \end{cases}$$

Hence the doubly indexed list $\{e_j \otimes f_k\}_{j=1,\dots,m;k=1,\dots,n}$ is indeed an orthonormal basis of $V \otimes W$. ∎

See Exercise 11 for an example of how the inner product structure on $V \otimes W$ interacts with operators on V and W.

ההגדרה להלן של מכפלה פנימית טבעית מעל $V \otimes W$ כעת מוצדקת לפי 9.80. לא היינו יכולים פשוט להגדיר את $\langle v \otimes w, u \otimes x \rangle$ כ־ $\langle v, u \rangle \langle w, x \rangle$ (ואז להשתמש באדיטיביות בכל רכיב בנפרד להרחבת ההגדרה ל־ $V \otimes W$) ללא הוכחה כלשהי מאחר שלאיברים של $V \otimes W$ אין ייצוג יחיד כסכום סופי של איברים מהצורה $v \otimes w$.

9.82 הגדרה: מכפלה פנימית מעל מכפלה טנזורית של שני מרחבי מכפלה פנימית

נניח ש־ V ו־ W מרחבי מכפלה פנימית. המכפלה הפנימית מעל $V \otimes W$ היא הפונקציה היחידה $\langle \cdot, \cdot \rangle$ מ־ $(V \otimes W) \times (V \otimes W)$ אל $\mathbf{F}$ כך ש־

$$\langle v \otimes w, u \otimes x \rangle = \langle v, u \rangle \langle w, x \rangle$$

לכל $v, u \in V$ ו־ $w, x \in W$.

ניקח $u = v$ ו־ $x = w$ במשוואה לעיל ואז נוציא שורשים ריבועיים להראות כי

$$\|v \otimes w\| = \|v\| \, \|w\|$$

לכל $v \in V$ ולכל $w \in W$.

מבנה המכפלה הפנימית בהוכחת 9.80 תלוי בבסיס אורתונורמלי $e_1, \dots, e_m$ של V ובבסיס אורתונורמלי $f_1, \dots, f_n$ של W. נוסחה 9.81 גוררת ש־ $\{e_j \otimes f_k\}_{j=1,\dots,m;k=1,\dots,n}$ היא רשימה אורתונורמלית בעלת אינדקס כפול ב־ $V \otimes W$ ולכן היא בסיס אורתונורמלי של $V \otimes W$ [לפי 9.74(ב)]. חשיבות המסקנה הבאה מתעוררת מאחר שהבסיסים האורתונורמליים שבשימוש יכולים להיות שונים מהבסיסים האורתונורמליים שבשימוש במכפלה הפנימית ב־ 9.80. למרות הסימון הזהה עבור הבסיסים בהוכחת 9.80 ובמסקנה להלן, חשבו עליהם כשתי קבוצות שונות של בסיסים אורתונורמליים.

9.83 בסיס אורתונורמלי של $V \otimes W$

נניח ש־ V ו־ W מרחבי מכפלה פנימית, ו־ $e_1, \dots, e_m$ בסיס אורתונורמלי של V ו־ $f_1, \dots, f_n$ בסיס אורתונורמלי של W. אזי

$$\{e_j \otimes f_k\}_{j=1,\dots,m;k=1,\dots,n}$$

בסיס אורתונורמלי של $V \otimes W$.

הוכחה ידוע לנו כי $\{e_j \otimes f_k\}_{j=1,\dots,m;k=1,\dots,n}$ הוא בסיס של $V \otimes W$ [לפי 9.74(ב)]. לפיכך כל שעלינו לוודא הוא האורתונורמליות. לצורך כך, נניח ש־ $j, M \in \{1, \dots, m\}$ ו־ $k, N \in \{1, \dots, n\}$. אזי

$$\langle e_j \otimes f_k, e_M \otimes f_N \rangle = \langle e_j, e_M \rangle \langle f_k, f_N \rangle = \begin{cases} 1 & \text{if } j = M \text{ and } k = N, \\ 0 & \text{otherwise.} \end{cases}$$

לכן הרשימה בעלת האינדקסים הכפולים $\{e_j \otimes f_k\}_{j=1,\dots,m;k=1,\dots,n}$ היא בהחלט בסיס אורתונורמלי של $V \otimes W$. ■

ראו תרגיל 11 לדוגמה לאופן שבו מבנה המכפלה הפנימית מעל $V \otimes W$ משתלב עם אופרטורים מעל V ו־ W.

Tensor Product of Multiple Vector Spaces

We have been discussing properties of the tensor product of two finite-dimensional vector spaces. Now we turn our attention to the tensor product of multiple finite-dimensional vector spaces. This generalization requires no new ideas, only some slightly more complicated notation. Readers with a good understanding of the tensor product of two vector spaces should be able to make the extension to the tensor product of more than two vector spaces.

Thus in this subsection, no proofs will be provided. The definitions and the statements of results that will be provided should be enough information to enable readers to fill in the details, using what has already been learned about the tensor product of two vector spaces.

We begin with the following notational assumption.

9.84 notation: $V_1, \dots, V_m$

For the rest of this subsection, m denotes an integer greater than 1 and $V_1, \dots, V_m$ denote finite-dimensional vector spaces.

The notion of an m-linear functional, which we are about to define, generalizes the notion of a bilinear functional (see 9.68). Recall that the use of the word "functional" indicates that we are mapping into the scalar field $\mathbf{F}$. Recall also that the terminology "m-linear form" is used in the special case $V_1 = \cdots = V_m$ (see 9.25). The notation $\mathcal{B}(V_1, \dots, V_m)$ generalizes our previous notation $\mathcal{B}(V, W)$.

9.85 definition: *m-linear functional, the vector space $\mathcal{B}(V_1, \dots, V_m)$*

- An *m-linear functional* on $V_1 \times \cdots \times V_m$ is a function $\beta \colon V_1 \times \cdots \times V_m \to \mathbf{F}$ that is a linear functional in each slot when the other slots are held fixed.
- The vector space of m-linear functionals on $V_1 \times \cdots \times V_m$ is denoted by $\mathcal{B}(V_1, \dots, V_m)$.

9.86 example: *m-linear functional*

Suppose $\varphi_k \in V_k'$ for each $k \in \{1, \dots, m\}$. Define $\beta \colon V_1 \times \cdots \times V_m \to \mathbf{F}$ by

$$\beta(v_1, \dots, v_m) = \varphi_1(v_1) \times \cdots \times \varphi_m(v_m).$$

Then β is an m-linear functional on $V_1 \times \cdots \times V_m$.

The next result can be proved by imitating the proof of 9.70.

9.87 *dimension of the vector space of m-linear functionals*

$$\dim \mathcal{B}(V_1, \dots, V_m) = (\dim V_1) \cdots (\dim V_m).$$

מכפלה טנזורית של מספר מרחבים וקטוריים

עד כה עסקנו בתכונות המכפלה הטנזורית של שני מרחבים וקטוריים מממד סופי. כעת נפנה את תשומת ליבנו למכפלה טנזורית של מספר מרחבים וקטוריים מממד סופי. הכללה זו אינה דורשת רעיונות חדשים, מלבד סימון מעט יותר מסובך. קוראים בעלי הבנה טובה של המכפלה הטנזורית של שני מרחבים וקטוריים יוכלו לבצע את ההרחבה למכפלה טנזורית של יותר משני מרחבים וקטוריים מממד סופי.

לפיכך בתת־יחידה זו, לא יסופקו הוכחות. ההגדרות והניסוחים של המסקנות שיסופקו להלן אמורים להיות מספקים כדי לאפשר לקוראים להשלים את הפרטים, תוך שימוש במה שכבר נלמד בנוגע למכפלה טנזורית של שני מרחבים וקטוריים.

אנו נפתח עם הנחות הסימון הבאות.

9.84 סימון: $V_1, \dots, V_m$

עד תום יחידה זו, m מסמן מספר שלם הגדול מ־ 1 ו־ $V_1, \dots, V_m$ מסמנים מרחבים וקטוריים מממד סופי.

הרעיון של פונקציונל m־ליניארי, שאותו אנו עומדים להגדיר, מכליל את רעיון הפונקציונל הבילינארי (ראו 9.68). ניזכר כי השימוש במילה "פונקציונל" הוא חיווי לכך שאנו מעתיקים אל שדה הסקלרים $\mathbf{F}$. נזכיר בנוסף כי הטרמינולוגיה "תבנית m־ליניארית" גם היא בשימוש במקרה הפרטי $V_1 = \cdots = V_m$ (ראו 9.25). הסימון $\mathcal{B}(V_1, \dots, V_m)$ הוא הכללה של הסימון $\mathcal{B}(V, W)$ לעיל.

9.85 הגדרה: פונקציונל m־ליניארי, המרחב הווקטורי $\mathcal{B}(V_1, \dots, V_m)$

- פונקציונל m־ליניארי מעל $V_1 \times \cdots \times V_m$ הוא פונקציה $\beta \colon V_1 \times \cdots \times V_m \to \mathbf{F}$ שהיא פונקציונל ליניארי בכל רכיב כאשר הרכיבים האחרים מקובעים.
- המרחב הווקטורי של פונקציונלים m־ליניאריים מעל $V_1 \times \cdots \times V_m$ מסומן על ידי $\mathcal{B}(V_1, \dots, V_m)$.

9.86 דוגמה: פונקציונל m־ליניארי

נניח ש־ $\varphi_k \in V_k'$ לכל $k \in \{1, \dots, m\}$. נגדיר $\beta \colon V_1 \times \cdots \times V_m \to \mathbf{F}$ על ידי

$$\beta(v_1, \dots, v_m) = \varphi_1(v_1) \times \cdots \times \varphi_m(v_m).$$

אזי β פונקציונל m־ליניארי מעל $V_1 \times \cdots \times V_m$.

את המסקנה הבאה ניתן להוכיח באמצעות חיקוי של הוכחת 9.70.

9.87 ממד המרחב הווקטורי של פונקציונלים m־ליניאריים

$$\dim \mathcal{B}(V_1, \dots, V_m) = (\dim V_1) \cdots (\dim V_m).$$

Now we can define the tensor product of multiple vector spaces and the tensor product of elements of those vector spaces. The following definition is completely analogous to our previous definition (9.71) in the case $m = 2$.

9.88 definition: *tensor product*, $V_1 \otimes \cdots \otimes V_m$, $v_1 \otimes \cdots \otimes v_m$

- The *tensor product* $V_1 \otimes \cdots \otimes V_m$ is defined to be $\mathcal{B}(V_1', \ldots, V_m')$.
- For $v_1 \in V_1, \ldots, v_m \in V_m$, the *tensor product* $v_1 \otimes \cdots \otimes v_m$ is the element of $V_1 \otimes \cdots \otimes V_m$ defined by

$$(v_1 \otimes \cdots \otimes v_m)(\varphi_1, \ldots, \varphi_m) = \varphi_1(v_1) \cdots \varphi_m(v_m)$$

for all $(\varphi_1, \ldots, \varphi_m) \in V_1' \times \cdots \times V_m'$.

The next result can be proved by following the pattern of the proof of the analogous result when $m = 2$ (see 9.72).

9.89 *dimension of the tensor product*

$$\dim(V_1 \otimes \cdots \otimes V_m) = (\dim V_1) \cdots (\dim V_m).$$

Our next result generalizes 9.74.

9.90 *basis of* $V_1 \otimes \cdots \otimes V_m$

Suppose $\dim V_k = n_k$ and $e_1^k, \ldots, e_{n_k}^k$ is a basis of V_k for $k = 1, \ldots, m$. Then

$$\{e_{j_1}^1 \otimes \cdots \otimes e_{j_m}^m\}_{j_1=1,\ldots,n_1;\cdots;j_m=1,\ldots,n_m}$$

is a basis of $V_1 \otimes \cdots \otimes V_m$.

Suppose $m = 2$ and $e_1^1, \ldots, e_{n_1}^1$ is a basis of V_1 and $e_1^2, \ldots, e_{n_2}^2$ is a basis of V_2. Then with respect to the basis $\{e_{j_1}^1 \otimes e_{j_2}^2\}_{j_1=1,\ldots,n_1;j_2=1,\ldots,n_2}$ in the result above, the coefficients of an element of $V_1 \otimes V_2$ can be represented by an n_1-by-n_2 matrix that contains the coefficient of $e_{j_1}^1 \otimes e_{j_2}^2$ in row j_1, column j_2. Thus we need a matrix, which is an array specified by two indices, to represent an element of $V_1 \otimes V_2$.

If $m > 2$, then the result above shows that we need an array specified by m indices to represent an arbitrary element of $V_1 \otimes \cdots \otimes V_m$. Thus tensor products may appear when we deal with objects specified by arrays with multiple indices.

The next definition generalizes the notion of a bilinear map (see 9.77). As with bilinear maps, the target space can be an arbitrary vector space.

9.91 definition: *m-linear map*

An *m-linear map* from $V_1 \times \cdots \times V_m$ to a vector space U is a function $\Gamma : V_1 \times \cdots \times V_m \to U$ that is a linear map in each slot when the other slots are held fixed.

כעת באפשרותנו להגדיר את המכפלה הטנזורית של מספר מרחבים וקטוריים ואת המכפלה הטנזורית של איברי מרחבים וקטוריים אלו. ההגדרה הבאה היא אנלוגית לחלוטין להגדרה (9.71) עבור המקרה של $m = 2$.

9.88 הגדרה: מכפלה טנזורית, $V_1 \otimes \cdots \otimes V_m$, $v_1 \otimes \cdots \otimes v_m$

- המכפלה הטנזורית $V_1 \otimes \cdots \otimes V_m$ מוגדרת כ־ $\mathcal{B}(V_1', \ldots, V_m')$.
- עבור $v_1 \in V_1, \ldots, v_m \in V_m$, המכפלה הטנזורית $v_1 \otimes \cdots \otimes v_m$ היא האיבר של $V_1 \otimes \cdots \otimes V_m$ המוגדר על ידי
$$(v_1 \otimes \cdots \otimes v_m)(\varphi_1, \ldots, \varphi_m) = \varphi_1(v_1) \cdots \varphi_m(v_m)$$
לכל $(\varphi_1, \ldots, \varphi_m) \in V_1' \times \cdots \times V_m'$.

את המסקנה הבאה ניתן להוכיח באמצעות התחקות אחרי ההוכחה שניתנה במסקנה האנלוגית כאשר $m = 2$ (ראו 9.72).

9.89 ממד המכפלה הטנזורית

$$\dim(V_1 \otimes \cdots \otimes V_m) = (\dim V_1) \cdots (\dim V_m).$$

המסקנה הבאה מכלילה את 9.74.

9.90 בסיס של $V_1 \otimes \cdots \otimes V_m$

נניח ש־ $\dim V_k = n_k$ ו־ $e_1^k, \ldots, e_{n_k}^k$ בסיס של V_k עבור $k = 1, \ldots, m$. אזי
$$\{e_{j_1}^1 \otimes \cdots \otimes e_{j_m}^m\}_{j_1=1,\ldots,n_1;\cdots;j_m=1,\ldots,n_m}$$
בסיס של $V_1 \otimes \cdots \otimes V_m$.

נניח ש־ $m = 2$ ו־ $e_1^1, \ldots, e_{n_1}^1$ בסיס של V_1 ו־ $e_1^2, \ldots, e_{n_2}^2$ בסיס של V_2. אזי ביחס לבסיס $\{e_{j_1}^1 \otimes e_{j_2}^2\}_{j_1=1,\ldots,n_1;j_2=1,\ldots,n_2}$ במסקנה לעיל, מקדמי האיברים של $V_1 \otimes V_2$ ניתנים לייצוג באמצעות מטריצה $n_1 \times n_2$ המכילה את המקדם של $e_{j_1}^1 \otimes e_{j_2}^2$ בשורה j_1, עמודה j_2. לפיכך נצטרך מטריצה, שהיא מערך המיוצג באמצעות שני אינדקסים, כדי לייצג איבר של $V_1 \otimes V_2$.

אם $m > 2$, אזי המסקנה לעיל מראה כי אנו צריכים מערך המוגדר באמצעות m אינדקסים כדי לייצג איבר שרירותי של $V_1 \otimes \cdots \otimes V_m$. לפיכך ייתכן שמכפלה טנזורית תופיע כאשר נדון באובייקטים המוגדרים על ידי מערכים בעלי אינדקסים מרובים.

ההגדרה הבאה מכלילה את רעיון ההעתקה הבילינארית (ראו 9.77). בדומה להעתקות הבילינאריות, מרחב היעד יכול להיות מרחב וקטורי שרירותי.

9.91 הגדרה: העתקה m־ליניארית

העתקה m־ליניארית מ־ $V_1 \times \cdots \times V_m$ למרחב וקטורי U היא פונקציה $\Gamma : V_1 \times \cdots \times V_m \to U$ שהיא העתקה ליניארית בכל רכיב כאשר כל הרכיבים האחרים מקובעים.

The next result can be proved by following the pattern of the proof of 9.79.

9.92 *converting m-linear maps to linear maps*

Suppose U is a vector space.

(a) Suppose that $\Gamma : V_1 \times \cdots \times V_m \to U$ is an m-linear map. Then there exists a unique linear map $\hat{\Gamma} : V_1 \otimes \cdots \otimes V_m \to U$ such that

$$\hat{\Gamma}(v_1 \otimes \cdots \otimes v_m) = \Gamma(v_1, \ldots, v_m)$$

for all $(v_1, \ldots, v_m) \in V_1 \times \cdots \times V_m$.

(b) Conversely, suppose $T : V_1 \otimes \cdots \otimes V_m \to U$ is a linear map. Then there exists a unique m-linear map $T^{\#} : V_1 \times \cdots \times V_m \to U$ such that

$$T^{\#}(v_1, \ldots, v_m) = T(v_1 \otimes \cdots \otimes v_m)$$

for all $(v_1, \ldots, v_m) \in V_1 \times \cdots \times V_m$.

See Exercises 12 and 13 for tensor products of multiple inner product spaces.

Exercises 9D

1 Suppose $v \in V$ and $w \in W$. Prove that $v \otimes w = 0$ if and only if $v = 0$ or $w = 0$.

2 Give an example of six distinct vectors $v_1, v_2, v_3, w_1, w_2, w_3$ in $\mathbf{R}^3$ such that

$$v_1 \otimes w_1 + v_2 \otimes w_2 + v_3 \otimes w_3 = 0$$

but none of $v_1 \otimes w_1, v_2 \otimes w_2, v_3 \otimes w_3$ is a scalar multiple of another element of this list.

3 Suppose that $v_1, \ldots, v_m$ is a linearly independent list in V. Suppose also that $w_1, \ldots, w_m$ is a list in W such that

$$v_1 \otimes w_1 + \cdots + v_m \otimes w_m = 0.$$

Prove that $w_1 = \cdots = w_m = 0$.

4 Suppose $\dim V > 1$ and $\dim W > 1$. Prove that

$$\{v \otimes w : (v, w) \in V \times W\}$$

is not a subspace of $V \otimes W$.

This exercise implies that if $\dim V > 1$ *and* $\dim W > 1$, *then*

$$\{v \otimes w : (v, w) \in V \times W\} \neq V \otimes W.$$

את המסקנה הבאה ניתן להוכיח בכך שנתחקה אחרי תבנית ההוכחה של 9.79.

9.92 המרת העתקות m־ליניאריות להעתקות ליניאריות

נניח ש־ U מרחב וקטורי.

א. נניח ש־ $\Gamma : V_1 \times \cdots \times V_m \to U$ העתקה m־ליניארית. אזי קיימת העתקה ליניארית יחידה $\hat{\Gamma} : V_1 \otimes \cdots \otimes V_m \to U$ כך ש־

$$\hat{\Gamma}(v_1 \otimes \cdots \otimes v_m) = \Gamma(v_1, \dots, v_m)$$

לכל $(v_1, \dots, v_m) \in V_1 \times \cdots \times V_m$.

ב. בכיוון ההפוך, נניח ש־ $T : V_1 \otimes \cdots \otimes V_m \to U$ העתקה ליניארית. אזי קיימת העתקה m־ליניארית יחידה $T^\# : V_1 \times \cdots \times V_m \to U$ כך ש־

$$T^\#(v_1, \dots, v_m) = T(v_1 \otimes \cdots \otimes v_m)$$

לכל $(v_1, \dots, v_m) \in V_1 \times \cdots \times V_m$.

ראו תרגילים 12 ו־ 13 עבור מכפלות טנזוריות של מרחבי מכפלה פנימית.

תרגילים 9ט

1 נניח ש־ $v \in V$ ו־ $w \in W$. הוכיחו כי $v \otimes w = 0$ אם־ורק־אם $v = 0$ או $w = 0$.

2 הביאו דוגמה של שישה וקטורים שונים $v_1, v_2, v_3, w_1, w_2, w_3$ ב־ $\mathbf{R}^3$ כך ש־

$$v_1 \otimes w_1 + v_2 \otimes w_2 + v_3 \otimes w_3 = 0$$

אך אף אחד מ־ $v_1 \otimes w_1$, $v_2 \otimes w_2$, $v_3 \otimes w_3$ אינו כפולת סקלר של איבר אחר ברשימה זו.

3 נניח ש־ $v_1, \dots, v_m$ רשימה בלתי תלויה ליניארית ב־ V. נניח בנוסף כי $w_1, \dots, w_m$ רשימה ב־ W כך ש־

$$v_1 \otimes w_1 + \cdots + v_m \otimes w_m = 0.$$

הוכיחו כי $w_1 = \cdots = w_m = 0$.

4 נניח ש־ $\dim V > 1$ ו־ $\dim W > 1$. הוכיחו כי

$$\{v \otimes w : (v, w) \in V \times W\}$$

אינו תת־מרחב של $V \otimes W$.

תרגיל זה גורר שאם $\dim V > 1$ ו־ $\dim W > 1$, אזי

$$\{v \otimes w : (v, w) \in V \times W\} \neq V \otimes W.$$

5 Suppose m and n are positive integers. For $v \in \mathbf{F}^m$ and $w \in \mathbf{F}^n$, identify $v \otimes w$ with an m-by-n matrix as in Example 9.76. With that identification, show that the set

$$\{v \otimes w : v \in \mathbf{F}^m \text{ and } w \in \mathbf{F}^n\}$$

is the set of m-by-n matrices (with entries in $\mathbf{F}$) that have rank at most one.

6 Suppose m and n are positive integers. Give a description, analogous to Exercise 5, of the set of m-by-n matrices (with entries in $\mathbf{F}$) that have rank at most two.

7 Suppose $\dim V > 2$ and $\dim W > 2$. Prove that

$$\{v_1 \otimes w_1 + v_2 \otimes w_2 : v_1, v_2 \in V \text{ and } w_1, w_2 \in W\} \neq V \otimes W.$$

8 Suppose $v_1, \dots, v_m \in V$ and $w_1, \dots, w_m \in W$ are such that

$$v_1 \otimes w_1 + \cdots + v_m \otimes w_m = 0.$$

Suppose that U is a vector space and $\Gamma : V \times W \to U$ is a bilinear map. Show that

$$\Gamma(v_1, w_1) + \cdots + \Gamma(v_m, w_m) = 0.$$

9 Suppose $S \in \mathcal{L}(V)$ and $T \in \mathcal{L}(W)$. Prove that there exists a unique operator on $V \otimes W$ that takes $v \otimes w$ to $Sv \otimes Tw$ for all $v \in V$ and $w \in W$.

In an abuse of notation, the operator on $V \otimes W$ given by this exercise is often called $S \otimes T$.

10 Suppose $S \in \mathcal{L}(V)$ and $T \in \mathcal{L}(W)$. Prove that $S \otimes T$ is an invertible operator on $V \otimes W$ if and only if both S and T are invertible operators. Also, prove that if both S and T are invertible operators, then $(S \otimes T)^{-1} = S^{-1} \otimes T^{-1}$, where we are using the notation from the comment after Exercise 9.

11 Suppose V and W are inner product spaces. Prove that if $S \in \mathcal{L}(V)$ and $T \in \mathcal{L}(W)$, then $(S \otimes T)^* = S^* \otimes T^*$, where we are using the notation from the comment after Exercise 9.

12 Suppose that $V_1, \dots, V_m$ are finite-dimensional inner product spaces. Prove that there is a unique inner product on $V_1 \otimes \cdots \otimes V_m$ such that

$$\langle v_1 \otimes \cdots \otimes v_m, u_1 \otimes \cdots \otimes u_m \rangle = \langle v_1, u_1 \rangle \cdots \langle v_m, u_m \rangle$$

for all $(v_1, \dots, v_m)$ and $(u_1, \dots, u_m)$ in $V_1 \times \cdots \times V_m$.

Note that the equation above implies that

$$\|v_1 \otimes \cdots \otimes v_m\| = \|v_1\| \times \cdots \times \|v_m\|$$

for all $(v_1, \dots, v_m) \in V_1 \times \cdots \times V_m$.

5 נניח ש־ m ו־ n מספרים שלמים חיוביים. עבור $v \in \mathbf{F}^m$ ו־ $w \in \mathbf{F}^n$, זהו את $v \otimes w$ עם מטריצה $m \times n$ כמו בדוגמה 9.76. עם זיהוי זה, הראו כי הקבוצה

$$\{v \otimes w : v \in \mathbf{F}^m \text{ and } w \in \mathbf{F}^n\}$$

היא קבוצת המטריצות $m \times n$ (עם איברים ב־ $\mathbf{F}$) שדרגתן היא לכל היותר 1.

6 נניח ש־ m ו־ n מספרים שלמים חיוביים. תנו תיאור אנלוגי לתרגיל 5, של קבוצת המטריצות $m \times n$ (עם איברים ב־ $\mathbf{F}$) שדרגתן לכל היותר 2.

7 נניח ש־ $\dim V > 2$ ו־ $\dim W > 2$. הוכיחו כי

$$\{v_1 \otimes w_1 + v_2 \otimes w_2 : v_1, v_2 \in V \text{ and } w_1, w_2 \in W\} \neq V \otimes W.$$

8 נניח ש־ $v_1, \dots, v_m \in V$ ו־ $w_1, \dots, w_m \in W$ הם כך ש־

$$v_1 \otimes w_1 + \cdots + v_m \otimes w_m = 0.$$

נניח ש־ U מרחב וקטורי ו־ $\Gamma : V \times W \to U$ העתקה בילינארית. הראו כי

$$\Gamma(v_1, w_1) + \cdots + \Gamma(v_m, w_m) = 0.$$

9 נניח ש־ $S \in \mathcal{L}(V)$ ו־ $T \in \mathcal{L}(W)$. הוכיחו כי קיים אופרטור יחיד מעל $V \otimes W$ המעתיק $v \otimes w$ אל $Sv \otimes Tw$ לכל $v \in V$ ו־ $w \in W$.

בשימוש חופשי בסימון, האופרטור מעל $V \otimes W$ המוצג בתרגיל זה לעיתים קרובות נקרא $S \otimes T$.

10 נניח ש־ $S \in \mathcal{L}(V)$ ו־ $T \in \mathcal{L}(W)$. הוכיחו כי $S \otimes T$ הוא אופרטור הפיך מעל $V \otimes W$ אם־ורק־אם S ו־ T שניהם אופרטורים הפיכים. בנוסף, הוכיחו כי אם S ו־ T שניהם אופרטורים הפיכים, אזי $(S \otimes T)^{-1} = S^{-1} \otimes T^{-1}$, כאשר אנו משתמשים בסימון מההערה שאחרי תרגיל 9.

11 נניח ש־ V ו־ W שניהם מרחבי מכפלה פנימית. הוכיחו כי אם $S \in \mathcal{L}(V)$ ו־ $T \in \mathcal{L}(W)$, אזי $(S \otimes T)^* = S^* \otimes T^*$, כאשר אנו משתמשים בסימון מההערה שאחרי תרגיל 9.

12 נניח ש־ $V_1, \dots, V_m$ מרחבי מכפלה פנימית מממד סופי. הוכיחו כי קיימת מכפלה פנימית יחידה מעל $V_1 \otimes \cdots \otimes V_m$ כך ש־

$$\langle v_1 \otimes \cdots \otimes v_m, u_1 \otimes \cdots \otimes u_m \rangle = \langle v_1, u_1 \rangle \cdots \langle v_m, u_m \rangle$$

לכל $(v_1, \dots, v_m)$ ו־ $(u_1, \dots, u_m)$ ב־ $V_1 \times \cdots \times V_m$.

שימו לב שהמשוואה לעיל גוררת ש־

$$\|v_1 \otimes \cdots \otimes v_m\| = \|v_1\| \times \cdots \times \|v_m\|$$

לכל $(v_1, \dots, v_m) \in V_1 \times \cdots \times V_m$.

13 Suppose that $V_1, \dots, V_m$ are finite-dimensional inner product spaces and $V_1 \otimes \cdots \otimes V_m$ is made into an inner product space using the inner product from Exercise 12. Suppose $e_1^k, \dots, e_{n_k}^k$ is an orthonormal basis of V_k for each $k = 1, \dots, m$. Show that the list

$$\{e_{j_1}^1 \otimes \cdots \otimes e_{j_m}^m\}_{j_1=1,\dots,n_1;\cdots;j_m=1,\dots,n_m}$$

is an orthonormal basis of $V_1 \otimes \cdots \otimes V_m$.

13 נניח ש־ $V_1, \dots, V_m$ מרחבי מכפלה פנימית מממד סופי ו־ $V_1 \otimes \cdots \otimes V_m$ הפך למרחב מכפלה פנימית בשימוש במכפלה הפנימית מתרגיל 12. נניח ש־ $e_1^k, \dots, e_{n_k}^k$ בסיס אורתונורמלי של V_k לכל $k = 1, \dots, m$. הראו כי הרשימה

$$\{e_{j_1}^1 \otimes \cdots \otimes e_{j_m}^m\}_{j_1=1,\dots,n_1;\cdots;j_m=1,\dots,n_m}$$

היא בסיס אורתונורמלי של $V_1 \otimes \cdots \otimes V_m$.

Photo Credits

- page v: Photos by Carrie Heeter and Bishnu Sarangi. Public domain image.
- page 1: Original painting by Pierre Louis Dumesnil; 1884 copy by Nils Forsberg. Public domain image downloaded on 29 March 2022 from https://commons.wikimedia.org/wiki/File:René_Descartes_i_samtal_med_Sveriges_drottning,_Kristina.jpg.
- page 27: Public domain image downloaded on 4 February 2022 from https://commons.wikimedia.org/wiki/File:IAS_Princeton.jpg.
- page 51: Photo by Stefan Schäfer; Creative Commons Attribution Share Alike license. Downloaded on 28 March 2022 from https://commons.wikimedia.org/wiki/File:BurgDankwarderode2016.jpg.
- page 119: Photo by Alireza Javaheri. Creative Commons Attribution license. Downloaded on 12 March 2023 from https://commons.wikimedia.org/wiki/File:Hakim_Omar_Khayam_-_panoramio.jpg.
- page 132: Statue completed by Giovanni Paganucci in 1863. Photo by Hans-Peter Postel; Creative Commons Attribution license. Downloaded on 14 March 2022 from https://commons.wikimedia.org/wiki/File:Leonardo_da_Pisa.jpg.
- page 181: Photo by Matthew Petroff; Creative Commons Attribution Share Alike license. Downloaded on 31 March 2022 from https://commons.wikimedia.org/wiki/File:George-peabody-library.jpg.
- page 227: Photo by Petar Milošević; Creative Commons Attribution Share Alike license. Downloaded on 30 March 2022 from https://en.wikipedia.org/wiki/Lviv.
- page 297: Photo by David Iliff; Creative Commons Attribution Share Alike license. Downloaded on 30 March 2022 from https://en.wikipedia.org/wiki/File:Long_Room_Interior,_Trinity_College_Dublin,_Ireland_-_Diliff.jpg.
- page 332: Photo by Daniel Schwen; Creative Commons Attribution Share Alike license. Downloaded on 9 July 2019 from https://commons.wikimedia.org/wiki/File:Mathematik_Göttingen.jpg.

Symbol Index

A^{-1}, 91
$A_{j,\cdot}$, 74
$A_{j,k}$, 69
$A_{\cdot,k}$, 74
α_T, 354
A^*, 231
A^t, 77

$\hat{\Gamma}$, 375, 380
B, 287
$\mathcal{B}(V_1, \ldots, V_m)$, 378
$\mathcal{B}(V, W)$, 370

$\mathbf{C}$, 2
$\circ$, 55

deg, 31
Δ, 196
$\det A$, 355
$\det T$, 354
dim, 44
$\oplus$, 21

$E(s_1 f_1, \ldots, s_n f_n)$, 287
$E(\lambda, T)$, 164

$\mathbf{F}$, 4
$\mathbf{F}^\infty$, 13
$\mathbf{F}^{m,n}$, 72
$\mathbf{F}^n$, 6
$\mathbf{F}^S$, 13

$G(\lambda, T)$, 308

I, 52, 90

$\iff$, 23
Im, 120
$-\infty$, 31

$\mathcal{L}(V)$, 52
$\mathcal{L}(V, W)$, 52

$\mathcal{M}(\beta)$, 334
$\mathcal{M}(T)$, 69, 154
$\mathcal{M}(v)$, 88

perm, 348
$\mathcal{P}(\mathbf{F})$, 30
π, 101
$\mathcal{P}_m(\mathbf{F})$, 31
$p(T)$, 137
P_U, 214

q_β, 341
$\blacksquare$, 7

$\mathbf{R}$, 2
Re, 120

$S \otimes T$, 381
$\subsetneq$, 299

$\sqrt{T}$, 253
$\widetilde{T}$, 102
T', 107
T^*, 228
T^{-1}, 82
$T(\Omega)$, 288
$T_{\mathbf{C}}$, 68
$T^\dagger$, 221
T^m, 137
$\|T\|$, 280
$T^\#$, 375, 380
$\operatorname{tr} A$, 326
$\operatorname{tr} T$, 327
$T|_U$, 133
T/U, 142

$U^\perp$, 211
U^0, 109
$\langle u, v \rangle$, 184

V, 15
V', 105, 204
V/U, 99
$-v$, 15
$V_1 \otimes \cdots \otimes V_m$, 379
$v_1 \otimes \cdots \otimes v_m$, 379
$V^{(2)}$, 334
$V^{(2)}_{\text{alt}}$, 339
$V^{(2)}_{\text{sym}}$, 337
$V_{\mathbf{C}}$, 17
V^m, 103, 346
$V^{(m)}$, 346
$V^{(m)}_{\text{alt}}$, 347
$V \otimes W$, 372
$v \otimes w$, 372
$v + U$, 98

$\|v\|$, 186

$\overline{z}$, 120
$|z|$, 120

Index

Abbott, Edwin A., 6
absolute value, 120
addition
    in quotient space, 100
    of complex numbers, 2
    of functions, 13
    of linear maps, 55
    of matrices, 71
    of subspaces, 19
    of vectors, 12
    of vectors in $\mathbf{F}^n$, 6
additive inverse
    in $\mathbf{C}$, 3, 4
    in $\mathbf{F}^n$, 9
    in vector space, 12, 15
additivity, 52
adjoint of a linear map, 228
algebraic multiplicity, 311
alternating bilinear form, 339
alternating m-linear form, 347
annihilator, 109
Apollonius's identity, 195
Artin, Emil, 80
associativity, 3, 12, 56

backward shift, 53, 59, 84, 140
ball, 287
Banach, Stefan, 227
basis, 39
    of eigenvectors, 165, 245, 246, 250
    of generalized eigenvectors, 301
Bernstein polynomials, 49
Bessel's inequality, 198
bilinear form, 333
bilinear functional, 370
bilinear map, 374
block diagonal matrix, 314
Bunyakovsky, Viktor, 189

C^*-algebras, 295
Carroll, Lewis, 11
Cauchy, Augustin-Louis, 189
Cauchy–Schwarz inequality, 189
Cayley, Arthur, 312
Cayley–Hamilton theorem, 364
    on complex vector space, 312
change-of-basis formula
    for bilinear forms, 336
    for operators, 93
characteristic polynomial, 363
    on complex vector space, 311
ChatGPT, 196, 279
Cholesky factorization, 267
Cholesky, André-Louis, 267
Christina, Queen of Sweden, 1
closed under addition, 18
closed under scalar multiplication, 18
column rank of a matrix, 77, 114, 239
column–row factorization, 78
commutativity, 3, 7, 12, 25, 56, 73, 80
commuting operators, 138, 175–180, 209, 235, 248–249, 256
companion matrix, 152
complex conjugate, 120
complex number, 2
complex spectral theorem, 246
complex vector space, 13
complexification
    eigenvalues of, 140
    generalized eigenvectors of, 318

minimal polynomial of, 153
multiplicity of eigenvalues, 318
of a linear map, 68
of a vector space, 17, 43
of an inner product space, 194
conjugate symmetry, 183
conjugate transpose of a matrix, 231
coordinate, 6
cube root of an operator, 248

De Moivre's theorem, 125
degree of a polynomial, 31
Descartes, René, 1
determinant
of matrix, 355
of operator, 354
of positive operator, 362
of unitary operator, 362
diagonal matrix, 163, 274
diagonal of a square matrix, 155
diagonalizable, 163, 172, 176, 245, 246, 294, 307, 316
differentiation linear map, 53, 56, 59, 61, 62, 67, 70, 79, 138, 208, 304
dimension, 44
of a sum of subspaces, 47
direct sum, 21, 42, 98
of a subspace and its orthogonal complement, 212
of null $T^{\dim V}$ and range $T^{\dim V}$, 299
discrete Fourier transform, 269
distributive property, 3, 12, 15, 56, 80
division algorithm for polynomials, 124
division of complex numbers, 4
dot product, 182
double dual space, 118
dual
of a basis, 106
of a linear map, 107
of a vector space, 105, 204
of an operator, 140, 153, 162, 174

eigenspace, 164
eigenvalue
of adjoint, 239
of dual of an operator, 140
of operator, 134
of positive operator, 252
of self-adjoint operator, 233
of unitary operator, 262
on odd-dimensional space, 150, 318, 367
eigenvector, 135
ellipsoid, 287
Euclidean inner product, 184
Euler, Leonhard, 2

Fibonacci, 132
Fibonacci sequence, 174
field, 10
finite-dimensional vector space, 30
Flatland, 6
forward shift, 140
Fourier series, 207
Frankenstein, 50
Frobenius norm, 331
Fuglede's theorem, 248
fundamental theorem of algebra, 125
fundamental theorem of linear maps, 62

Gauss, Carl Friedrich, 51
Gaussian elimination, 51, 65, 361
generalized eigenspace, 308
generalized eigenvector, 300
geometric multiplicity, 311
Gershgorin disk, 170
Gershgorin disk theorem, 171
Gershgorin, Semyon Aronovich, 171
Gram, Jørgen, 200
Gram–Schmidt procedure, 200
graph of a linear map, 103

Hadamard's inequality, 365
Halmos, Paul, 27

Hamilton, William, 297
harmonic function, 196
Hilbert matrix, 256
Hilbert–Schmidt norm, 331
homogeneity, 52
homogeneous system of linear equations, 65, 95
hyponormal operator, 241

identity matrix, 90
identity operator, 52, 56
imaginary part, 120
infinite-dimensional vector space, 31
injective, 60
inner product, 183
inner product space, 184
Institute for Advanced Study, 27
invariant subspace, 133
inverse
  of a linear map, 82
  of a matrix, 91
invertible linear map, 82
invertible matrix, 91
isometry, 258
isomorphic vector spaces, 86
isomorphism, 86

Jordan basis, 322
Jordan form, 324
Jordan, Camille, 324

kernel, 59
Khayyam, Omar, 119
Klein, Felix, 369

Laplacian, 196
length of list, 5
Leonardo of Pisa, 132
linear combination, 28
linear dependence lemma, 33
linear equations, 64–65, 95
linear functional, 105, 204
linear map, 52
linear map lemma, 54
linear span, 29
linear subspace, 18
linear transformation, 52
linearly dependent, 33
linearly independent, 32
list, 5
  of vectors, 28
lower-triangular matrix, 162, 267
Lviv, 227
Lwów, 227

matrix, 69
  multiplication, 73
  of bilinear form, 334
  of linear map, 69
  of nilpotent operator, 305
  of operator, 154
  of product of linear maps, 74, 91
  of T', 113
  of T^*, 232
  of vector, 88
minimal polynomial
  and basis of generalized eigenvectors, 306
  and characteristic polynomial, 312
  and diagonalizability, 169
  and generalized eigenspace decomposition, 316
  and generalized eigenspaces, 317
  and invertibility, 149
  and upper-triangular matrices, 159, 203
  computing, 145
  definition of, 145
  gcd with its derivative, 173
  no direct sum decomposition, 325
  of adjoint, 241
  of companion matrix, 152
  of complexification, 153
  of dual map, 153
  of nilpotent operator, 305, 324

of normal operator, 241
of quotient operator, 153
of restriction operator, 148
of self-adjoint operator, 244
polynomial multiple of, 148
zeros of, 146
minimizing distance, 217
m-linear form, 346
m-linear functional, 378
m-linear map, 379
monic polynomial, 144
Moon, v, xvii
Moore–Penrose inverse, 221
multilinear form, 346
multiplication, *see* product
multiplicity of an eigenvalue, 310

nilpotent operator, 303, 322
Noether, Emmy, 332
nonsingular matrix, 91
norm of a linear map, 280
norm of a vector, 182, 186
normal operator, 235
null space, 59
of powers of an operator, 298
of T', 111
of T^*, 231

one-to-one, 60
onto, 62
operator, 133
orthogonal
complement, 211
projection, 214
vectors, 187
orthonormal
basis, 199
list, 197

parallelogram equality, 191
Parseval's identity, 200
partial differentiation operator, 175
Peabody Library, 181
permutation, 348
photo credits, 383
point, 12
polar decomposition, 286
polynomial, 30
positive definite, 266
positive operator, 251
positive semidefinite operator, 251
principal axes, 287
product
of complex numbers, 2
of linear maps, 55
of matrices, 73
of polynomials, 138
of scalar and linear map, 55
of scalar and vector, 12
of scalar and vector in $\mathbf{F}^n$, 9
of vector spaces, 96
pseudoinverse, 221, 250, 255, 275, 279
Pythagorean theorem, 187

QR factorization, 264, 365
quadratic form, 341
quotient
map, 101
operator, 142, 153, 162, 173
space, 99

range, 61
of powers of an operator, 306
of T', 112
of T^*, 231
rank of a matrix, 79, 114, 239
real part, 120
real spectral theorem, 245
real vector space, 13
reverse triangle inequality, 129, 193, 294
Riesz representation theorem, 205, 210, 216, 224, 225
Riesz, Frigyes, 205

row rank of a matrix, 77, 114, 239

scalar, 4
scalar multiplication, 9, 12
    in quotient space, 100
    of linear maps, 55
    of matrices, 71
Schmidt pair, 278
Schmidt, Erhard, 200, 278
Schur's theorem, 204
Schur, Issai, 204
Schwarz, Hermann, 189
self-adjoint operator, 233
Shelley, Mary Wollstonecraft, 50
sign of a permutation, 349
simultaneous diagonalization, 176
simultaneously upper triangularizable, 178
singular matrix, 91
singular value decomposition
    of adjoint, 275
    of linear map, 273
    of pseudoinverse, 275
singular values, 271, 362
skew operator, 240, 247, 269
span, 29
spans, 29
spectral theorem, 245, 246
square matrix, 91
square root of an operator, 248, 251, 253, 320
standard basis
    of $\mathbf{F}^n$, 39
    of $\mathcal{P}_m(\mathbf{F})$, 39
subspace, 18
subtraction of complex numbers, 4
sum, *see* addition
sum of subspaces, 19
Supreme Court, 210
surjective, 62
SVD, *see* singular value decomposition
Sylvester, James, 181

symmetric bilinear form, 337
symmetric matrix, 269, 337

Taylor polynomial, 219
tensor product, 372, 379
Through the Looking Glass, 11
trace
    of a matrix, 326
    of an operator, 327
translate, 99
transpose of a matrix, 77, 231
triangle inequality, 121, 190, 281
tuple, 5
two-sided ideal, 58

unit circle in $\mathbf{C}$, 262, 269
unitary matrix, 263
unitary operator, 260
University of Dublin, 297
University of Göttingen, 332
upper-triangular matrix, 155–160, 264, 267, 314

Vandermonde matrix, 366
vector, 8, 12
vector space, 12
volume, 292, 363
    of a box, 291

zero of a polynomial, 122

Colophon: Notes on Typesetting

- This book was typeset in LuaLaTeX by the author, who wrote the LaTeX code to implement the book's design.
- The LaTeX software used for this book was written by Leslie Lamport. The TeX software, which forms the base for LaTeX, was written by Donald Knuth.
- The main text font in this book is the Open Type Format version of TeX Gyre Termes, a font based on Times, which was designed by Stanley Morison and Victor Lardent for the British newspaper *The Times* in 1931.
- The main math font in this book is the Open Type Format version of TeX Gyre Pagella Math, a font based on Palatino, which was designed by Hermann Zapf.
- The sans serif font used for page headings and some other design elements is the Open Type Format version of TeX Gyre Heros, a font based on Helvetica, which was designed by Max Miedinger and Eduard Hoffmann.
- The LuaLaTeX packages fontspec and unicode-math, both written by Will Robertson, were used to manage fonts.
- The LaTeX package fontsize, written by Ivan Valbusa, was used to gracefully change the main fonts to 10.5 point size.
- The figures in the book were produced by *Mathematica*, using *Mathematica* code written by the author. *Mathematica* was created by Stephen Wolfram. The *Mathematica* package MaTeX, written by Szabolcs Horvát, was used to place LaTeX-generated labels in the *Mathematica* figures.
- The LaTeX package graphicx, written by David Carlisle and Sebastian Rahtz, was used to include photos and figures.
- The LaTeX package multicol, written by Frank Mittelbach, was used to get around LaTeX's limitation that two-column format must start on a new page (needed for the Symbol Index and the Index).
- The LaTeX packages TikZ, written by Till Tantau, and tcolorbox, written by Thomas Sturm, were used to produce the definition boxes and result boxes.
- The LaTeX package color, written by David Carlisle, was used to add appropriate color to various design elements.
- The LaTeX package wrapfig, written by Donald Arseneau, was used to wrap text around the comment boxes.
- The LaTeX package microtype, written by Robert Schlicht, was used to reduce hyphenation and produce more pleasing right justification.

אינדקס סמלים וסימנים

A^{-1}, 91
$A_{j,\cdot}$, 74
$A_{j,k}$, 69
$A_{\cdot,k}$, 74
α_T, 354
A^*, 231
A^t, 77

$\hat{\Gamma}$, 375, 380

B, 287
$\mathcal{B}(V_1, \dots, V_m)$, 378
$\mathcal{B}(V, W)$, 370

$\mathbf{C}$, 2
$\circ$, 55

deg, 31
Δ, 196
$\det A$, 355
$\det T$, 354
dim, 44
$\oplus$, 21

$E(s_1 f_1, \dots, s_n f_n)$, 287
$E(\lambda, T)$, 164

$\mathbf{F}$, 4
$\mathbf{F}^\infty$, 13
$\mathbf{F}^{m,n}$, 72
$\mathbf{F}^n$, 6
$\mathbf{F}^S$, 13

$G(\lambda, T)$, 308

I, 52, 90

$\iff$, 23
Im, 120
$-\infty$, 31

$\mathcal{L}(V)$, 52
$\mathcal{L}(V, W)$, 52

$\mathcal{M}(\beta)$, 334
$\mathcal{M}(T)$, 69, 154
$\mathcal{M}(v)$, 88

perm, 348
$\mathcal{P}(\mathbf{F})$, 30
π, 101
$\mathcal{P}_m(\mathbf{F})$, 31
$p(T)$, 137
P_U, 214

q_β, 341

$\mathbf{R}$, 2
Re, 120

$S \otimes T$, 381
$\subsetneq$, 299

$\sqrt{T}$, 253
$\widetilde{T}$, 102
T', 107
T^*, 228
T^{-1}, 82
$T(\Omega)$, 288
$T_{\mathbf{C}}$, 68
$T^\dagger$, 221
T^m, 137

$\|T\|$, 280
$T^\#$, 375, 380
$\operatorname{tr} A$, 326
$\operatorname{tr} T$, 327
$T|_U$, 133
T/U, 142

$U^\perp$, 211
U^0, 109
$\langle u, v \rangle$, 183

V, 15
V', 105, 204
V/U, 99
$-v$, 15
$V_1 \otimes \cdots \otimes V_m$, 379
$v_1 \otimes \cdots \otimes v_m$, 379
$V^{(2)}$, 334
$V^{(2)}_{\text{alt}}$, 339
$V^{(2)}_{\text{sym}}$, 337
$V_{\mathbf{C}}$, 17
V^m, 103
V^m, 346
$V^{(m)}$, 346
$V^{(m)}_{\text{alt}}$, 347
$V \otimes W$, 372
$v \otimes w$, 372
$v + U$, 98

$\|v\|$, 186

$\bar{z}$, 120
$|z|$, 120

אינדקס

n־יה, 5

one-to-one, 60

span, 29
SVD, ראה פירוק לערכים סינגולריים

אִיזוֹמוֹרְפִיזְם, 86
אבוט, אדווין אבוט, 6
אדיטיביות, 52
אוילר, לאונרד, 2
אוניברסיטת גטינגן, 332
אוניברסיטת דבלין, 297
אופרטור, 133
אופרטור אוניטרי, 260
אופרטור אנטי־צמוד, 240, 247, 269
אופרטור הזהות, 52, 56
אופרטור היפו־נורמלי, 241
אופרטור חיובי, 251
אופרטור נגזרת חלקית, 175
אופרטור נורמלי, 235
אופרטור נילפוטנטי, 303, 322
אופרטור צמוד לעצמו, 233
אופרטור קומוטטיבי, 138
אופרטורים מתחלפים, 175–180, 209, 235, 248–249, 256
אורך רשימה, 5
אורתוגונלי
היטל, 214
משלים, 211
אורתוגונליים
וקטורים, 187
אורתונורמלי
בסיס, 199
אורתונורמלית
רשימה, 197
אי־שוויון אָדָמָר, 365
אי־שוויון בֶּסֶל, 198
אי־שוויון המשולש, 121, 190, 281
אי־שוויון המשולש ההפוך, 129, 193, 294
אי־שוויון קושי־שוורץ, 189
איבר נגדי לחיבור
ב־ $\mathbf{C}$, 3, 4
ב־ $\mathbf{F}^n$, 9
במרחב וקטורי, 12, 15
אידיאל דו־צדדי, 58
איזומטריה, 258
אלגברת סי כוכב, 295
אלגוריתם חלוקת פולינומים, 124
אליפסואיד, 287
אלכסון מטריצה ריבועית, 155
אסוציאטיביות, 3, 12, 56
ארטין, אמיל, 80
בּוּנְיָקוֹבְסְקִי, ויקטור, 189
בית המשפט העליון, 210
בלתי תלויה ליניארית, 32
בנך, סטפן, 227
בסיס, 39
של וקטורים עצמיים, 165, 245, 246, 250
של וקטורים עצמיים מוכללים, 301
בסיס זוֹרדן, 322
בסיס סטנדרטי
של $\mathbf{F}^n$, 39
של $\mathcal{P}_m(\mathbf{F})$, 39
גְּרָם, יוֹרְגֶן פֶּדֶרְסֶן, 200
גאוס, קרל פרידריך, 51

אלגברה ליניארית בדרך הישר, מהדורה רביעית, מחבר: שלדון אקסלר
Linear Algebra Done Right, fourth edition, by Sheldon Axler

גרעין, 59
גרף של העתקה ליניארית, 103
גרשגורן, סמיון ארונוביץ', 171
דואלי
  מרחב וקטורי, 105, 204
  של אופרטור, 140, 153, 162, 174
  של בסיס, 106
  של העתקה ליניארית, 107
דטרמיננטה
  של אופרטור, 354
  של אופרטור אוניטרי, 362
  של אופרטור חיובי, 362
  של מטריצה, 355
דקארט, רנה, 1
דרגה של מטריצה, 79
דרגת העמודות של מטריצה, 77, 114
דרגת השורות של מטריצה, 77, 114
דרגת מטריצה, 114, 239
דרגת עמודות של מטריצה, 239
דרגת שורות של מטריצה, 239
הומוגניות, 52
הופכי מור–פנרוז, 221
הופכית
  של העתקה ליניארית, 82
  של מטריצה, 91
הזזה, 99
הלמוש, פול, 27
המילטון, ויליאם, 297
המכון למחקר מתקדם, 27
המשפט היסודי של האלגברה, 125
המשפט היסודי של ההעתקות
  הליניאריות, 62
הסטה לאחור, 53, 59, 84, 140
הסטה לפנים, 140
העתקה
  מנה, 101
העתקה m-ליניארית, 379
העתקה ביליניארית, 374
העתקה ליניארית, 52
העתקה ליניארית הפיכה, 82
העתקת הגזירה, 53, 56, 59, 61, 62, 67, 70, 79, 138, 208, 304
הפולינום האופייני, 363
הצמודה של העתקה ליניארית, 228
הצמודה־משוחלפת של מטריצה, 231
התמרת פורייה בדידה, 269
וקטור, 8, 12
וקטור עצמי, 135
וקטור עצמי מוכלל, 300
ז'ורדן, קאמי, 324
ח'יאם, עומר, 119
חד־חד־ערכית, 60
חיבור
  במרחב מנה, 100
  וקטורים, 12
  וקטורים ב־ $\mathbf{F}^n$, 6
  מטריצות, 71
  מספרים מרוכבים, 2
  פונקציות, 13
  של העתקות ליניאריות, 55
  תתי־מרחב, 19
חיובית לחלוטין, 266
חילוק מרוכבים, 4
חיסור מרוכבים, 4
חלק מדומה, 120
חלק ממשי, 120
טווח, 61
  של T', 112
  של T^*, 231
  של חזקות של אופרטור, 306
טור פורייה, 207
טרנספורמציה ליניארית, 52
כדור, 287
כלל המקבילית, 191
כפל
  וקטור בסקלר, 12
  מספרים מרוכבים, 2
  כפל בסקלר, 9, 12
  במרחב מנה, 100
  של העתקות ליניאריות, 55

כפל סקלר
במטריצה, 71
כריסטינה, מלכת שוודיה, 1
לֶמַת ההעתקה הליניארית, 54
לֶמַת התלות הליניארית, 33
לאונרדו מפיזה, 132
לבוב, 227
לביב, 227
לכסין, 163, 172, 176, 245, 246, 294, 307, 316
לכסינות בו־זמנית, 176
למברג, 227
לפלסיאן, 196
מְאַפֵּס, 109
מבעד למראה ומה אליס מצאה שם, 11
מון, v, xii
מזעור מרחק, 217
מטריצה, 69
כפל, 73
מכפלת העתקות ליניאריות, 91
של אופרטור, 154
של אופרטור נילפוטנטי, 305
של T', 113
של T^*, 232
של העתקה ליניארית, 69
של וקטור, 88
של מכפלת העתקות ליניאריות, 74
של תבנית ביליניארית, 334
מטריצה אוניטרית, 263
מטריצה אי־סינגולרית, 91
מטריצה אלכסונית, 163, 274
מטריצה הפיכה, 91
מטריצה מלווה, 152
מטריצה משוחלפת, 77
מטריצה משולשית עליונה, 155–160, 264, 267, 314
סימולטנית לאופרטור, 178
מטריצה משולשית תחתונה, 162, 267
מטריצה סימטרית, 269, 337
מטריצה סינגולרית, 91
מטריצה ריבועית, 91
מטריצת בלוקים אלכסונית, 314
מטריצת הזהות, 90
מטריצת היחידה, 90
מטריצת הילברט, 256
מטריצת וַנְדֶרמוֹנְד, 366
מירכוב
וקטורים עצמיים מוכללים של, 318
ערך עצמי של, 140
פולינום מינימלי של, 153
ריבוי של ערכים עצמיים, 318
של העתקה ליניארית, 68
של מרחב וקטורי, 17, 43
של מרחב מכפלה פנימית, 194
מכפלה
מרחבים וקטוריים, 96
של העתקות ליניאריות, 55
של מטריצות, 73
של סקלר והעתקות ליניאריות, 55
של סקלר ווקטור ב־ $\mathbf{F}^n$, 9
של פולינומים, 138
מכפלה טנזורית, 372, 379
מכפלה סקלרית, 182
מכפלה פנימית, 183
מכפלה פנימית אוקלידית, 184
ממד, 44
סכום תתי־מרחב, 47
מנה
אופרטור, 142, 153, 162, 173
מספר מרוכב, 2
מעגל היחידה ב־ $\mathbf{C}$, 262, 269
מעלת פולינום, 31
מערכת הומוגנית של משוואות ליניאריות, 65, 95
מרחב
מנה, 99
מרחב דואלי כפול, 118
מרחב האפס, 59
של T', 111
של T^*, 231

של חזקות אופרטור, 298
מרחב וקטורי, 12
מרחב וקטורי מממד אינסופי, 31
מרחב וקטורי מממד סופי, 30
מרחב וקטורי ממשי, 13
מרחב וקטורי מרוכב, 13
מרחב מכפלה פנימית, 184
מרחב עצמי, 164
מרחב עצמי מוכלל, 308
מרחבים וקטוריים אִיזוֹמוֹרְפִיִים, 86
משוואות ליניאריות, 64–65, 95
משוחלפת של מטריצה, 231
משפט אפולוניוס, 195
משפט דה־מואבר, 125
משפט ההצגה של ריס, 205, 210, 216, 224, 225
משפט הפירוק הספקטרלי, 245, 246
משפט הפירוק הספקטרלי הממשי, 245
משפט הפירוק הספקטרלי המרוכב, 246
משפט הפירוק של שור, 204
משפט עיגולי גרשגורן, 171
משפט פיתגורס, 187
משפט קיילי־המילטון, 364
מעל מרחב וקטורי מרוכב, 312
נֶתֶר, אֵמִי, 332
נוסחת שינוי בסיס
לאופרטורים, 93
של תבניות ביליניאריות, 336
נורמה של העתקה ליניארית, 280
נורמה של וקטור, 182, 186
נורמת הילברט־שמידט, 331
נורמת פרובניוס, 331
נפח, 292, 363
של תיבה, 291
נקודה, 12
סגירות תחת חיבור, 18
סגירות תחת כפל בסקלר, 18
סדרת פיבונאצ'י, 174
סורג'קטיבית, 62
סילבסטר, ג'יימס, 181
סימטריית צמוד (הרמיטיות), 183
סימן של תמורה, 349
סכום, ראה חיבור
סכום ישר, 21, 42, 98
של $\operatorname{null} T^{\dim V}$ ו־ $\operatorname{range} T^{\dim V}$, 299
של תת־מרחב והמשלים האורתוגונלי שלו, 212
סכום תתי־מרחב, 19
ספריית פיבודי, 181
סקלר, 4
עָל, 62
עיגולי גרשגורן, 170
עקבה
של אופרטור, 327
של מטריצה, 326
ערך מוחלט, 120
ערך עצמי
במרחב מממד אי־זוגי, 150
מעל מרחב מממד אי־זוגי, 318
מרחב מממד אי־זוגי, 367
של אופרטור, 134
של אופרטור אוניטרי, 262
של אופרטור חיובי, 252
של אופרטור צמוד לעצמו, 233
של דואלי של אופרטור, 140
של הצמוד, 239
ערכים סינגולריים, 271, 362
פְּרַנְקְשְׁטַיְן, 50
פְּסִידוֹ־הופכי, 221
פוּגְלֶדֶה, 248
פולינום, 30
פולינום אופייני
מעל מרחב וקטורי מרוכב, 311
פולינום טיילור, 219
פולינום מינימלי
הגדרה, 145
ובסיס של וקטורים עצמיים מוכללים, 306
והפיכות, 149

והפירוק למרחבים עצמיים מוכללים, 316
ולכסינות, 169
ומטריצה משולשית עליונה, 159, 203
ומרחבים עצמיים מוכללים, 317
ופולינום אופייני, 312
חישוב, 145
כפולת פולינום של, 148
ללא פירוק סכום ישר, 325
מחלק משותף מרבי עם נגזרתו, 173
שורשים של, 146
של אופרטור מנה, 153
של אופרטור מצומצם, 148
של אופרטור נורמלי, 241
של אופרטור נילפוטנטי, 305, 324
של אופרטור צמוד לעצמו, 244
של העתקה דואלית, 153
של הצמוד, 241
של מטריצה מלווה, 152
של מירכוב, 153
פולינום מתוקן, 144
פולינומי ברנשטיין, 49
פונקציה הרמונית, 196
פונקציונל m־ליניארי, 378
פונקציונל ביליניארי, 370
פונקציונל ליניארי, 105, 204
פורסת, 29
פיבונאצ'י, 132
פירוק QR, 264, 365
פירוק לערכים סינגולריים
של העתקה ליניארית, 273
של פסידו־הופכי, 275
של צמודה, 275
פירוק עמודה־שורה, 78
פירוק פולרי, 286
פירוק שולסקי, 267
פסידו־הופכי, 250, 255, 275, 279
פריסה, 29
פריסה ליניארית, 29
צורת ז'ורדן, 324
צירוף ליניארי, 28
צירים ראשיים, 287
צמד שמידט, 278
צמוד מרוכב, 120
קואורדינטה, 6
קומוטטיביות, 3, 7, 12, 25, 56, 73, 80
קושי, אוגוסטן־לואי, 189
קיילי, ארתור, 312
קלוד, 196, 279
קרדיטים לתמונות, 383
קרול, לואיס, 11
ריבוי אלגברי, 311
ריבוי גאומטרי, 311
ריבוי של ערך עצמי, 310
ריס, פרידייש, 205
רשימה, 5
וקטורים, 28
שְׁמִידְט, אֶרְהַרְד, 200
שֶׁלִי, מַרִי ווּלְסְטוֹנקְרַאפְט, 50
שׁוּר, ישי, 204
שדה, 10
שוויון פָּרְסֶבַל, 200
שוורץ, הרמן, 189
שולסקי, אנדרה־לואי, 267
שורש פולינום, 122
שורש ריבועי של אופרטור, 248, 251, 253, 320
שורש שלישי של אופרטור, 248
שטוחלנדיה, 6
שיטת החילוץ של גאוס, 51, 65, 361
שמידט, ארהרד, 278
תבנית m־ליניארית, 346
תבנית m־ליניארית מתחלפת, 347
תבנית ביליניארית, 333
תבנית ביליניארית מתחלפת, 339
תבנית ביליניארית סימטרית, 337
תבנית מולטי־ליניארית, 346
תבנית ריבועית, 341
תהליך גרם־שמידט, 200

תכונת הדיסטריבוטיביות, 3, 12, 15, 56, 80
תלויה ליניארית, 33
תמורה, 348
תת־מרחב, 18
תת־מרחב ליניארי, 18
תת־מרחב שמור, 133

קולופון: הערות בנוגע לסִדָר

- ספר זה סודר ב־LuaLaTeX על ידי המחבר, אשר כתב את קוד ה־LaTeX המממש את עיצוב הספר.
- תוכנת ה־LaTeX ששימשה עבור ספר זה נכתבה על ידי Leslie Lamport. תוכנת ה־TeX, המהווה את הבסיס של LaTeX, נכתבה על ידי Donald Knuth.
- גופן הטקסט הלטיני בספר זה הוא גרסת ה־Open Type Format של TeX Gyre Termes, גופן המבוסס על Times, שעוצב על ידי Stanley Morison ו־Victor Lardent עבור העיתון הבריטי The Times בשנת 1931.
- גופן המתמטיקה הראשי בספר זה הוא גרסת ה־Open Type Format של TeX Gyre Pagella Math, גופן המבוסס על Palatino, שעוצב על ידי Hermann Zapf.
- הגופן הסאנס־סריף הלטיני, ששימש לרכיבי עיצוב ולטקסט באנגלית, הוא גרסת ה־Open Type Format של TeX Gyre Heros, גופן המבוסס על Helvetica, שעוצב על ידי Max Miedinger ו־Eduard Hoffmann.
- חבילות ה־LuaLaTeX בשם fontspec ו־unicode-math, שתיהן נכתבו על ידי Will Robertson, שימשו לניהול הגופנים.
- חבילת ה־LaTeX בשם fontsize, שנכתבה על ידי Ivan Valbusa, שימשה לשינוי מעודן של הגופנים הראשיים לגודל 10.5 נקודות.
- האיורים בספר הופקו באמצעות Mathematica, בעזרת קוד Mathematica שנכתב על ידי המחבר. Mathematica נוצרה על ידי Stephen Wolfram. חבילת ה־Mathematica בשם MaTeX, שנכתבה על ידי Szabolcs Horvát, שימשה למיקום תוויות שיוצרו על ידי LaTeX באיורי ה־Mathematica.
- חבילת ה־LaTeX בשם graphicx, שנכתבה על ידי David Carlisle ו־Sebastian Rahtz, שימשה לשילוב תמונות ואיורים.
- חבילת ה־LaTeX בשם multicol, שנכתבה על ידי Frank Mittelbach, שימשה לעקיפת המגבלה של LaTeX לפיה תבנית דו־טורית חייבת להתחיל בעמוד חדש (נדרש עבור אינדקס הסמלים והאינדקס).
- חבילות ה־LaTeX בשם TikZ, שנכתבה על ידי Till Tantau, ו־tcolorbox, שנכתבה על ידי Thomas Sturm, שימשו להפקת תיבות ההגדרה ותיבות המסקנה.
- חבילת ה־LaTeX בשם color, שנכתבה על ידי David Carlisle, שימשה להוספת צבע מתאים לרכיבי עיצוב שונים.
- חבילת ה־LaTeX בשם wrapfig, שנכתבה על ידי Donald Arseneau, שימשה לעטיפת טקסט סביב תיבות ההערות.

אלגברה ליניארית בדרך הישר, מהדורה רביעית, מחבר: שלדון אקסלר
Linear Algebra Done Right, fourth edition, by Sheldon Axler

- חבילת ה־LaTeX בשם microtype, שנכתבה על ידי Robert Schlicht, שימשה להפחתת מיקוף ולהפקת יישור נאה יותר.
- המהדורה העברית תורגמה וסודרה על ידי צורי בר יוחאי. קוד ה־LaTeX של המהדורה האנגלית הותאם לסדר ימין־לשמאל עם תמיכה דו־כיוונית בנוסחאות מתמטיות.
- חבילת ה־LaTeX בשם babel, המתוחזקת על ידי Javier Bezos ושנכתבה במקור על ידי Johannes Braams, שימשה עם האפשרות bidi=basic-r לכיוון פסקה ימין־לשמאל, לטיפול דו־כיווני בטקסט, ולמיקום נכון של נוסחאות מתמטיות בתוך טקסט עברי.
- גופן הטקסט העברי במהדורה זו הוא David CLM, מפרויקט קולמוס (Culmus). גופן דוד עוצב במקור על ידי איתמר (איסמר) דוד בשנת 1954.
- לתרגום הספר פותחה על ידי צורי בר יוחאי מערכת ייעודית לפיצול קוד המקור של LaTeX ליחידות תרגום תוך שמירה על כל התוכן המתמטי, הפקודות, ומבנה המסמך. היחידות המתורגמות הורכבו מחדש ללא אובדן לקובצי ה־LaTeX העבריים.
- חבילת ה־LaTeX בשם varioref, שנכתבה על ידי Frank Mittelbach, שימשה עם הגדרות תבנית מותאמות אישית להצגת מספרי הפרקים כאותיות עבריות ולעיצוב ההפניות לפי מוסכמת מהדורה זו — אות הפרק ולאחריה מספר יחידה.