CHAPTER

5

Statue of Italian mathematician Leonardo of Pisa (1170–1250, approximate dates), also known as Fibonacci. Exercise 16 in Section 5.C shows how linear algebra can be used to find an explicit formula for the Fibonacci sequence.

Eigenvalues, Eigenvectors, and Invariant Subspaces

Linear maps from one vector space to another vector space were the objects of study in Chapter 3. Now we begin our investigation of linear maps from a finite-dimensional vector space to itself. Their study constitutes the most important part of linear algebra.

Our standing assumptions are as follows:

5.1 **Notation** $\mathbf{F}$, V

- $\mathbf{F}$ denotes $\mathbf{R}$ or $\mathbf{C}$.
- V denotes a vector space over $\mathbf{F}$.

LEARNING OBJECTIVES FOR THIS CHAPTER

- invariant subspaces
- eigenvalues, eigenvectors, and eigenspaces
- each operator on a finite-dimensional complex vector space has an eigenvalue and an upper-triangular matrix with respect to some basis

5.A *Invariant Subspaces*

In this chapter we develop the tools that will help us understand the structure of operators. Recall that an operator is a linear map from a vector space to itself. Recall also that we denote the set of operators on V by $\mathcal{L}(V)$; in other words, $\mathcal{L}(V) = \mathcal{L}(V, V)$.

Let's see how we might better understand what an operator looks like. Suppose $T \in \mathcal{L}(V)$. If we have a direct sum decomposition

$$V = U_1 \oplus \cdots \oplus U_m,$$

where each U_j is a proper subspace of V, then to understand the behavior of T, we need only understand the behavior of each $T|_{U_j}$; here $T|_{U_j}$ denotes the restriction of T to the smaller domain U_j. Dealing with $T|_{U_j}$ should be easier than dealing with T because U_j is a smaller vector space than V.

However, if we intend to apply tools useful in the study of operators (such as taking powers), then we have a problem: $T|_{U_j}$ may not map U_j into itself; in other words, $T|_{U_j}$ may not be an operator on U_j. Thus we are led to consider only decompositions of V of the form above where T maps each U_j into itself.

The notion of a subspace that gets mapped into itself is sufficiently important to deserve a name.

5.2 Definition ***invariant subspace***

Suppose $T \in \mathcal{L}(V)$. A subspace U of V is called ***invariant*** under T if $u \in U$ implies $Tu \in U$.

In other words, U is invariant under T if $T|_U$ is an operator on U.

5.3 Example Suppose $T \in \mathcal{L}(V)$. Show that each of the following subspaces of V is invariant under T:

(a) $\{0\}$;

(b) V;

(c) null T;

(d) range T.

The most famous unsolved problem in functional analysis is called the ***invariant subspace problem****. It deals with invariant subspaces of operators on infinite-dimensional vector spaces.*

Solution

(a) If $u \in \{0\}$, then $u = 0$ and hence $Tu = 0 \in \{0\}$. Thus $\{0\}$ is invariant under T.

(b) If $u \in V$, then $Tu \in V$. Thus V is invariant under T.

(c) If $u \in \operatorname{null} T$, then $Tu = 0$, and hence $Tu \in \operatorname{null} T$. Thus $\operatorname{null} T$ is invariant under T.

(d) If $u \in \operatorname{range} T$, then $Tu \in \operatorname{range} T$. Thus $\operatorname{range} T$ is invariant under 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 5.21 and 9.8.

Although $\operatorname{null} T$ and $\operatorname{range} T$ are invariant under T, they do not necessarily provide easy answers to the question 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).

5.4 **Example** Suppose that $T \in \mathcal{L}\big(\mathcal{P}(\mathbf{R})\big)$ 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.

Eigenvalues and Eigenvectors

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 with dimension 1.

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 1-dimensional subspace of V (and every 1-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 1-dimensional subspace of V invariant under T.

The equation

$$Tv = \lambda v,$$

which we have just seen is intimately connected with 1-dimensional invariant subspaces, is important enough that the vectors v and scalars λ 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 adjective **eigen** means "own" in the sense of characterizing an intrinsic property. Some mathematicians use the term **characteristic value** instead of eigenvalue.*

The comments above show that T has a 1-dimensional invariant subspace if and only if T has an eigenvalue.

In the definition above, we require that $v \neq 0$ because every scalar $\lambda \in \mathbf{F}$ satisfies $T0 = \lambda 0$.

5.6 Equivalent conditions to be an eigenvalue

Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and $\lambda \in F$. 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.

Recall that $I \in \mathcal{L}(V)$ is the identity operator defined by $Iv = v$ 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 3.69. ∎

5.7 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 $Tv = \lambda v$.

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 $v \in \operatorname{null}(T - \lambda I)$.

5.8 **Example** Suppose $T \in \mathcal{L}(\mathbf{F}^2)$ is defined by

$$T(w, z) = (-z, w).$$

(a) Find the eigenvalues and eigenvectors of T if $\mathbf{F} = \mathbf{R}$.

(b) Find the eigenvalues and eigenvectors of T if $\mathbf{F} = \mathbf{C}$.

Solution

(a) If $\mathbf{F} = \mathbf{R}$, 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 $\mathbf{R}^2$ obviously never equals a scalar multiple of itself. Conclusion: if $\mathbf{F} = \mathbf{R}$, then T has no eigenvalues (and thus has no eigenvectors).

(b) 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 above is equivalent to the simultaneous equations

5.9 $$-z = \lambda w, \quad w = \lambda z.$$

Substituting the value for w given by the second equation into the first equation gives

$$-z = \lambda^2 z.$$

Now z cannot equal 0 [otherwise 5.9 implies that $w = 0$; we are looking for solutions to 5.9 where (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 should be able to verify easily 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 $w \neq 0$, and the eigenvectors corresponding to the eigenvalue $-i$ are the vectors of the form (w, wi), with $w \in \mathbf{C}$ and $w \neq 0$.

Now we show that eigenvectors corresponding to distinct eigenvalues are linearly independent.

5.10 Linearly independent eigenvectors

Let $T \in \mathcal{L}(V)$. Suppose $\lambda_1, \dots, \lambda_m$ are distinct eigenvalues of T and $v_1, \dots, v_m$ are corresponding eigenvectors. Then $v_1, \dots, v_m$ is linearly independent.

Proof Suppose $v_1, \dots, v_m$ is linearly dependent. Let k be the smallest positive integer such that

5.11 $$v_k \in \operatorname{span}(v_1, \dots, v_{k-1});$$

the existence of k with this property follows from the Linear Dependence Lemma (2.21). Thus there exist $a_1, \dots, a_{k-1} \in \mathbf{F}$ such that

5.12 $$v_k = a_1 v_1 + \cdots + a_{k-1} v_{k-1}.$$

Apply T to both sides of this equation, getting

$$\lambda_k v_k = a_1 \lambda_1 v_1 + \cdots + a_{k-1} \lambda_{k-1} v_{k-1}.$$

Multiply both sides of 5.12 by λ_k and then subtract the equation above, getting

$$0 = a_1(\lambda_k - \lambda_1) v_1 + \cdots + a_{k-1}(\lambda_k - \lambda_{k-1}) v_{k-1}.$$

Because we chose k to be the smallest positive integer satisfying 5.11, $v_1, \dots, v_{k-1}$ is linearly independent. Thus the equation above implies that all the a's are 0 (recall that λ_k is not equal to any of $\lambda_1, \dots, \lambda_{k-1}$). However, this means that v_k equals 0 (see 5.12), contradicting our hypothesis that v_k is an eigenvector. Therefore our assumption that $v_1, \dots, v_m$ is linearly dependent was false. ■

The corollary below states that an operator cannot have more distinct eigenvalues than the dimension of the vector space on which it acts.

5.13 Number of eigenvalues

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 5.10 implies that the list $v_1, \dots, v_m$ is linearly independent. Thus $m \le \dim V$ (see 2.23), as desired. ■

Restriction and Quotient Operators

If $T \in \mathcal{L}(V)$ and U is a subspace of V invariant under T, then U determines two other operators $T|_U \in \mathcal{L}(U)$ and $T/U \in \mathcal{L}(V/U)$ in a natural way, as defined below.

5.14 Definition $T|_U$ ***and*** T/U

Suppose $T \in \mathcal{L}(V)$ and U is a subspace of V invariant under T.

- The ***restriction operator*** $T|_U \in \mathcal{L}(U)$ is defined by
$$T|_U(u) = Tu$$
for $u \in U$.

- The ***quotient operator*** $T/U \in \mathcal{L}(V/U)$ is defined by
$$(T/U)(v + U) = Tv + U$$
for $v \in V$.

For both the operators defined above, it is worthwhile to pay attention to their domains and to spend a moment thinking about why they are well defined as operators on their domains. First consider the restriction operator $T|_U \in \mathcal{L}(U)$, which is T with its domain restricted to U, thought of as mapping into U instead of into V. The condition that U is invariant under T is what allows us to think of $T|_U$ as an operator on U, meaning a linear map into the same space as the domain, rather than as simply a linear map from one vector space to another vector space.

To show that the definition above of the quotient operator makes sense, we need to verify that if $v + U = w + U$, then $Tv + U = Tw + U$. Hence suppose $v + U = w + U$. Thus $v - w \in U$ (see 3.85). Because U is invariant under T, we also have $T(v - w) \in U$, which implies that $Tv - Tw \in U$, which implies that $Tv + U = Tw + U$, as desired.

Suppose T is an operator on a finite-dimensional vector space V and U is a subspace of V invariant under T, with $U \neq \{0\}$ and $U \neq V$. In some sense, we can learn about T by studying the operators $T|_U$ and T/U, each of which is an operator on a vector space with smaller dimension than V. For example, proof 2 of 5.27 makes nice use of T/U.

However, sometimes $T|_U$ and T/U do not provide enough information about T.In the next example, both $T|_U$ and T/U are 0 even though T is not the 0 operator.

5.15 **Example** Define an operator $T \in \mathcal{L}(\mathbf{F}^2)$ by $T(x, y) = (y, 0)$. Let $U = \{(x, 0) : x \in \mathbf{F}\}$. Show that

(a) U is invariant under T and $T|_U$ is the 0 operator on U;

(b) there does not exist a subspace W of $\mathbf{F}^2$ that is invariant under T and such that $\mathbf{F}^2 = U \oplus W$;

(c) T/U is the 0 operator on $\mathbf{F}^2/U$.

Solution

(a) For $(x, 0) \in U$, we have $T(x, 0) = (0, 0) \in U$. Thus U is invariant under T and $T|_U$ is the 0 operator on U.

(b) Suppose W is a subspace of V such that $\mathbf{F}^2 = U \oplus W$. Because $\dim \mathbf{F}^2 = 2$ and $\dim U = 1$, we have $\dim W = 1$. If W were invariant under T, then each nonzero vector in W would be an eigenvector of T. However, it is easy to see that 0 is the only eigenvalue of T and that all eigenvectors of T are in U. Thus W is not invariant under T.

(c) For $(x, y) \in \mathbf{F}^2$, we have

$$\begin{aligned}(T/U)\big((x, y) + U\big) &= T(x, y) + U \\ &= (y, 0) + U \\ &= 0 + U,\end{aligned}$$

where the last equality holds because $(y, 0) \in U$. The equation above shows that T/U is the 0 operator.

EXERCISES 5.A

1 Suppose $T \in \mathcal{L}(V)$ and U is a subspace of V.

(a) Prove that if $U \subset \operatorname{null} T$, then U is invariant under T.

(b) Prove that if $\operatorname{range} T \subset U$, then U is invariant under T.

2 Suppose $S, T \in \mathcal{L}(V)$ are such that $ST = TS$. Prove that $\operatorname{null} S$ is invariant under T.

3 Suppose $S, T \in \mathcal{L}(V)$ are such that $ST = TS$. Prove that range S is invariant under T.

4 Suppose that $T \in \mathcal{L}(V)$ and $U_1, \dots, U_m$ are subspaces of V invariant under T. Prove that $U_1 + \cdots + U_m$ is invariant under T.

5 Suppose $T \in \mathcal{L}(V)$. Prove that the intersection of every collection of subspaces of V invariant under T is invariant under T.

6 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 $U = V$.

7 Suppose $T \in \mathcal{L}(\mathbf{R}^2)$ is defined by $T(x, y) = (-3y, x)$. Find the eigenvalues of T.

8 Define $T \in \mathcal{L}(\mathbf{F}^2)$ by

$$T(w, z) = (z, w).$$

Find all eigenvalues and eigenvectors of T.

9 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.

10 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 invariant subspaces of T.

11 Define $T\colon \mathcal{P}(\mathbf{R}) \to \mathcal{P}(\mathbf{R})$ by $Tp = p'$. Find all eigenvalues and eigenvectors of T.

12 Define $T \in \mathcal{L}\big(\mathcal{P}_4(\mathbf{R})\big)$ by

$$(Tp)(x) = xp'(x)$$

for all $x \in \mathbf{R}$. Find all eigenvalues and eigenvectors of T.

13 Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and $\lambda \in \mathbf{F}$. Prove that there exists $\alpha \in \mathbf{F}$ such that $|\alpha - \lambda| < \frac{1}{1000}$ and $T - \alpha I$ is invertible.

14 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 $u \in U$ and $w \in W$. Find all eigenvalues and eigenvectors of P.

15 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 $S^{-1}TS$?

16 Suppose V is a complex vector space, $T \in \mathcal{L}(V)$, and the matrix of T with respect to some basis of V contains only real entries. Show that if λ is an eigenvalue of T, then so is $\bar{\lambda}$.

17 Give an example of an operator $T \in \mathcal{L}(\mathbf{R}^4)$ such that T has no (real) eigenvalues.

18 Show that the operator $T \in \mathcal{L}(\mathbf{C}^\infty)$ defined by

$$T(z_1, z_2, \ldots) = (0, z_1, z_2, \ldots)$$

has no eigenvalues.

19 Suppose n is a positive integer and $T \in \mathcal{L}(\mathbf{F}^n)$ is defined by

$$T(x_1, \ldots, x_n) = (x_1 + \cdots + x_n, \ldots, x_1 + \cdots + x_n);$$

in other words, T is the operator whose matrix (with respect to the standard basis) consists of all 1's. Find all eigenvalues and eigenvectors of T.

20 Find all eigenvalues and eigenvectors of the backward shift operator $T \in \mathcal{L}(\mathbf{F}^\infty)$ defined by

$$T(z_1, z_2, z_3, \ldots) = (z_2, z_3, \ldots).$$

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 v and w in V such that

$$Tv = 3w \quad \text{and} \quad Tw = 3v.$$

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, v are eigenvectors of T such that $u + v$ is also an eigenvector of T. Prove that u and v 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 V is finite-dimensional and $T \in \mathcal{L}(V)$ is such that every subspace of V with dimension $\dim V - 1$ is invariant under T. Prove that T is a scalar multiple of the identity operator.

28 Suppose V is finite-dimensional with $\dim V \geq 3$ and $T \in \mathcal{L}(V)$ is such that every 2-dimensional subspace of V is invariant under T. Prove that T is a scalar multiple of the identity operator.

29 Suppose $T \in \mathcal{L}(V)$ and $\dim \operatorname{range} T = k$. Prove that T has at most $k + 1$ distinct eigenvalues.

30 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 $Tx - 9x = (-4, 5, \sqrt{7})$.

31 Suppose V is finite-dimensional and $v_1, \dots, v_m$ is a list of vectors in V. Prove that $v_1, \dots, v_m$ 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.

32 Suppose $\lambda_1, \dots, \lambda_n$ 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 $V = \operatorname{span}\bigl(e^{\lambda_1 x}, \dots, e^{\lambda_n x}\bigr)$, and define an operator $T \in \mathcal{L}(V)$ by $Tf = f'$. Find eigenvalues and eigenvectors of T.

33 Suppose $T \in \mathcal{L}(V)$. Prove that $T/(\operatorname{range} T) = 0$.

34 Suppose $T \in \mathcal{L}(V)$. Prove that $T/(\operatorname{null} T)$ is injective if and only if $(\operatorname{null} T) \cap (\operatorname{range} T) = \{0\}$.

35 Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and U is invariant under T. Prove that each eigenvalue of T/U is an eigenvalue of T.
[*The exercise below asks you to verify that the hypothesis that V is finite-dimensional is needed for the exercise above.*]

36 Give an example of a vector space V, an operator $T \in \mathcal{L}(V)$, and a subspace U of V that is invariant under T such that T/U has an eigenvalue that is not an eigenvalue of T.

5.B Eigenvectors and Upper-Triangular Matrices

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. We begin this section by defining that notion and the key concept of applying a polynomial to an operator.

If $T \in \mathcal{L}(V)$, then TT makes sense and is also in $\mathcal{L}(V)$. We usually write T^2 instead of TT. More generally, we have the following definition.

5.16 Definition T^m

Suppose $T \in \mathcal{L}(V)$ and m is a positive integer.

- T^m 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} 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 allowed to be arbitrary integers if T is invertible and nonnegative integers if T is not invertible.

5.17 Definition $p(T)$

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 $z \in \mathbf{F}$. Then $p(T)$ is the operator 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 it to operators, not just elements of $\mathbf{F}$.

5.18 Example 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

$$\big(p(D)\big)q = 7q - 3q' + 5q''$$

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.

5.19 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 $z \in \mathbf{F}$.

Any two polynomials of an operator commute, as shown below.

5.20 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) $p(T)q(T) = q(T)p(T)$.

Part (a) holds because when expanding a product of polynomials 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 $z \in \mathbf{F}$. 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) Part (a) implies $p(T)q(T) = (pq)(T) = (qp)(T) = q(T)p(T)$. ■

Existence of Eigenvalues

Now we come to one of the central results about operators on complex vector spaces.

5.21 Operators on complex vector spaces have an eigenvalue

Every operator on a finite-dimensional, nonzero, complex vector space has an eigenvalue.

Proof Suppose V is a complex vector space with dimension $n > 0$ and $T \in \mathcal{L}(V)$. Choose $v \in V$ with $v \neq 0$. Then

$$v, Tv, T^2v, \ldots, T^n v$$

is not linearly independent, because V has dimension n and we have $n + 1$ vectors. Thus there exist complex numbers $a_0, \ldots, a_n$, not all 0, such that

$$0 = a_0 v + a_1 T v + \cdots + a_n T^n v.$$

Note that $a_1, \ldots, a_n$ cannot all be 0, because otherwise the equation above would become $0 = a_0 v$, which would force a_0 also to be 0.

Make the a's the coefficients of a polynomial, which by the Fundamental Theorem of Algebra (4.14) has a factorization

$$a_0 + a_1 z + \cdots + a_n z^n = c(z - \lambda_1) \cdots (z - \lambda_m),$$

where c is a nonzero complex number, each λ_j is in $\mathbf{C}$, and the equation holds for all $z \in \mathbf{C}$ (here m is not necessarily equal to n, because a_n may equal 0). We then have

$$\begin{aligned} 0 &= a_0 v + a_1 T v + \cdots + a_n T^n v \\ &= (a_0 I + a_1 T + \cdots + a_n T^n) v \\ &= c(T - \lambda_1 I) \cdots (T - \lambda_m I) v. \end{aligned}$$

Thus $T - \lambda_j I$ is not injective for at least one j. In other words, T has an eigenvalue. ■

The proof above depends on the Fundamental Theorem of Algebra, which is typical of proofs of this result. See Exercises 16 and 17 for possible ways to rewrite the proof above using the idea of the proof in a slightly different form.

Upper-Triangular Matrices

In Chapter 3 we discussed the matrix of a linear map from one vector space to another vector space. That matrix depended on a choice of a 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.

5.22 Definition ***matrix of an operator, $\mathcal{M}(T)$***

Suppose $T \in \mathcal{L}(V)$ and $v_1, \ldots, v_n$ is a basis of V. The ***matrix of*** T with respect to this basis is the n-by-n matrix

$$\mathcal{M}(T) = \begin{pmatrix} A_{1,1} & \ldots & A_{1,n} \\ \vdots & & \vdots \\ A_{n,1} & \ldots & A_{n,n} \end{pmatrix}$$

whose entries $A_{j,k}$ are defined by

$$Tv_k = A_{1,k}v_1 + \cdots + A_{n,k}v_n.$$

If the basis is not clear from the context, then the notation $\mathcal{M}\big(T, (v_1, \ldots, v_n)\big)$ is used.

Note that the matrices of operators are square arrays, rather than the more general rectangular arrays that we considered earlier for linear maps.

The k^{th} column of the matrix $\mathcal{M}(T)$ is formed from the coefficients used to write Tv_k as a linear combination of $v_1, \ldots, v_n$.

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 j^{th} basis vector is 1 in the j^{th} slot and 0 in all the other slots). You can then think of the j^{th} column of $\mathcal{M}(T)$ as T applied to the j^{th} basis vector.

5.23 Example Define $T \in \mathcal{L}(\mathbf{F}^3)$ by $T(x, y, z) = (2x+y, 5y+3z, 8z)$. Then

$$\mathcal{M}(T) = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 5 & 3 \\ 0 & 0 & 8 \end{pmatrix}.$$

A central goal of linear algebra is to show that given an operator $T \in \mathcal{L}(V)$, 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.

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 the $*$ denotes the entries in all the columns other than the first column. To prove this, let λ be an eigenvalue of T (one exists by 5.21) 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.

5.24 Definition ***diagonal of a matrix***

The ***diagonal*** of a square matrix consists of the entries along the line from the upper left corner to the bottom right corner.

For example, the diagonal of the matrix in 5.23 consists of the entries $2, 5, 8$.

5.25 Definition ***upper-triangular matrix***

A matrix is called ***upper triangular*** if all the entries below the diagonal equal 0.

For example, the matrix in 5.23 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—for n large, almost half its entries in an n-by-n upper-triangular matrix are 0.

We often use $*$ to denote matrix entries that we do not know about or that are irrelevant to the questions being discussed.

The following proposition demonstrates a useful connection between upper-triangular matrices and invariant subspaces.

5.26 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) $Tv_j \in \operatorname{span}(v_1, \dots, v_j)$ for each $j = 1, \dots, n$;

(c) $\operatorname{span}(v_1, \dots, v_j)$ is invariant under T for each $j = 1, \dots, n$.

Proof The equivalence of (a) and (b) follows easily from the definitions and a moment's thought. Obviously (c) implies (b). Hence to complete the proof, we need only prove that (b) implies (c).

Thus suppose (b) holds. Fix $j \in \{1, \dots, n\}$. From (b), we know that

$$\begin{aligned} Tv_1 &\in \operatorname{span}(v_1) \subset \operatorname{span}(v_1, \dots, v_j); \\ Tv_2 &\in \operatorname{span}(v_1, v_2) \subset \operatorname{span}(v_1, \dots, v_j); \\ &\vdots \\ Tv_j &\in \operatorname{span}(v_1, \dots, v_j). \end{aligned}$$

Thus if v is a linear combination of $v_1, \dots, v_j$, then

$$Tv \in \operatorname{span}(v_1, \dots, v_j).$$

In other words, $\operatorname{span}(v_1, \dots, v_j)$ is invariant under T, completing the proof. ■

The next result does not hold on real vector spaces, because the first vector in a basis with respect to which an operator has an upper-triangular matrix is an eigenvector of the operator. Thus if an operator on a real vector space has no eigenvalues [see 5.8(a) for an example], then there is no basis with respect to which the operator has an upper-triangular matrix.

Now we can prove 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 only 0's below the diagonal. In Chapter 8 we will improve even this result.

Sometimes more insight comes from seeing more than one proof of a theorem. Thus two proofs are presented of the next result. Use whichever appeals more to you.

5.27 Over $\mathbf{C}$, every operator 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 1 We will use induction on the dimension of V. Clearly the desired result holds if $\dim V = 1$.

Suppose now that $\dim V > 1$ and the desired result holds for all complex vector spaces whose dimension is less than the dimension of V. Let λ be any eigenvalue of T (5.21 guarantees that T has an eigenvalue). Let

$$U = \operatorname{range}(T - \lambda I).$$

Because $T - \lambda I$ is not surjective (see 3.69), $\dim U < \dim V$. Furthermore, U is invariant under T. To prove this, suppose $u \in U$. Then

$$Tu = (T - \lambda I)u + \lambda u.$$

Obviously $(T - \lambda I)u \in U$ (because U equals the range of $T - \lambda I$) and $\lambda u \in U$. Thus the equation above shows that $Tu \in U$. Hence U is invariant under T, as claimed.

Thus $T|_U$ is an operator on U. By our induction hypothesis, there is a basis $u_1, \dots, u_m$ of U with respect to which $T|_U$ has an upper-triangular matrix. Thus for each j we have (using 5.26)

5.28 $$Tu_j = (T|_U)(u_j) \in \operatorname{span}(u_1, \dots, u_j).$$

Extend $u_1, \dots, u_m$ to a basis $u_1, \dots, u_m, v_1, \dots, v_n$ of V. For each k, we have

$$Tv_k = (T - \lambda I)v_k + \lambda v_k.$$

The definition of U shows that $(T - \lambda I)v_k \in U = \operatorname{span}(u_1, \dots, u_m)$. Thus the equation above shows that

5.29 $$Tv_k \in \operatorname{span}(u_1, \dots, u_m, v_1, \dots, v_k).$$

From 5.28 and 5.29, we conclude (using 5.26) that T has an upper-triangular matrix with respect to the basis $u_1, \dots, u_m, v_1, \dots, v_n$ of V, as desired. ■

Proof 2 We will use induction on the dimension of V. Clearly the desired result holds if $\dim V = 1$.

Suppose now that $\dim V = n > 1$ and the desired result holds for all complex vector spaces whose dimension is $n - 1$. Let v_1 be any eigenvector of T (5.21 guarantees that T has an eigenvector). Let $U = \operatorname{span}(v_1)$. Then U is an invariant subspace of T and $\dim U = 1$.

Because $\dim V/U = n - 1$ (see 3.89), we can apply our induction hypothesis to $T/U \in \mathcal{L}(V/U)$. Thus there is a basis $v_2 + U, \dots, v_n + U$ of V/U such that T/U has an upper-triangular matrix with respect to this basis. Hence by 5.26,

$$(T/U)(v_j + U) \in \operatorname{span}(v_2 + U, \dots, v_j + U)$$

for each $j = 2, \dots, n$. Unraveling the meaning of the inclusion above, we see that

$$Tv_j \in \operatorname{span}(v_1, \dots, v_j)$$

for each $j = 1, \dots, n$. Thus by 5.26, T has an upper-triangular matrix with respect to the basis $v_1, \dots, v_n$ of V, as desired (it is easy to verify that $v_1, \dots, v_n$ is a basis of V; see Exercise 13 in Section 3.E for a more general result). ■

How does one determine from looking at the matrix of an operator whether the operator is invertible? If we are fortunate enough to have a basis with respect to which the matrix of the operator is upper triangular, then this problem becomes easy, as the following proposition shows.

5.30 Determination of invertibility from upper-triangular matrix

Suppose $T \in \mathcal{L}(V)$ has an upper-triangular matrix with respect to some basis of V. Then T is invertible if and only if all the entries on the diagonal of that upper-triangular matrix are nonzero.

Proof Suppose $v_1, \dots, v_n$ is a basis of V with respect to which T has an upper-triangular matrix

5.31
$$\mathcal{M}(T) = \begin{pmatrix} \lambda_1 & & & * \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}.$$

We need to prove that T is invertible if and only if all the λ_j's are nonzero.

First suppose the diagonal entries $\lambda_1, \ldots, \lambda_n$ are all nonzero. The upper-triangular matrix in 5.31 implies that $Tv_1 = \lambda_1 v_1$. Because $\lambda_1 \neq 0$, we have $T(v_1/\lambda_1) = v_1$; thus $v_1 \in \operatorname{range} T$.

Now

$$T(v_2/\lambda_2) = av_1 + v_2$$

for some $a \in \mathbf{F}$. The left side of the equation above and av_1 are both in $\operatorname{range} T$; thus $v_2 \in \operatorname{range} T$.

Similarly, we see that

$$T(v_3/\lambda_3) = bv_1 + cv_2 + v_3$$

for some $b, c \in \mathbf{F}$. The left side of the equation above and bv_1, cv_2 are all in $\operatorname{range} T$; thus $v_3 \in \operatorname{range} T$.

Continuing in this fashion, we conclude that $v_1, \ldots, v_n \in \operatorname{range} T$. Because $v_1, \ldots, v_n$ is a basis of V, this implies that $\operatorname{range} T = V$. In other words, T is surjective. Hence T is invertible (by 3.69), as desired.

To prove the other direction, now suppose that T is invertible. This implies that $\lambda_1 \neq 0$, because otherwise we would have $Tv_1 = 0$.

Let $1 < j \leq n$, and suppose $\lambda_j = 0$. Then 5.31 implies that T maps $\operatorname{span}(v_1, \ldots, v_j)$ into $\operatorname{span}(v_1, \ldots, v_{j-1})$. Because

$$\dim \operatorname{span}(v_1, \ldots, v_j) = j \quad \text{and} \quad \dim \operatorname{span}(v_1, \ldots, v_{j-1}) = j - 1,$$

this implies that T restricted to $\operatorname{span}(v_1, \ldots, v_j)$ is not injective (by 3.23). Thus there exists $v \in \operatorname{span}(v_1, \ldots, v_j)$ such that $v \neq 0$ and $Tv = 0$. Thus T is not injective, which contradicts our hypothesis (for this direction) that T is invertible. This contradiction means that our assumption that $\lambda_j = 0$ must be false. Hence $\lambda_j \neq 0$, as desired. ■

As an example of the result above, we see that the operator in Example 5.23 is invertible.

Powerful numeric techniques exist for finding good approximations to the eigenvalues of an operator from its matrix.

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 following proposition shows.

5.32 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

$$\mathcal{M}(T) = \begin{pmatrix} \lambda_1 & & & * \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}.$$

Let $\lambda \in \mathbf{F}$. Then

$$\mathcal{M}(T - \lambda I) = \begin{pmatrix} \lambda_1 - \lambda & & & * \\ & \lambda_2 - \lambda & & \\ & & \ddots & \\ 0 & & & \lambda_n - \lambda \end{pmatrix}.$$

Hence $T - \lambda I$ is not invertible if and only if λ equals one of the numbers $\lambda_1, \dots, \lambda_n$ (by 5.30). Thus λ is an eigenvalue of T if and only if λ equals one of the numbers $\lambda_1, \dots, \lambda_n$. ■

5.33 Example Define $T \in \mathcal{L}(\mathbf{F}^3)$ by $T(x, y, z) = (2x + y, 5y + 3z, 8z)$. What are the eigenvalues of T?

Solution The matrix of T with respect to the standard basis is

$$\mathcal{M}(T) = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 5 & 3 \\ 0 & 0 & 8 \end{pmatrix}.$$

Thus $\mathcal{M}(T)$ is an upper-triangular matrix. Now 5.32 implies that the eigenvalues of T are 2, 5, and 8.

Once the eigenvalues of an operator on $\mathbf{F}^n$ are known, the eigenvectors can be found easily using Gaussian elimination.

EXERCISES 5.B

1 Suppose $T \in \mathcal{L}(V)$ and there exists a positive integer n such that $T^n = 0$.

(a) Prove that $I - T$ is invertible and that

$$(I-T)^{-1} = I + T + \cdots + T^{n-1}.$$

(b) Explain how you would guess the formula above.

2 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 $\lambda = 4$.

3 Suppose $T \in \mathcal{L}(V)$ and $T^2 = I$ and -1 is not an eigenvalue of T. Prove that $T = I$.

4 Suppose $P \in \mathcal{L}(V)$ and $P^2 = P$. Prove that $V = \operatorname{null} P \oplus \operatorname{range} P$.

5 Suppose $S, T \in \mathcal{L}(V)$ and S is invertible. Suppose $p \in \mathcal{P}(\mathbf{F})$ is a polynomial. Prove that

$$p(STS^{-1}) = Sp(T)S^{-1}.$$

6 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 $p \in \mathcal{P}(\mathbf{F})$.

7 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.

8 Give an example of $T \in \mathcal{L}(\mathbf{R}^2)$ such that $T^4 = -I$.

9 Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and $v \in V$ with $v \neq 0$. Let p be a nonzero polynomial of smallest degree such that $p(T)v = 0$. Prove that every zero of p is an eigenvalue of T.

10 Suppose $T \in \mathcal{L}(V)$ and v is an eigenvector of T with eigenvalue λ. Suppose $p \in \mathcal{P}(\mathbf{F})$. Prove that $p(T)v = p(\lambda)v$.

11 Suppose $\mathbf{F} = \mathbf{C}$, $T \in \mathcal{L}(V)$, $p \in \mathcal{P}(\mathbf{C})$ is a 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.

12 Show that the result in the previous exercise does not hold if $\mathbf{C}$ is replaced with $\mathbf{R}$.

13 Suppose W is a complex vector space and $T \in \mathcal{L}(W)$ has no eigenvalues. Prove that every subspace of W invariant under T is either $\{0\}$ or infinite-dimensional.

14 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.
[*The exercise above and the exercise below show that 5.30 fails without the hypothesis that an upper-triangular matrix is under consideration.*]

15 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.

16 Rewrite the proof of 5.21 using the linear map that sends $p \in \mathcal{P}_n(\mathbf{C})$ to $\bigl(p(T)\bigr)v \in V$ (and use 3.23).

17 Rewrite the proof of 5.21 using the linear map that sends $p \in \mathcal{P}_{n^2}(\mathbf{C})$ to $p(T) \in \mathcal{L}(V)$ (and use 3.23).

18 Suppose V is a finite-dimensional complex vector space and $T \in \mathcal{L}(V)$. Define a function $f : \mathbf{C} \to \mathbf{R}$ by

$$f(\lambda) = \dim \operatorname{range}(T - \lambda I).$$

Prove that f is not a continuous function.

19 Suppose V is finite-dimensional with $\dim V > 1$ and $T \in \mathcal{L}(V)$. Prove that

$$\{p(T) : p \in \mathcal{P}(\mathbf{F})\} \neq \mathcal{L}(V).$$

20 Suppose V is a finite-dimensional complex vector space and $T \in \mathcal{L}(V)$. Prove that T has an invariant subspace of dimension k for each $k = 1, \dots, \dim V$.

5.C Eigenspaces and Diagonal Matrices

5.34 Definition ***diagonal matrix***

A ***diagonal matrix*** is a square matrix that is 0 everywhere except possibly along the diagonal.

5.35 Example

$$\begin{pmatrix} 8 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 5 \end{pmatrix}$$

is a diagonal matrix.

Obviously every diagonal matrix is upper triangular. In general, a diagonal matrix has many more 0's than an upper-triangular matrix.

If an operator has a diagonal matrix with respect to some basis, then the entries along the diagonal are precisely the eigenvalues of the operator; this follows from 5.32 (or find an easier proof for diagonal matrices).

5.36 Definition ***eigenspace,*** $E(\lambda, T)$

Suppose $T \in \mathcal{L}(V)$ and $\lambda \in \mathbf{F}$. The ***eigenspace*** of T corresponding to λ, denoted $E(\lambda, T)$, is defined by

$$E(\lambda, T) = \operatorname{null}(T - \lambda I).$$

In other words, $E(\lambda, T)$ 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 eigenspace $E(\lambda, T)$ 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 $E(\lambda, T) \neq \{0\}$.

5.37 Example 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 5.35 above. Then

$$E(8, T) = \operatorname{span}(v_1), \qquad E(5, T) = \operatorname{span}(v_2, v_3).$$

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.

5.38 Sum of eigenspaces is a direct sum

Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Suppose also that $\lambda_1, \dots, \lambda_m$ are distinct eigenvalues of T. Then

$$E(\lambda_1, T) + \cdots + E(\lambda_m, T)$$

is a direct sum. Furthermore,

$$\dim E(\lambda_1, T) + \cdots + \dim E(\lambda_m, T) \leq \dim V.$$

Proof To show that $E(\lambda_1, T) + \cdots + E(\lambda_m, T)$ is a direct sum, suppose

$$u_1 + \cdots + u_m = 0,$$

where each u_j is in $E(\lambda_j, T)$. Because eigenvectors corresponding to distinct eigenvalues are linearly independent (see 5.10), this implies that each u_j equals 0. This implies (using 1.44) that $E(\lambda_1, T) + \cdots + E(\lambda_m, T)$ is a direct sum, as desired.

Now

$$\begin{aligned} \dim E(\lambda_1, T) + \cdots + \dim E(\lambda_m, T) &= \dim\bigl(E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T)\bigr) \\ &\leq \dim V, \end{aligned}$$

where the equality above follows from Exercise 16 in Section 2.C. ■

5.39 Definition *diagonalizable*

An operator $T \in \mathcal{L}(V)$ is called ***diagonalizable*** if the operator has a diagonal matrix with respect to some basis of V.

5.40 Example 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, because the matrix of T with respect to the basis $(1, 4), (7, 5)$ is

$$\begin{pmatrix} 69 & 0 \\ 0 & 46 \end{pmatrix},$$

as you should verify.

5.41 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) there exist 1-dimensional subspaces $U_1, \dots, U_n$ of V, each invariant under T, such that

$$V = U_1 \oplus \cdots \oplus U_n;$$

(d) $V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T)$;

(e) $\dim V = \dim E(\lambda_1, T) + \cdots + \dim E(\lambda_m, T)$.

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 $T v_j = \lambda_j v_j$ for each j. Thus (a) and (b) are equivalent.

Suppose (b) holds; thus V has a basis $v_1, \dots, v_n$ consisting of eigenvectors of T. For each j, let $U_j = \operatorname{span}(v_j)$. Obviously each U_j is a 1-dimensional subspace of V that is invariant under T. Because $v_1, \dots, v_n$ is a basis of V, each vector in V can be written uniquely as a linear combination of $v_1, \dots, v_n$. In other words, each vector in V can be written uniquely as a sum $u_1 + \cdots + u_n$, where each u_j is in U_j. Thus $V = U_1 \oplus \cdots \oplus U_n$. Hence (b) implies (c).

Suppose now that (c) holds; thus there are 1-dimensional subspaces $U_1, \dots, U_n$ of V, each invariant under T, such that $V = U_1 \oplus \cdots \oplus U_n$. For each j, let v_j be a nonzero vector in U_j. Then each v_j is an eigenvector of T. Because each vector in V can be written uniquely as a sum $u_1 + \cdots + u_n$, where each u_j is in U_j (so each u_j is a scalar multiple of v_j), we see that $v_1, \dots, v_n$ is a basis of V. Thus (c) implies (b).

At this stage of the proof we know that (a), (b), and (c) are all equivalent. We will finish the proof by showing that (b) implies (d), that (d) implies (e), and that (e) implies (b).

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 5.38 shows that (d) holds.

That (d) implies (e) follows immediately from Exercise 16 in Section 2.C.

Finally, suppose (e) holds; thus

5.42 $$\dim V = \dim E(\lambda_1, T) + \cdots + \dim E(\lambda_m, T).$$

Choose a basis of each $E(\lambda_j, T)$; put all these bases together to form a list $v_1, \dots, v_n$ of eigenvectors of T, where $n = \dim V$ (by 5.42). 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 $j = 1, \dots, m$, let u_j denote the sum of all the terms $a_k v_k$ such that $v_k \in E(\lambda_j, T)$. Thus each u_j is in $E(\lambda_j, T)$, and

$$u_1 + \cdots + u_m = 0.$$

Because eigenvectors corresponding to distinct eigenvalues are linearly independent (see 5.10), this implies that each u_j equals 0. Because each u_j is a sum of terms $a_k v_k$, where the v_k's were chosen to be a basis of $E(\lambda_j, T)$, this implies that all the a_k's equal 0. Thus $v_1, \dots, v_n$ is linearly independent and hence is a basis of V (by 2.39). Thus (e) implies (b), completing the proof. ■

Unfortunately not every operator is diagonalizable. This sad state of affairs can arise even on complex vector spaces, as shown by the next example.

5.43 **Example** Show that the operator $T \in \mathcal{L}(\mathbf{C}^2)$ defined by

$$T(w, z) = (z, 0)$$

is not diagonalizable.

Solution As you should verify, 0 is the only eigenvalue of T and furthermore $E(0, T) = \{(w, 0) \in \mathbf{C}^2 : w \in \mathbf{C}\}$.

Thus conditions (b), (c), (d), and (e) of 5.41 are easily seen to fail (of course, because these conditions are equivalent, it is only necessary to check that one of them fails). Thus condition (a) of 5.41 also fails, and hence T is not diagonalizable.

The next result shows that if an operator has as many distinct eigenvalues as the dimension of its domain, then the operator is diagonalizable.

5.44 Enough eigenvalues implies diagonalizability

If $T \in \mathcal{L}(V)$ has $\dim V$ distinct eigenvalues, then T is diagonalizable.

Proof Suppose $T \in \mathcal{L}(V)$ has $\dim V$ distinct eigenvalues $\lambda_1, \dots, \lambda_{\dim V}$. For each j, let $v_j \in V$ be an eigenvector corresponding to the eigenvalue λ_j. Because eigenvectors corresponding to distinct eigenvalues are linearly independent (see 5.10), $v_1, \dots, v_{\dim V}$ is linearly independent. A linearly independent list of $\dim V$ vectors in V is a basis of V (see 2.39); 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. ■

5.45 **Example** Define $T \in \mathcal{L}(\mathbf{F}^3)$ by $T(x, y, z) = (2x + y, 5y + 3z, 8z)$. Find a basis of $\mathbf{F}^3$ with respect to which T has a diagonal matrix.

Solution 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 upper triangular. Thus by 5.32, the eigenvalues of T are 2, 5, and 8. Because T is an operator on a vector space with dimension 3 and T has three distinct eigenvalues, 5.44 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 $\lambda = 8$. These simple equations are easy to solve: for $\lambda = 2$ we have the eigenvector $(1, 0, 0)$; for $\lambda = 5$ we have the eigenvector $(1, 3, 0)$; for $\lambda = 8$ we have the eigenvector $(1, 6, 6)$.

Thus $(1, 0, 0), (1, 3, 0), (1, 6, 6)$ is a basis of $\mathbf{F}^3$, and with respect to this basis the matrix of T is

$$\begin{pmatrix} 2 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 8 \end{pmatrix}.$$

The converse of 5.44 is not true. For example, the operator T defined on the three-dimensional space $\mathbf{F}^3$ by

$$T(z_1, z_2, z_3) = (4z_1, 4z_2, 5z_3)$$

has only two eigenvalues (4 and 5), but this operator has a diagonal matrix with respect to the standard basis.

In later chapters we will find additional conditions that imply that certain operators are diagonalizable.

EXERCISES 5.C

1 Suppose $T \in \mathcal{L}(V)$ is diagonalizable. Prove that $V = \operatorname{null} T \oplus \operatorname{range} T$.

2 Prove the converse of the statement in the exercise above or give a counterexample to the converse.

3 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) $\operatorname{null} T \cap \operatorname{range} T = \{0\}$.

4 Give an example to show that the exercise above is false without the hypothesis that V is finite-dimensional.

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 V is finite-dimensional, $T \in \mathcal{L}(V)$ has $\dim V$ distinct eigenvalues, and $S \in \mathcal{L}(V)$ has the same eigenvectors as T (not necessarily with the same eigenvalues). Prove that $ST = TS$.

7 Suppose $T \in \mathcal{L}(V)$ has a diagonal matrix A with respect to some basis of V and that $\lambda \in \mathbf{F}$. Prove that λ appears on the diagonal of A precisely $\dim E(\lambda, T)$ times.

8 Suppose $T \in \mathcal{L}(\mathbf{F}^5)$ and $\dim E(8, T) = 4$. Prove that $T - 2I$ or $T - 6I$ is invertible.

9 Suppose $T \in \mathcal{L}(V)$ is invertible. Prove that $E(\lambda, T) = E(\frac{1}{\lambda}, T^{-1})$ for every $\lambda \in \mathbf{F}$ with $\lambda \neq 0$.

10 Suppose that V is finite-dimensional and $T \in \mathcal{L}(V)$. Let $\lambda_1, \dots, \lambda_m$ denote the distinct nonzero eigenvalues of T. Prove that

$$\dim E(\lambda_1, T) + \cdots + \dim E(\lambda_m, T) \leq \dim \operatorname{range} T.$$

11 Verify the assertion in Example 5.40.

12 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 $R = S^{-1}TS$.

13 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 $R = S^{-1}TS$.

14 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 $\mathbf{C}^3$.

15 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 $(x, y, z) \in \mathbf{C}^3$ such that $T(x, y, z) = (17 + 8x, \sqrt{5} + 8y, 2\pi + 8z)$.

16 The ***Fibonacci sequence*** $F_1, F_2, \dots$ is defined by

$$F_1 = 1,\ F_2 = 1, \quad \text{and} \quad F_n = F_{n-2} + F_{n-1} \text{ for } n \geq 3.$$

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 positive 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 part (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 positive integer n.

(e) Use part (d) to conclude that for each positive integer n, the Fibonacci number F_n is the integer that is closest to

$$\frac{1}{\sqrt{5}}\left(\frac{1 + \sqrt{5}}{2}\right)^n.$$