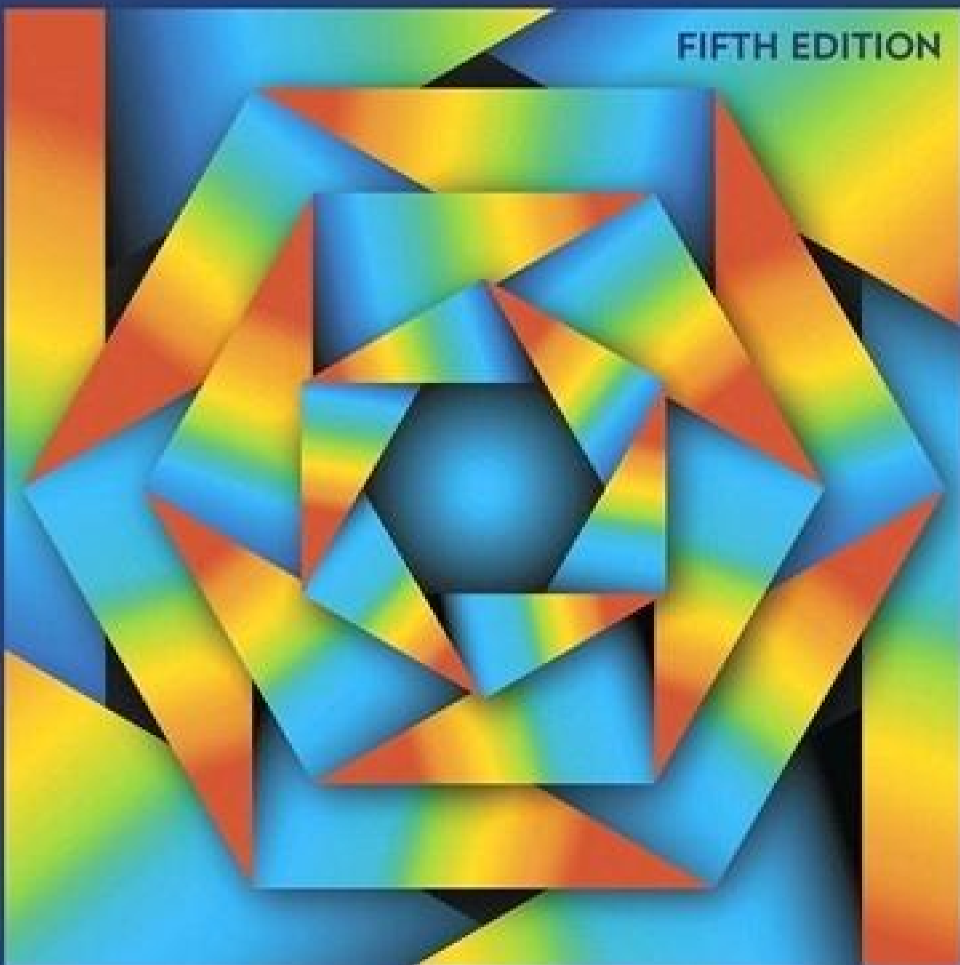
LINEAR ALGEBRA

FIFTH EDITION



Stephen H. **FRIEDBERG**

Arnold J. **INSEL**

Lawrence E. **SPENCE**

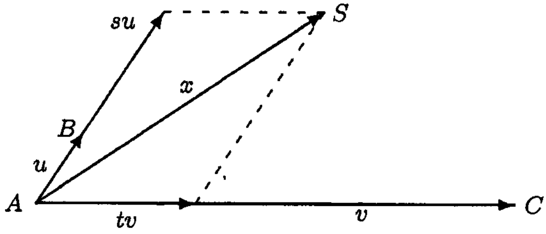


Figure 1.4

locates the endpoint of tv. Moreover, for any real numbers s and t, the vector $su + tv$ lies in the plane containing A, B, and C. It follows that an equation of the plane containing A, B, and C is

$$x = A + su + tv,$$

where s and t are arbitrary real numbers and x denotes an arbitrary point in the plane.

Example 2

Let A, B, and C be the points having coordinates $(1, 0, 2)$, $(-3, -2, 4)$, and $(1, 8, -5)$, respectively. The endpoint of the vector emanating from the origin and having the same length and direction as the vector beginning at A and terminating at B is

$$(-3, -2, 4) - (1, 0, 2) = (-4, -2, 2).$$

Similarly, the endpoint of a vector emanating from the origin and having the same length and direction as the vector beginning at A and terminating at C is $(1, 8, -5) - (1, 0, 2) = (0, 8, -7)$. Hence the equation of the plane containing the three given points is

$$x = (1, 0, 2) + s(-4, -2, 2) + t(0, 8, -7). \qquad \blacklozenge$$

Any mathematical structure possessing the eight properties on page 3 is called a *vector space*. In the next section we formally define a vector space and consider many examples of vector spaces other than the ones mentioned above.

EXERCISES

1. Determine whether the vectors emanating from the origin and terminating at the following pairs of points are parallel.

(a) $(3, 1, 2)$ and $(6, 4, 2)$
(b) $(-3, 1, 7)$ and $(9, -3, -21)$
(c) $(5, -6, 7)$ and $(-5, 6, -7)$
(d) $(2, 0, -5)$ and $(5, 0, -2)$

2. Find the equations of the lines through the following pairs of points in space.

(a) $(3, -2, 4)$ and $(-5, 7, 1)$
(b) $(2, 4, 0)$ and $(-3, -6, 0)$
(c) $(3, 7, 2)$ and $(3, 7, -8)$
(d) $(-2, -1, 5)$ and $(3, 9, 7)$

3. Find the equations of the planes containing the following points in space.

(a) $(2, -5, -1)$, $(0, 4, 6)$, and $(-3, 7, 1)$
(b) $(3, -6, 7)$, $(-2, 0, -4)$, and $(5, -9, -2)$
(c) $(-8, 2, 0)$, $(1, 3, 0)$, and $(6, -5, 0)$
(d) $(1, 1, 1)$, $(5, 5, 5)$, and $(-6, 4, 2)$

4. What are the coordinates of the vector 0 in the Euclidean plane that satisfies property 3 on page 3? Justify your answer.

5. Prove that if the vector x emanates from the origin of the Euclidean plane and terminates at the point with coordinates (a_1, a_2), then the vector tx that emanates from the origin terminates at the point with coordinates (ta_1, ta_2). Visit goo.gl/eYTxuU for a solution.

6. Show that the midpoint of the line segment joining the points (a, b) and (c, d) is $((a + c)/2, (b + d)/2)$.

7. Prove that the diagonals of a parallelogram bisect each other.

1.2 VECTOR SPACES

In Section 1.1, we saw that with the natural definitions of vector addition and scalar multiplication, the vectors in a plane satisfy the eight properties listed on page 3. Many other familiar algebraic systems also permit definitions of addition and scalar multiplication that satisfy the same eight properties. In this section, we introduce some of these systems, but first we formally define this type of algebraic structure.

Definitions. *A **vector space** (or **linear space**) V over a field*[2] *F consists of a set on which two operations (called **addition** and **scalar multiplication**, respectively) are defined so that for each pair of elements x, y,*

[2] Fields are discussed in Appendix C.

(b) The vector $-(ax)$ is the unique element of V such that $ax + [-(ax)] = 0$. Thus if $ax + (-a)x = 0$, Corollary 2 to Theorem 1.1 implies that $(-a)x = -(ax)$. But by (VS 8),

$$ax + (-a)x = [a + (-a)]x = 0x = 0$$

by (a). Consequently $(-a)x = -(ax)$. In particular, $(-1)x = -x$. So, by (VS 6),

$$a(-x) = a[(-1)x] = [a(-1)]x = (-a)x.$$

The proof of (c) is similar to the proof of (a). ∎

EXERCISES

1. Label the following statements as true or false.
   (a) Every vector space contains a zero vector.
   (b) A vector space may have more than one zero vector.
   (c) In any vector space, $ax = bx$ implies that $a = b$.
   (d) In any vector space, $ax = ay$ implies that $x = y$.
   (e) A vector in F^n may be regarded as a matrix in $\mathsf{M}_{n\times 1}(F)$.
   (f) An $m \times n$ matrix has m columns and n rows.
   (g) In $\mathsf{P}(F)$, only polynomials of the same degree may be added.
   (h) If f and g are polynomials of degree n, then $f + g$ is a polynomial of degree n.
   (i) If f is a polynomial of degree n and c is a nonzero scalar, then cf is a polynomial of degree n.
   (j) A nonzero scalar of F may be considered to be a polynomial in $\mathsf{P}(F)$ having degree zero.
   (k) Two functions in $\mathcal{F}(S, F)$ are equal if and only if they have the same value at each element of S.

2. Write the zero vector of $\mathsf{M}_{3\times 4}(F)$.

3. If

$$M = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix},$$

   what are M_{13}, M_{21}, and M_{22}?

4. Perform the indicated operations.
   (a) $\begin{pmatrix} 2 & 5 & -3 \\ 1 & 0 & 7 \end{pmatrix} + \begin{pmatrix} 4 & -2 & 5 \\ -5 & 3 & 2 \end{pmatrix}$
   (b) $\begin{pmatrix} -6 & 4 \\ 3 & -2 \\ 1 & 8 \end{pmatrix} + \begin{pmatrix} 7 & -5 \\ 0 & -3 \\ 2 & 0 \end{pmatrix}$
   (c) $4\begin{pmatrix} 2 & 5 & -3 \\ 1 & 0 & 7 \end{pmatrix}$

(d) $-5\begin{pmatrix} -6 & 4 \\ 3 & -2 \\ 1 & 8 \end{pmatrix}$

(e) $(2x^4 - 7x^3 + 4x + 3) + (8x^3 + 2x^2 - 6x + 7)$
(f) $(-3x^3 + 7x^2 + 8x - 6) + (2x^3 - 8x + 10)$
(g) $5(2x^7 - 6x^4 + 8x^2 - 3x)$
(h) $3(x^5 - 2x^3 + 4x + 2)$

Exercises 5 and 6 show why the definitions of matrix addition and scalar multiplication (as defined in Example 2) are the appropriate ones.

5. Richard Gard ("Effects of Beaver on Trout in Sagehen Creek, California," *J. Wildlife Management*, **25**, 221-242) reports the following number of trout having crossed beaver dams in Sagehen Creek.

Upstream Crossings

	Fall	Spring	Summer
Brook trout	8	3	1
Rainbow trout	3	0	0
Brown trout	3	0	0

Downstream Crossings

	Fall	Spring	Summer
Brook trout	9	1	4
Rainbow trout	3	0	0
Brown trout	1	1	0

Record the upstream and downstream crossings in two 3×3 matrices, and verify that the sum of these matrices gives the total number of crossings (both upstream and downstream) categorized by trout species and season.

6. At the end of May, a furniture store had the following inventory.

	Early American	Spanish	Mediterranean	Danish
Living room suites	4	2	1	3
Bedroom suites	5	1	1	4
Dining room suites	3	1	2	6

Record these data as a 3×4 matrix M. To prepare for its June sale, the store decided to double its inventory on each of the items listed in the preceding table. Assuming that none of the present stock is sold until the additional furniture arrives, verify that the inventory on hand after the order is filled is described by the matrix $2M$. If the inventory at the end of June is described by the matrix

$$A = \begin{pmatrix} 5 & 3 & 1 & 2 \\ 6 & 2 & 1 & 5 \\ 1 & 0 & 3 & 3 \end{pmatrix},$$

interpret $2M - A$. How many suites were sold during the June sale?

7. Let $S = \{0, 1\}$ and $F = R$. In $\mathcal{F}(S, R)$, show that $f = g$ and $f + g = h$, where $f(t) = 2t + 1$, $g(t) = 1 + 4t - 2t^2$, and $h(t) = 5^t + 1$.

8. In any vector space V, show that $(a + b)(x + y) = ax + ay + bx + by$ for any $x, y \in$ V and any $a, b \in F$.

$\boxed{9.}$ Prove Corollaries 1 and 2 of Theorem 1.1 and Theorem 1.2(c). Visit `goo.gl/WFWgzX` for a solution.

10. Let V denote the set of all differentiable real-valued functions defined on the real line. Prove that V is a vector space with the operations of addition and scalar multiplication defined in Example 3.

11. Let V $= \{\mathit{0}\}$ consist of a single vector $\mathit{0}$ and define $\mathit{0} + \mathit{0} = \mathit{0}$ and $c\mathit{0} = \mathit{0}$ for each scalar c in F. Prove that V is a vector space over F. (V is called the **zero vector space.**)

12. A real-valued function f defined on the real line is called an **even function** if $f(-t) = f(t)$ for each real number t. Prove that the set of even functions defined on the real line with the operations of addition and scalar multiplication defined in Example 3 is a vector space.

13. Let V denote the set of ordered pairs of real numbers. If (a_1, a_2) and (b_1, b_2) are elements of V and $c \in R$, define

$$(a_1, a_2) + (b_1, b_2) = (a_1 + b_1, a_2 b_2) \quad \text{and} \quad c(a_1, a_2) = (ca_1, a_2).$$

Is V a vector space over R with these operations? Justify your answer.

14. Let V $= \{(a_1, a_2, \ldots, a_n) \colon a_i \in C \text{ for } i = 1, 2, \ldots, n\}$; so V is a vector space over C by Example 1. Is V a vector space over the field of real numbers with the operations of coordinatewise addition and multiplication?

15. Let $\mathsf{V} = \{(a_1, a_2, \ldots, a_n) : a_i \in R \text{ for } i = 1, 2, \ldots, n\}$; so V is a vector space over R by Example 1. Is V a vector space over the field of complex numbers with the operations of coordinatewise addition and multiplication?

16. Let V denote the set of all $m \times n$ matrices with real entries; so V is a vector space over R by Example 2. Let F be the field of rational numbers. Is V a vector space over F with the usual definitions of matrix addition and scalar multiplication?

17. Let $\mathsf{V} = \{(a_1, a_2) : a_1, a_2 \in F\}$, where F is a field. Define addition of elements of V coordinatewise, and for $c \in F$ and $(a_1, a_2) \in \mathsf{V}$, define

$$c(a_1, a_2) = (a_1, 0).$$

Is V a vector space over F with these operations? Justify your answer.

18. Let $\mathsf{V} = \{(a_1, a_2) : a_1, a_2 \in R\}$. For $(a_1, a_2), (b_1, b_2) \in \mathsf{V}$ and $c \in R$, define

$$(a_1, a_2) + (b_1, b_2) = (a_1 + 2b_1, a_2 + 3b_2) \quad \text{and} \quad c(a_1, a_2) = (ca_1, ca_2).$$

Is V a vector space over R with these operations? Justify your answer.

19. Let $\mathsf{V} = \{(a_1, a_2) : a_1, a_2 \in R\}$. Define addition of elements of V coordinatewise, and for (a_1, a_2) in V and $c \in R$, define

$$c(a_1, a_2) = \begin{cases} (0, 0) & \text{if } c = 0 \\ \left(ca_1, \dfrac{a_2}{c}\right) & \text{if } c \neq 0. \end{cases}$$

Is V a vector space over R with these operations? Justify your answer.

20. Let V denote the set of all real-valued functions f defined on the real line such that $f(1) = 0$. Prove that V is a vector space with the operations of addition and scalar multiplication defined in Example 3.

21. Let V and W be vector spaces over a field F. Let

$$\mathsf{Z} = \{(v, w) : v \in \mathsf{V} \text{ and } w \in \mathsf{W}\}.$$

Prove that Z is a vector space over F with the operations

$$(v_1, w_1) + (v_2, w_2) = (v_1 + v_2, w_1 + w_2) \quad \text{and} \quad c(v_1, w_1) = (cv_1, cw_1).$$

22. How many matrices are there in the vector space $\mathsf{M}_{m \times n}(Z_2)$? (See Appendix C.)

Theorem 1.4. *Any intersection of subspaces of a vector space* V *is a subspace of* V.

Proof. Let $\mathcal{C}$ be a collection of subspaces of V, and let W denote the intersection of the subspaces in $\mathcal{C}$. Since every subspace contains the zero vector, $0 \in$ W. Let $a \in F$ and $x, y \in$ W. Then x and y are contained in each subspace in $\mathcal{C}$. Because each subspace in $\mathcal{C}$ is closed under addition and scalar multiplication, it follows that $x+y$ and ax are contained in each subspace in $\mathcal{C}$. Hence $x+y$ and ax are also contained in W, so that W is a subspace of V by Theorem 1.3. ∎

Having shown that the intersection of subspaces of a vector space V is a subspace of V, it is natural to consider whether or not the union of subspaces of V is a subspace of V. It is easily seen that the union of subspaces must contain the zero vector and be closed under scalar multiplication, but in general the union of subspaces of V need not be closed under addition. In fact, it can be readily shown that the union of two subspaces of V is a subspace of V if and only if one of the subspaces contains the other. (See Exercise 19.) There is, however, a natural way to combine two subspaces W_1 and W_2 to obtain a subspace that contains both W_1 and W_2. As we already have suggested, the key to finding such a subspace is to assure that it must be closed under addition. This idea is explored in Exercise 23.

EXERCISES

1. Label the following statements as true or false.
   (a) If V is a vector space and W is a subset of V that is a vector space, then W is a subspace of V.
   (b) The empty set is a subspace of every vector space.
   (c) If V is a vector space other than the zero vector space, then V contains a subspace W such that W $\neq$ V.
   (d) The intersection of any two subsets of V is a subspace of V.
   (e) An $n \times n$ diagonal matrix can never have more than n nonzero entries.
   (f) The trace of a square matrix is the product of its diagonal entries.
   (g) Let W be the xy-plane in R^3; that is, W $= \{(a_1, a_2, 0)\colon a_1, a_2 \in R\}$. Then W $= \mathsf{R}^2$.

2. Determine the transpose of each of the matrices that follow. In addition, if the matrix is square, compute its trace.

(a) $\begin{pmatrix} -4 & 2 \\ 5 & -1 \end{pmatrix}$ (b) $\begin{pmatrix} 0 & 8 & -6 \\ 3 & 4 & 7 \end{pmatrix}$

(c) $\begin{pmatrix} -3 & 9 \\ 0 & -2 \\ 6 & 1 \end{pmatrix}$ (d) $\begin{pmatrix} 10 & 0 & -8 \\ 2 & -4 & 3 \\ -5 & 7 & 6 \end{pmatrix}$

(e) $\begin{pmatrix} 1 & -1 & 3 & 5 \end{pmatrix}$ (f) $\begin{pmatrix} -2 & 5 & 1 & 4 \\ 7 & 0 & 1 & -6 \end{pmatrix}$

(g) $\begin{pmatrix} 5 \\ 6 \\ 7 \end{pmatrix}$ (h) $\begin{pmatrix} -4 & 0 & 6 \\ 0 & 1 & -3 \\ 6 & -3 & 5 \end{pmatrix}$

3. Prove that $(aA + bB)^t = aA^t + bB^t$ for any $A, B \in \mathsf{M}_{m\times n}(F)$ and any $a, b \in F$.

4. Prove that $(A^t)^t = A$ for each $A \in \mathsf{M}_{m\times n}(F)$.

5. Prove that $A + A^t$ is symmetric for any square matrix A.

6. Prove that $\operatorname{tr}(aA + bB) = a\operatorname{tr}(A) + b\operatorname{tr}(B)$ for any $A, B \in \mathsf{M}_{n\times n}(F)$.

7. Prove that diagonal matrices are symmetric matrices.

8. Determine whether the following sets are subspaces of R^3 under the operations of addition and scalar multiplication defined on R^3. Justify your answers.

(a) $\mathsf{W}_1 = \{(a_1, a_2, a_3) \in \mathsf{R}^3 \colon a_1 = 3a_2 \text{ and } a_3 = -a_2\}$
(b) $\mathsf{W}_2 = \{(a_1, a_2, a_3) \in \mathsf{R}^3 \colon a_1 = a_3 + 2\}$
(c) $\mathsf{W}_3 = \{(a_1, a_2, a_3) \in \mathsf{R}^3 \colon 2a_1 - 7a_2 + a_3 = 0\}$
(d) $\mathsf{W}_4 = \{(a_1, a_2, a_3) \in \mathsf{R}^3 \colon a_1 - 4a_2 - a_3 = 0\}$
(e) $\mathsf{W}_5 = \{(a_1, a_2, a_3) \in \mathsf{R}^3 \colon a_1 + 2a_2 - 3a_3 = 1\}$
(f) $\mathsf{W}_6 = \{(a_1, a_2, a_3) \in \mathsf{R}^3 \colon 5a_1^2 - 3a_2^2 + 6a_3^2 = 0\}$

9. Let W_1, W_3, and W_4 be as in Exercise 8. Describe $\mathsf{W}_1 \cap \mathsf{W}_3$, $\mathsf{W}_1 \cap \mathsf{W}_4$, and $\mathsf{W}_3 \cap \mathsf{W}_4$, and observe that each is a subspace of R^3.

10. Prove that $\mathsf{W}_1 = \{(a_1, a_2, \ldots, a_n) \in \mathsf{F}^n \colon a_1 + a_2 + \cdots + a_n = 0\}$ is a subspace of F^n, but $\mathsf{W}_2 = \{(a_1, a_2, \ldots, a_n) \in \mathsf{F}^n \colon a_1 + a_2 + \cdots + a_n = 1\}$ is not.

11. Is the set $\mathsf{W} = \{f(x) \in \mathsf{P}(F) \colon f(x) = 0 \text{ or } f(x) \text{ has degree } n\}$ a subspace of $\mathsf{P}(F)$ if $n \geq 1$? Justify your answer.

12. Prove that the set of $m \times n$ upper triangular matrices is a subspace of $\mathsf{M}_{m\times n}(F)$.

13. Let S be a nonempty set and F a field. Prove that for any $s_0 \in S$, $\{f \in \mathcal{F}(S, F) \colon f(s_0) = 0\}$, is a subspace of $\mathcal{F}(S, F)$.

14. Let S be a nonempty set and F a field. Let $\mathcal{C}(S,F)$ denote the set of all functions $f \in \mathcal{F}(S,F)$ such that $f(s) = 0$ for all but a finite number of elements of S. Prove that $\mathcal{C}(S,F)$ is a subspace of $\mathcal{F}(S,F)$.

15. Is the set of all differentiable real-valued functions defined on R a subspace of $\mathsf{C}(R)$? Justify your answer.

16. Let $\mathsf{C}^n(R)$ denote the set of all real-valued functions defined on the real line that have a continuous nth derivative. Prove that $\mathsf{C}^n(R)$ is a subspace of $\mathcal{F}(R,R)$.

17. Prove that a subset W of a vector space V is a subspace of V if and only if $\mathsf{W} \neq \varnothing$, and, whenever $a \in F$ and $x, y \in \mathsf{W}$, then $ax \in \mathsf{W}$ and $x + y \in \mathsf{W}$.

18. Prove that a subset W of a vector space V is a subspace of V if and only if $0 \in \mathsf{W}$ and $ax + y \in \mathsf{W}$ whenever $a \in F$ and $x, y \in W$.

19. Let W_1 and W_2 be subspaces of a vector space V. Prove that $\mathsf{W}_1 \cup \mathsf{W}_2$ is a subspace of V if and only if $\mathsf{W}_1 \subseteq \mathsf{W}_2$ or $\mathsf{W}_2 \subseteq \mathsf{W}_1$.

20.† Prove that if W is a subspace of a vector space V and $w_1, w_2, \ldots, w_n$ are in W, then $a_1w_1 + a_2w_2 + \cdots + a_nw_n \in \mathsf{W}$ for any scalars $a_1, a_2, \ldots, a_n$. Visit `goo.gl/KTg35w` for a solution.

21. Let V denote the vector space of sequences in R, as defined in Example 5 of Section 1.2. Show that the set of convergent sequences (a_n) (that is, those for which $\lim_{n\to\infty} a_n$ exists) is a subspace of V.

22. Let F_1 and F_2 be fields. A function $g \in \mathcal{F}(F_1, F_2)$ is called an **even function** if $g(-t) = g(t)$ for each $t \in F_1$ and is called an **odd function** if $g(-t) = -g(t)$ for each $t \in F_1$. Prove that the set of all even functions in $\mathcal{F}(F_1, F_2)$ and the set of all odd functions in $\mathcal{F}(F_1, F_2)$ are subspaces of $\mathcal{F}(F_1, F_2)$.

The following definitions are used in Exercises 23–30.

Definition. *If S_1 and S_2 are nonempty subsets of a vector space V, then the* **sum** *of S_1 and S_2, denoted $S_1 + S_2$, is the set $\{x + y \colon x \in S_1 \text{ and } y \in S_2\}$.*

Definition. *A vector space V is called the* **direct sum** *of W_1 and W_2 if W_1 and W_2 are subspaces of V such that $\mathsf{W}_1 \cap \mathsf{W}_2 = \{0\}$ and $\mathsf{W}_1 + \mathsf{W}_2 = \mathsf{V}$. We denote that V is the direct sum of W_1 and W_2 by writing $\mathsf{V} = \mathsf{W}_1 \oplus \mathsf{W}_2$.*

---

†A dagger means that this exercise is essential for a later section.

23. Let W_1 and W_2 be subspaces of a vector space V.

(a) Prove that $\mathsf{W}_1 + \mathsf{W}_2$ is a subspace of V that contains both W_1 and W_2.

(b) Prove that any subspace of V that contains both W_1 and W_2 must also contain $\mathsf{W}_1 + \mathsf{W}_2$.

24. Show that F^n is the direct sum of the subspaces

$$\mathsf{W}_1 = \{(a_1, a_2, \ldots, a_n) \in \mathsf{F}^n : a_n = 0\}$$

and

$$\mathsf{W}_2 = \{(a_1, a_2, \ldots, a_n) \in \mathsf{F}^n : a_1 = a_2 = \cdots = a_{n-1} = 0\}.$$

25. Let W_1 denote the set of all polynomials $f(x)$ in $\mathsf{P}(F)$ such that in the representation

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0,$$

we have $a_i = 0$ whenever i is even. Likewise let W_2 denote the set of all polynomials $g(x)$ in $\mathsf{P}(F)$ such that in the representation

$$g(x) = b_m x^m + b_{m-1} x^{m-1} + \cdots + b_1 x + b_0,$$

we have $b_i = 0$ whenever i is odd. Prove that $\mathsf{P}(F) = \mathsf{W}_1 \oplus \mathsf{W}_2$.

26. In $\mathsf{M}_{m\times n}(F)$ define $\mathsf{W}_1 = \{A \in \mathsf{M}_{m\times n}(F) : A_{ij} = 0 \text{ whenever } i > j\}$ and $\mathsf{W}_2 = \{A \in \mathsf{M}_{m\times n}(F) : A_{ij} = 0 \text{ whenever } i \le j\}$. ($\mathsf{W}_1$ is the set of all upper triangular matrices as defined on page 19.) Show that $\mathsf{M}_{m\times n}(F) = \mathsf{W}_1 \oplus \mathsf{W}_2$.

27. Let V denote the vector space of all upper triangular $n \times n$ matrices (as defined on page 19), and let W_1 denote the subspace of V consisting of all diagonal matrices. Define $\mathsf{W}_2 = \{A \in \mathsf{V} : A_{ij} = 0 \text{ whenever } i \ge j\}$. Show that $\mathsf{V} = \mathsf{W}_1 \oplus \mathsf{W}_2$.

28. A matrix M is called **skew-symmetric** if $M^t = -M$. Clearly, a skew-symmetric matrix is square. Let F be a field. Prove that the set W_1 of all skew-symmetric $n \times n$ matrices with entries from F is a subspace of $\mathsf{M}_{n\times n}(F)$. Now assume that F is not of characteristic two (see page 549), and let W_2 be the subspace of $\mathsf{M}_{n\times n}(F)$ consisting of all symmetric $n \times n$ matrices. Prove that $\mathsf{M}_{n\times n}(F) = \mathsf{W}_1 \oplus \mathsf{W}_2$.

29. Let F be a field that is not of characteristic two. Define

$$\mathsf{W}_1 = \{A \in \mathsf{M}_{n\times n}(F) : A_{ij} = 0 \text{ whenever } i \le j\}$$

and W_2 to be the set of all symmetric $n \times n$ matrices with entries from F. Both W_1 and W_2 are subspaces of $\mathsf{M}_{n\times n}(F)$. Prove that $\mathsf{M}_{n\times n}(F) = \mathsf{W}_1 \oplus \mathsf{W}_2$. Compare this exercise with Exercise 28.

30. Let W_1 and W_2 be subspaces of a vector space V. Prove that V is the direct sum of W_1 and W_2 if and only if each vector in V can be *uniquely* written as $x_1 + x_2$, where $x_1 \in \mathsf{W}_1$ and $x_2 \in \mathsf{W}_2$.

31. Let W be a subspace of a vector space V over a field F. For any $v \in \mathsf{V}$ the set $\{v\} + \mathsf{W} = \{v + w : w \in \mathsf{W}\}$ is called the **coset of W containing** v. It is customary to denote this coset by $v + \mathsf{W}$ rather than $\{v\} + \mathsf{W}$.

(a) Prove that $v + \mathsf{W}$ is a subspace of V if and only if $v \in \mathsf{W}$.
(b) Prove that $v_1 + \mathsf{W} = v_2 + \mathsf{W}$ if and only if $v_1 - v_2 \in \mathsf{W}$.

Addition and scalar multiplication by scalars of F can be defined in the collection $S = \{v + \mathsf{W} : v \in \mathsf{V}\}$ of all cosets of W as follows:

$$(v_1 + \mathsf{W}) + (v_2 + \mathsf{W}) = (v_1 + v_2) + \mathsf{W}$$

for all $v_1, v_2 \in \mathsf{V}$ and

$$a(v + \mathsf{W}) = av + \mathsf{W}$$

for all $v \in \mathsf{V}$ and $a \in F$.

(c) Prove that the preceding operations are well defined; that is, show that if $v_1 + \mathsf{W} = v_1' + \mathsf{W}$ and $v_2 + \mathsf{W} = v_2' + \mathsf{W}$, then

$$(v_1 + \mathsf{W}) + (v_2 + \mathsf{W}) = (v_1' + \mathsf{W}) + (v_2' + \mathsf{W})$$

and

$$a(v_1 + \mathsf{W}) = a(v_1' + \mathsf{W})$$

for all $a \in F$.

(d) Prove that the set S is a vector space with the operations defined in (c). This vector space is called the **quotient space V modulo W** and is denoted by V/W.

1.4 LINEAR COMBINATIONS AND SYSTEMS OF LINEAR EQUATIONS

In Section 1.1, it was shown that the equation of the plane through three noncollinear points A, B, and C in space is $x = A + su + tv$, where u and v denote the vectors beginning at A and ending at B and C, respectively, and s and t denote arbitrary real numbers. An important special case occurs when A is the origin. In this case, the equation of the plane simplifies to $x = su + tv$, and the set of all points in this plane is a subspace of R^3. (This is proved as Theorem 1.5.) Expressions of the form $su + tv$, where s and t are scalars and u and v are vectors, play a central role in the theory of vector spaces. The appropriate generalization of such expressions is presented in the following definitions.

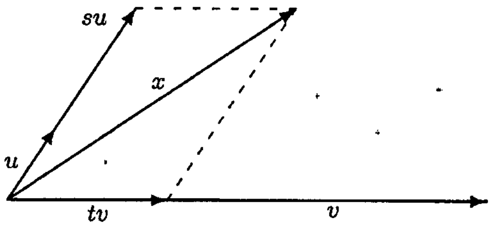


Figure 1.5

a linear combination of $u, v \in \mathsf{R}^3$ if and only if x lies in the plane containing u and v. (See Figure 1.5.)

Usually there are many different subsets that generate a subspace W. (See Exercise 13.) It is natural to seek a subset of W that generates W and is as small as possible. In the next section we explore the circumstances under which a vector can be removed from a generating set to obtain a smaller generating set.

EXERCISES

1. Label the following statements as true or false.

(a) The zero vector is a linear combination of any nonempty set of vectors.
(b) The span of $\varnothing$ is $\varnothing$.
(c) If S is a subset of a vector space V, then span(S) equals the intersection of all subspaces of V that contain S.
(d) In solving a system of linear equations, it is permissible to multiply an equation by any constant.
(e) In solving a system of linear equations, it is permissible to add any multiple of one equation to another.
(f) Every system of linear equations has a solution.

2. Solve the following systems of linear equations by the method introduced in this section.

(a)
$$\begin{aligned} 2x_1 - 2x_2 - 3x_3 \phantom{{}+5x_4} &= -2 \\ 3x_1 - 3x_2 - 2x_3 + 5x_4 &= 7 \\ x_1 - x_2 - 2x_3 - x_4 &= -3 \end{aligned}$$

(b)
$$\begin{aligned} 3x_1 - 7x_2 + 4x_3 &= 10 \\ x_1 - 2x_2 + x_3 &= 3 \\ 2x_1 - x_2 - 2x_3 &= 6 \end{aligned}$$

(c)
$$\begin{aligned} x_1 + 2x_2 - x_3 + x_4 &= 5 \\ x_1 + 4x_2 - 3x_3 - 3x_4 &= 6 \\ 2x_1 + 3x_2 - x_3 + 4x_4 &= 8 \end{aligned}$$

(d) $\begin{aligned} x_1 + 2x_2 + 2x_3 &= 2 \\ x_1 + 8x_3 + 5x_4 &= -6 \\ x_1 + x_2 + 5x_3 + 5x_4 &= 3 \end{aligned}$

(e) $\begin{aligned} x_1 + 2x_2 - 4x_3 - x_4 + x_5 &= 7 \\ -x_1 + 10x_3 - 3x_4 - 4x_5 &= -16 \\ 2x_1 + 5x_2 - 5x_3 - 4x_4 - x_5 &= 2 \\ 4x_1 + 11x_2 - 7x_3 - 10x_4 - 2x_5 &= 7 \end{aligned}$

(f) $\begin{aligned} x_1 + 2x_2 + 6x_3 &= -1 \\ 2x_1 + x_2 + x_3 &= 8 \\ 3x_1 + x_2 - x_3 &= 15 \\ x_1 + 3x_2 + 10x_3 &= -5 \end{aligned}$

3. For each of the following lists of vectors in R^3, determine whether the first vector can be expressed as a linear combination of the other two.

(a) $(-2,0,3), (1,3,0), (2,4,-1)$
(b) $(1,2,-3), (-3,2,1), (2,-1,-1)$
(c) $(3,4,1), (1,-2,1), (-2,-1,1)$
(d) $(2,-1,0), (1,2,-3), (1,-3,2)$
(e) $(5,1,-5), (1,-2,-3), (-2,3,-4)$
(f) $(-2,2,2), (1,2,-1), (-3,-3,3)$

4. For each list of polynomials in $\mathsf{P}_3(R)$, determine whether the first polynomial can be expressed as a linear combination of the other two.

(a) $x^3 - 3x + 5, x^3 + 2x^2 - x + 1, x^3 + 3x^2 - 1$
(b) $4x^3 + 2x^2 - 6, x^3 - 2x^2 + 4x + 1, 3x^3 - 6x^2 + x + 4$
(c) $-2x^3 - 11x^2 + 3x + 2, x^3 - 2x^2 + 3x - 1, 2x^3 + x^2 + 3x - 2$
(d) $x^3 + x^2 + 2x + 13, 2x^3 - 3x^2 + 4x + 1, x^3 - x^2 + 2x + 3$
(e) $x^3 - 8x^2 + 4x, x^3 - 2x^2 + 3x - 1, x^3 - 2x + 3$
(f) $6x^3 - 3x^2 + x + 2, x^3 - x^2 + 2x + 3, 2x^3 - 3x + 1$

5. In each part, determine whether the given vector is in the span of S.

(a) $(2,-1,1), \quad S = \{(1,0,2), (-1,1,1)\}$
(b) $(-1,2,1), \quad S = \{(1,0,2), (-1,1,1)\}$
(c) $(-1,1,1,2), \quad S = \{(1,0,1,-1), (0,1,1,1)\}$
(d) $(2,-1,1,-3), \quad S = \{(1,0,1,-1), (0,1,1,1)\}$
(e) $-x^3 + 2x^2 + 3x + 3, \quad S = \{x^3 + x^2 + x + 1, x^2 + x + 1, x + 1\}$
(f) $2x^3 - x^2 + x + 3, \quad S = \{x^3 + x^2 + x + 1, x^2 + x + 1, x + 1\}$
(g) $\begin{pmatrix} 1 & 2 \\ -3 & 4 \end{pmatrix}, \quad S = \left\{ \begin{pmatrix} 1 & 0 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \right\}$
(h) $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad S = \left\{ \begin{pmatrix} 1 & 0 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \right\}$

6. Show that the vectors $(1,1,0)$, $(1,0,1)$, and $(0,1,1)$ generate F^3.

7. In F^n, let e_j denote the vector whose jth coordinate is 1 and whose other coordinates are 0. Prove that $\{e_1, e_2, \ldots, e_n\}$ generates F^n.

8. Show that $\mathsf{P}_n(F)$ is generated by $\{1, x, \ldots, x^n\}$.

9. Show that the matrices

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad \text{and} \quad \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

generate $\mathsf{M}_{2\times 2}(F)$.

10. Show that if

$$M_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad M_2 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad \text{and} \quad M_3 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

then the span of $\{M_1, M_2, M_3\}$ is the set of all symmetric 2×2 matrices.

11.† Prove that $\text{span}(\{x\}) = \{ax \colon a \in F\}$ for any vector x in a vector space. Interpret this result geometrically in R^3.

12. Show that a subset W of a vector space V is a subspace of V if and only if $\text{span}(\mathsf{W}) = \mathsf{W}$.

13.† Show that if S_1 and S_2 are subsets of a vector space V such that $S_1 \subseteq S_2$, then $\text{span}(S_1) \subseteq \text{span}(S_2)$. In particular, if $S_1 \subseteq S_2$ and $\text{span}(S_1) = \mathsf{V}$, deduce that $\text{span}(S_2) = \mathsf{V}$. Visit `goo.gl/Fi8Epr` for a solution.

14. Show that if S_1 and S_2 are arbitrary subsets of a vector space V, then $\text{span}(S_1 \cup S_2) = \text{span}(S_1) + \text{span}(S_2)$. (The sum of two subsets is defined in the exercises of Section 1.3.)

15. Let S_1 and S_2 be subsets of a vector space V. Prove that $\text{span}(S_1 \cap S_2) \subseteq \text{span}(S_1) \cap \text{span}(S_2)$. Give an example in which $\text{span}(S_1 \cap S_2)$ and $\text{span}(S_1) \cap \text{span}(S_2)$ are equal and one in which they are unequal.

16. Let V be a vector space and S a subset of V with the property that whenever $v_1, v_2, \ldots, v_n \in S$ and $a_1v_1 + a_2v_2 + \cdots + a_nv_n = \mathit{0}$, then $a_1 = a_2 = \cdots = a_n = 0$. Prove that every vector in the span of S can be *uniquely* written as a linear combination of vectors of S.

17. Let W be a subspace of a vector space V. Under what conditions are there only a finite number of distinct subsets S of W such that S generates W?

Note that $v \neq v_i$ for $i = 1, 2, \ldots, m$ because $v \notin S$. Hence the coefficient of v in this linear combination is nonzero, and so the set $\{v_1, v_2, \ldots, v_m, v\}$ is linearly dependent. Thus $S \cup \{v\}$ is linearly dependent by Theorem 1.6. ∎

Linearly independent generating sets are investigated in detail in Section 1.6.

EXERCISES

1. Label the following statements as true or false.
   (a) If S is a linearly dependent set, then each vector in S is a linear combination of other vectors in S.
   (b) Any set containing the zero vector is linearly dependent.
   (c) The empty set is linearly dependent.
   (d) Subsets of linearly dependent sets are linearly dependent.
   (e) Subsets of linearly independent sets are linearly independent.
   (f) If $a_1x_1 + a_2x_2 + \cdots + a_nx_n = \mathit{0}$ and $x_1, x_2, \ldots, x_n$ are linearly independent, then all the scalars a_i are zero.

2.[3] Determine whether the following sets are linearly dependent or linearly independent.
   (a) $\left\{\begin{pmatrix} 1 & -3 \\ -2 & 4 \end{pmatrix}, \begin{pmatrix} -2 & 6 \\ 4 & -8 \end{pmatrix}\right\}$ in $\mathsf{M}_{2\times 2}(R)$

   (b) $\left\{\begin{pmatrix} 1 & -2 \\ -1 & 4 \end{pmatrix}, \begin{pmatrix} -1 & 1 \\ 2 & -4 \end{pmatrix}\right\}$ in $\mathsf{M}_{2\times 2}(R)$

   (c) $\{x^3 + 2x^2, -x^2 + 3x + 1, x^3 - x^2 + 2x - 1\}$ in $\mathsf{P}_3(R)$
   (d) $\{x^3 - x, 2x^2 + 4, -2x^3 + 3x^2 + 2x + 6\}$ in $\mathsf{P}_3(R)$
   (e) $\{(1, -1, 2), (1, -2, 1), (1, 1, 4)\}$ in R^3
   (f) $\{(1, -1, 2), (2, 0, 1), (-1, 2, -1)\}$ in R^3

   (g) $\left\{\begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} -1 & 2 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 2 & 1 \\ -4 & 4 \end{pmatrix}\right\}$ in $\mathsf{M}_{2\times 2}(R)$

   (h) $\left\{\begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} -1 & 2 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 2 & 1 \\ 2 & -2 \end{pmatrix}\right\}$ in $\mathsf{M}_{2\times 2}(R)$

   (i) $\{x^4 - x^3 + 5x^2 - 8x + 6, -x^4 + x^3 - 5x^2 + 5x - 3,$
   $x^4 + 3x^2 - 3x + 5, 2x^4 + 3x^3 + 4x^2 - x + 1, x^3 - x + 2\}$ in $\mathsf{P}_4(R)$
   (j) $\{x^4 - x^3 + 5x^2 - 8x + 6, -x^4 + x^3 - 5x^2 + 5x - 3,$
   $x^4 + 3x^2 - 3x + 5, 2x^4 + x^3 + 4x^2 + 8x\}$ in $\mathsf{P}_4(R)$

[3]The computations in Exercise 2(g), (h), (i), and (j) are tedious unless technology is used.

3. In $\mathsf{M}_{3\times 2}(F)$, prove that the set

$$\left\{ \begin{pmatrix} 1 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{pmatrix} \right\}$$

is linearly dependent.

4. In F^n, let e_j denote the vector whose jth coordinate is 1 and whose other coordinates are 0. Prove that $\{e_1, e_2, \ldots, e_n\}$ is linearly independent.

5. Show that the set $\{1, x, x^2, \ldots, x^n\}$ is linearly independent in $\mathsf{P}_n(F)$.

6. In $\mathsf{M}_{m\times n}(F)$, let E^{ij} denote the matrix whose only nonzero entry is 1 in the ith row and jth column. Prove that $\{E^{ij} : 1 \le i \le m,\ 1 \le j \le n\}$ is linearly independent.

7. Recall from Example 3 in Section 1.3 that the set of diagonal matrices in $\mathsf{M}_{2\times 2}(F)$ is a subspace. Find a linearly independent set that generates this subspace.

8. Let $S = \{(1,1,0), (1,0,1), (0,1,1)\}$ be a subset of the vector space F^3.
   (a) Prove that if $F = R$, then S is linearly independent.
   (b) Prove that if F has characteristic two, then S is linearly dependent.

9.† Let u and v be distinct vectors in a vector space V. Show that $\{u, v\}$ is linearly dependent if and only if u or v is a multiple of the other.

10. Give an example of three linearly dependent vectors in R^3 such that none of the three is a multiple of another.

11. Let $S = \{u_1, u_2, \ldots, u_n\}$ be a linearly independent subset of a vector space V over the field Z_2. How many vectors are there in span(S)? Justify your answer.

12. Prove Theorem 1.6 and its corollary.

13. Let V be a vector space over a field of characteristic not equal to two.
   (a) Let u and v be distinct vectors in V. Prove that $\{u, v\}$ is linearly independent if and only if $\{u+v, u-v\}$ is linearly independent.
   (b) Let u, v, and w be distinct vectors in V. Prove that $\{u, v, w\}$ is linearly independent if and only if $\{u+v, u+w, v+w\}$ is linearly independent.

14. Prove that a set S is linearly dependent if and only if $S = \{0\}$ or there exist distinct vectors $v, u_1, u_2, \ldots, u_n$ in S such that v is a linear combination of $u_1, u_2, \ldots, u_n$.

15. Let $S = \{u_1, u_2, \ldots, u_n\}$ be a finite set of vectors. Prove that S is linearly dependent if and only if $u_1 = 0$ or $u_{k+1} \in \operatorname{span}(\{u_1, u_2, \ldots, u_k\})$ for some k $(1 \leq k < n)$.

16. Prove that a set S of vectors is linearly independent if and only if each finite subset of S is linearly independent.

17. Let M be a square upper triangular matrix (as defined on page 19 of Section 1.3) with nonzero diagonal entries. Prove that the columns of M are linearly independent.

18. Let S be a set of nonzero polynomials in $\mathsf{P}(F)$ such that no two have the same degree. Prove that S is linearly independent.

19. Prove that if $\{A_1, A_2, \ldots, A_k\}$ is a linearly independent subset of $\mathsf{M}_{n\times n}(F)$, then $\{A_1^t, A_2^t, \ldots, A_k^t\}$ is also linearly independent.

20. Let $f, g \in \mathcal{F}(R, R)$ be the functions defined by $f(t) = e^{rt}$ and $g(t) = e^{st}$, where $r \neq s$. Prove that f and g are linearly independent in $\mathcal{F}(R, R)$.

$\boxed{\textbf{21.}}$ Let S_1 and S_2 be disjoint linearly independent subsets of V. Prove that $S_1 \cup S_2$ is linearly dependent if and only if $\operatorname{span}(S_1) \cap \operatorname{span}(S_2) \neq \{0\}$. Visit `goo.gl/Fi8Epr` for a solution.

1.6 BASES AND DIMENSION

We saw in Section 1.5 that if S is a generating set for a subspace W and no proper subset of S is a generating set for W, then S must be linearly independent. A linearly independent generating set for W possesses a very useful property—every vector in W can be expressed in one and only one way as a linear combination of the vectors in the set. (This property is proved below in Theorem 1.8.) It is this property that makes linearly independent generating sets the building blocks of vector spaces.

Definition. *A **basis** β for a vector space V is a linearly independent subset of V that generates V. If β is a basis for V, we also say that the vectors of β form a basis for V.*

Example 1

Recalling that $\operatorname{span}(\varnothing) = \{0\}$ and $\varnothing$ is linearly independent, we see that $\varnothing$ is a basis for the zero vector space. ♦

Example 2

In F^n, let $e_1 = (1, 0, 0, \ldots, 0), e_2 = (0, 1, 0, \ldots, 0), \ldots, e_n = (0, 0, \ldots, 0, 1)$; $\{e_1, e_2, \ldots, e_n\}$ is readily seen to be a basis for F^n and is called the **standard basis** for F^n. ♦

is the unique polynomial in $\mathsf{P}_n(F)$ such that $g(c_j) = b_j$. Thus we have found the unique polynomial of degree not exceeding n that has specified values b_j at given points c_j in its domain $(j = 0, 1, \ldots, n)$. For example, let us construct the real polynomial g of degree at most 2 whose graph contains the points $(1, 8), (2, 5)$, and $(3, -4)$. (Thus, in the notation above, $c_0 = 1$, $c_1 = 2$, $c_2 = 3$, $b_0 = 8$, $b_1 = 5$, and $b_2 = -4$.) The Lagrange polynomials associated with c_0, c_1, and c_2 are

$$f_0(x) = \frac{(x-2)(x-3)}{(1-2)(1-3)} = \frac{1}{2}(x^2 - 5x + 6),$$

$$f_1(x) = \frac{(x-1)(x-3)}{(2-1)(2-3)} = -1(x^2 - 4x + 3),$$

and

$$f_2(x) = \frac{(x-1)(x-2)}{(3-1)(3-2)} = \frac{1}{2}(x^2 - 3x + 2).$$

Hence the desired polynomial is

$$\begin{aligned} g(x) &= \sum_{i=0}^{2} b_i f_i(x) = 8f_0(x) + 5f_1(x) - 4f_2(x) \\ &= 4(x^2 - 5x + 6) - 5(x^2 - 4x + 3) - 2(x^2 - 3x + 2) \\ &= -3x^2 + 6x + 5. \end{aligned}$$

An important consequence of the Lagrange interpolation formula is the following result: If $f \in \mathsf{P}_n(F)$ and $f(c_i) = 0$ for $n+1$ distinct scalars $c_0, c_1, \ldots, c_n$ in F, then f is the zero function.

EXERCISES

1. Label the following statements as true or false.
   (a) The zero vector space has no basis.
   (b) Every vector space that is generated by a finite set has a basis.
   (c) Every vector space has a finite basis.
   (d) A vector space cannot have more than one basis.
   (e) If a vector space has a finite basis, then the number of vectors in every basis is the same.
   (f) The dimension of $\mathsf{P}_n(F)$ is n.
   (g) The dimension of $\mathsf{M}_{m \times n}(F)$ is $m + n$.
   (h) Suppose that V is a finite-dimensional vector space, that S_1 is a linearly independent subset of V, and that S_2 is a subset of V that generates V. Then S_1 cannot contain more vectors than S_2.

(i) If S generates the vector space V, then every vector in V can be written as a linear combination of vectors in S in only one way.
(j) Every subspace of a finite-dimensional space is finite-dimensional.
(k) If V is a vector space having dimension n, then V has exactly one subspace with dimension 0 and exactly one subspace with dimension n.
(l) If V is a vector space having dimension n, and if S is a subset of V with n vectors, then S is linearly independent if and only if S spans V.

2. Determine which of the following sets are bases for R^3.

(a) $\{(1,0,-1),(2,5,1),(0,-4,3)\}$
(b) $\{(2,-4,1),(0,3,-1),(6,0,-1)\}$
(c) $\{(1,2,-1),(1,0,2),(2,1,1)\}$
(d) $\{(-1,3,1),(2,-4,-3),(-3,8,2)\}$
(e) $\{(1,-3,-2),(-3,1,3),(-2,-10,-2)\}$

3. Determine which of the following sets are bases for $\mathsf{P}_2(R)$.

(a) $\{-1-x+2x^2, 2+x-2x^2, 1-2x+4x^2\}$
(b) $\{1+2x+x^2, 3+x^2, x+x^2\}$
(c) $\{1-2x-2x^2, -2+3x-x^2, 1-x+6x^2\}$
(d) $\{-1+2x+4x^2, 3-4x-10x^2, -2-5x-6x^2\}$
(e) $\{1+2x-x^2, 4-2x+x^2, -1+18x-9x^2\}$

4. Do the polynomials x^3-2x^2+1, $4x^2-x+3$, and $3x-2$ generate $\mathsf{P}_3(R)$? Justify your answer.

5. Is $\{(1,4,-6),(1,5,8),(2,1,1),(0,1,0)\}$ a linearly independent subset of R^3? Justify your answer.

6. Give three different bases for F^2 and for $\mathsf{M}_{2\times 2}(F)$.

7. The vectors $u_1=(2,-3,1)$, $u_2=(1,4,-2)$, $u_3=(-8,12,-4)$, $u_4=(1,37,-17)$, and $u_5=(-3,-5,8)$ generate R^3. Find a subset of the set $\{u_1,u_2,u_3,u_4,u_5\}$ that is a basis for R^3.

8. Let W denote the subspace of R^5 consisting of all the vectors having coordinates that sum to zero. The vectors

$$\begin{aligned} u_1 &= (2,-3,4,-5,2), & u_2 &= (-6,9,-12,15,-6),\\ u_3 &= (3,-2,7,-9,1), & u_4 &= (2,-8,2,-2,6),\\ u_5 &= (-1,1,2,1,-3), & u_6 &= (0,-3,-18,9,12),\\ u_7 &= (1,0,-2,3,-2), & u_8 &= (2,-1,1,-9,7) \end{aligned}$$

generate W. Find a subset of the set $\{u_1,u_2,\ldots,u_8\}$ that is a basis for W.

9. The vectors $u_1 = (1,1,1,1)$, $u_2 = (0,1,1,1)$, $u_3 = (0,0,1,1)$, and $u_4 = (0,0,0,1)$ form a basis for F^4. Find the unique representation of an arbitrary vector (a_1, a_2, a_3, a_4) in F^4 as a linear combination of u_1, u_2, u_3, and u_4.

10. In each part, use the Lagrange interpolation formula to construct the polynomial of smallest degree whose graph contains the following points.
    (a) $(-2,-6)$, $(-1,5)$, $(1,3)$
    (b) $(-4,24)$, $(1,9)$, $(3,3)$
    (c) $(-2,3)$, $(-1,-6)$, $(1,0)$, $(3,-2)$
    (d) $(-3,-30)$, $(-2,7)$, $(0,15)$, $(1,10)$

11. Let u and v be distinct vectors of a vector space V. Show that if $\{u, v\}$ is a basis for V and a and b are nonzero scalars, then both $\{u+v, au\}$ and $\{au, bv\}$ are also bases for V.

12. Let u, v, and w be distinct vectors of a vector space V. Show that if $\{u, v, w\}$ is a basis for V, then $\{u+v+w, v+w, w\}$ is also a basis for V.

13. The set of solutions to the system of linear equations

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ 2x_1 - 3x_2 + x_3 &= 0 \end{aligned}$$

is a subspace of R^3. Find a basis for this subspace.

14. Find bases for the following subspaces of F^5:

$$\mathsf{W}_1 = \{(a_1, a_2, a_3, a_4, a_5) \in \mathsf{F}^5 \colon a_1 - a_3 - a_4 = 0\}$$

and

$$\mathsf{W}_2 = \{(a_1, a_2, a_3, a_4, a_5) \in \mathsf{F}^5 \colon a_2 = a_3 = a_4 \text{ and } a_1 + a_5 = 0\}.$$

What are the dimensions of W_1 and W_2?

15. The set of all $n \times n$ matrices having trace equal to zero is a subspace W of $\mathsf{M}_{n\times n}(F)$ (see Example 4 of Section 1.3). Find a basis for W. What is the dimension of W?

16. The set of all upper triangular $n \times n$ matrices is a subspace W of $\mathsf{M}_{n\times n}(F)$ (see Exercise 12 of Section 1.3). Find a basis for W. What is the dimension of W?

17. The set of all skew-symmetric $n \times n$ matrices is a subspace W of $\mathsf{M}_{n\times n}(F)$ (see Exercise 28 of Section 1.3). Find a basis for W. What is the dimension of W?

18. Let V denote the vector space of all sequences in F, as defined in Example 5 of Section 1.2. Find a basis for the subspace W of V consisting of the sequences (a_n) that have only a finite number of nonzero terms a_n. Justify your answer.

19. Complete the proof of Theorem 1.8.

20.† Let V be a vector space having dimension n, and let S be a subset of V that generates V.

(a) Prove that there is a subset of S that is a basis for V. (Be careful not to assume that S is finite.)

(b) Prove that S contains at least n vectors.

Visit goo.gl/wE2wwA for a solution.

21. Prove that a vector space is infinite-dimensional if and only if it contains an infinite linearly independent subset.

22. Let W_1 and W_2 be subspaces of a finite-dimensional vector space V. Determine necessary and sufficient conditions on W_1 and W_2 so that $\dim(W_1 \cap W_2) = \dim(W_1)$.

23. Let $v_1, v_2, \ldots, v_k, v$ be vectors in a vector space V, and define $W_1 = \mathrm{span}(\{v_1, v_2, \ldots, v_k\})$, and $W_2 = \mathrm{span}(\{v_1, v_2, \ldots, v_k, v\})$.

(a) Find necessary and sufficient conditions on v such that $\dim(W_1) = \dim(W_2)$.

(b) State and prove a relationship involving $\dim(W_1)$ and $\dim(W_2)$ in the case that $\dim(W_1) \neq \dim(W_2)$.

24. Let $f(x)$ be a polynomial of degree n in $P_n(R)$. Prove that for any $g(x) \in P_n(R)$ there exist scalars $c_0, c_1, \ldots, c_n$ such that

$$g(x) = c_0 f(x) + c_1 f'(x) + c_2 f''(x) + \cdots + c_n f^{(n)}(x),$$

where $f^{(n)}(x)$ denotes the nth derivative of $f(x)$.

25. Let V, W, and Z be as in Exercise 21 of Section 1.2. If V and W are vector spaces over F of dimensions m and n, determine the dimension of Z.

26. For a fixed $a \in R$, determine the dimension of the subspace of $P_n(R)$ defined by $\{f \in P_n(R) \colon f(a) = 0\}$.

27. Let W_1 and W_2 be the subspaces of $P(F)$ defined in Exercise 25 in Section 1.3. Determine the dimensions of the subspaces $W_1 \cap P_n(F)$ and $W_2 \cap P_n(F)$.

28. Let V be a finite-dimensional vector space over C with dimension n. Prove that if V is now regarded as a vector space over R, then $\dim V = 2n$. (See Examples 11 and 12.)

Exercises 29–34 require knowledge of the sum and direct sum of subspaces, as defined in the exercises of Section 1.3.

29. (a) Prove that if W_1 and W_2 are finite-dimensional subspaces of a vector space V, then the subspace $W_1 + W_2$ is finite-dimensional, and $\dim(W_1 + W_2) = \dim(W_1) + \dim(W_2) - \dim(W_1 \cap W_2)$. *Hint:* Start with a basis $\{u_1, u_2, \ldots, u_k\}$ for $W_1 \cap W_2$ and extend this set to a basis $\{u_1, u_2, \ldots, u_k, v_1, v_2, \ldots, v_m\}$ for W_1 and to a basis $\{u_1, u_2, \ldots, u_k, w_1, w_2, \ldots, w_p\}$ for W_2.

(b) Let W_1 and W_2 be finite-dimensional subspaces of a vector space V, and let $V = W_1 + W_2$. Deduce that V is the direct sum of W_1 and W_2 if and only if $\dim(V) = \dim(W_1) + \dim(W_2)$.

30. Let

$$V = M_{2\times 2}(F), \quad W_1 = \left\{ \begin{pmatrix} a & b \\ c & a \end{pmatrix} \in V \colon a, b, c \in F \right\},$$

and

$$W_2 = \left\{ \begin{pmatrix} 0 & a \\ -a & b \end{pmatrix} \in V \colon a, b \in F \right\}.$$

Prove that W_1 and W_2 are subspaces of V, and find the dimensions of W_1, W_2, $W_1 + W_2$, and $W_1 \cap W_2$.

31. Let W_1 and W_2 be subspaces of a vector space V having dimensions m and n, respectively, where $m \geq n$.

(a) Prove that $\dim(W_1 \cap W_2) \leq n$.
(b) Prove that $\dim(W_1 + W_2) \leq m + n$.

32. Find examples of subspaces W_1 and W_2 of R^3 such that $\dim(W_1) > \dim(W_2) > 0$ and

(a) $\dim(W_1 \cap W_2) = \dim(W_2)$;
(b) $\dim(W_1 + W_2) = \dim(W_1) + \dim(W_2)$;
(c) $\dim(W_1 + W_2) < \dim(W_1) + \dim(W_2)$.

33. (a) Let W_1 and W_2 be subspaces of a vector space V such that $V = W_1 \oplus W_2$. If β_1 and β_2 are bases for W_1 and W_2, respectively, show that $\beta_1 \cap \beta_2 = \varnothing$ and $\beta_1 \cup \beta_2$ is a basis for V.

(b) Conversely, let β_1 and β_2 be disjoint bases for subspaces W_1 and W_2, respectively, of a vector space V. Prove that if $\beta_1 \cup \beta_2$ is a basis for V, then $V = W_1 \oplus W_2$.

34. (a) Prove that if W_1 is any subspace of a finite-dimensional vector space V, then there exists a subspace W_2 of V such that $\mathsf{V} = \mathsf{W}_1 \oplus \mathsf{W}_2$.
(b) Let $\mathsf{V} = \mathsf{R}^2$ and $\mathsf{W}_1 = \{(a_1, 0) : a_1 \in R\}$. Give examples of two different subspaces W_2 and W_2' such that $\mathsf{V} = \mathsf{W}_1 \oplus \mathsf{W}_2$ and $\mathsf{V} = \mathsf{W}_1 \oplus \mathsf{W}_2'$.

The following exercise requires familiarity with Exercise 31 of Section 1.3.

35. Let W be a subspace of a finite-dimensional vector space V, and consider the basis $\{u_1, u_2, \ldots, u_k\}$ for W. Let $\{u_1, u_2, \ldots, u_k, u_{k+1}, \ldots, u_n\}$ be an extension of this basis to a basis for V.
(a) Prove that $\{u_{k+1} + \mathsf{W}, u_{k+2} + \mathsf{W}, \ldots, u_n + \mathsf{W}\}$ is a basis for V/W.
(b) Derive a formula relating $\dim(\mathsf{V})$, $\dim(\mathsf{W})$, and $\dim(\mathsf{V}/\mathsf{W})$.

1.7* MAXIMAL LINEARLY INDEPENDENT SUBSETS

In this section, several significant results from Section 1.6 are extended to infinite-dimensional vector spaces. Our principal goal here is to prove that every vector space has a basis. This result is important in the study of infinite-dimensional vector spaces because it is often difficult to construct an explicit basis for such a space. Consider, for example, the vector space of real numbers over the field of rational numbers. There is no obvious way to construct a basis for this space, and yet it follows from the results of this section that such a basis does exist.

The difficulty that arises in extending the theorems of the preceding section to infinite-dimensional vector spaces is that the principle of mathematical induction, which played a crucial role in many of the proofs of Section 1.6, is no longer adequate. Instead, an alternate result called the *Hausdorff maximal principle* is needed. Before stating this principle, we need to introduce some terminology.

Definition. *Let $\mathcal{F}$ be a family of sets. A member M of $\mathcal{F}$ is called* **maximal** *(with respect to set inclusion) if M is contained in no member of $\mathcal{F}$ other than M itself.*

Example 1

Let $\mathcal{F}$ be the family of all subsets of a nonempty set S. (This family $\mathcal{F}$ is called the **power set** of S.) The set S is easily seen to be a maximal element of $\mathcal{F}$. ♦

Example 2

Let S and T be disjoint nonempty sets, and let $\mathcal{F}$ be the union of their power sets. Then S and T are both maximal elements of $\mathcal{F}$. ♦

The Hausdorff maximal principle implies that $\mathcal{F}$ has a maximal element. This element is easily seen to be a maximal linearly independent subset of V that contains S. ∎

Corollary. *Every vector space has a basis.*

It can be shown, analogously to Corollary 1 of the replacement theorem (p. 47), that every basis for an infinite-dimensional vector space has the same *cardinality.* (Sets have the same cardinality if there is a one-to-one and onto mapping between them.) (See, for example, N. Jacobson, *Lectures in Abstract Algebra*, vol. 2, Linear Algebra, D. Van Nostrand Company, New York, 1953, p. 240.)

Exercises 4–7 extend other results from Section 1.6 to infinite-dimensional vector spaces.

EXERCISES

1. Label the following statements as true or false.
   (a) Every family of sets contains a maximal element.
   (b) Every chain contains a maximal element.
   (c) If a family of sets has a maximal element, then that maximal element is unique.
   (d) If a chain of sets has a maximal element, then that maximal element is unique.
   (e) A basis for a vector space is a maximal linearly independent subset of that vector space.
   (f) A maximal linearly independent subset of a vector space is a basis for that vector space.

2. Show that the set of convergent sequences is an infinite-dimensional subspace of the vector space of all sequences of real numbers. (See Exercise 21 in Section 1.3.)

3. Let V be the set of real numbers regarded as a vector space over the field of rational numbers. Prove that V is infinite-dimensional. *Hint:* Use the fact that π is transcendental, that is, π is not a zero of any polynomial with rational coefficients.

4. Let W be a subspace of a (not necessarily finite-dimensional) vector space V. Prove that any basis for W is a subset of a basis for V.

5. Prove the following infinite-dimensional version of Theorem 1.8 (p. 44): Let β be a subset of an infinite-dimensional vector space V. Then β is a basis for V if and only if for each nonzero vector v in V, there exist unique vectors $u_1, u_2, \ldots, u_n$ in β and unique nonzero scalars $c_1, c_2, \ldots, c_n$ such that $v = c_1u_1 + c_2u_2 + \cdots + c_nu_n$. Visit `goo.gl/fNWSDM` for a solution.

6. Prove the following generalization of Theorem 1.9 (p. 45): Let S_1 and S_2 be subsets of a vector space V such that $S_1 \subseteq S_2$. If S_1 is linearly independent and S_2 generates V, then there exists a basis β for V such that $S_1 \subseteq \beta \subseteq S_2$. *Hint:* Apply the Hausdorff maximal principle to the family of all linearly independent subsets of S_2 that contain S_1, and proceed as in the proof of Theorem 1.13.

7. Prove the following generalization of the replacement theorem. Let β be a basis for a vector space V, and let S be a linearly independent subset of V. There exists a subset S_1 of β such that $S \cup S_1$ is a basis for V.

INDEX OF DEFINITIONS FOR CHAPTER 1

Example 14

Let $\mathsf{T}: \mathsf{R}^2 \to \mathsf{R}^2$ be the linear transformation defined by

$$\mathsf{T}(a_1, a_2) = (2a_2 - a_1, 3a_1),$$

and suppose that $\mathsf{U}: \mathsf{R}^2 \to \mathsf{R}^2$ is linear. If we know that $\mathsf{U}(1,2) = (3,3)$ and $\mathsf{U}(1,1) = (1,3)$, then $\mathsf{U} = \mathsf{T}$. This follows from the corollary and from the fact that $\{(1,2),(1,1)\}$ is a basis for R^2. ♦

EXERCISES

1. Label the following statements as true or false. In each part, V and W are finite-dimensional vector spaces (over F), and T is a function from V to W.

   (a) If T is linear, then T preserves sums and scalar products.
   (b) If $\mathsf{T}(x+y) = \mathsf{T}(x) + \mathsf{T}(y)$, then T is linear.
   (c) T is one-to-one if and only if the only vector x such that $\mathsf{T}(x) = 0$ is $x = 0$.
   (d) If T is linear, then $\mathsf{T}(0_{\mathsf{V}}) = 0_{\mathsf{W}}$.
   (e) If T is linear, then $\text{nullity}(\mathsf{T}) + \text{rank}(\mathsf{T}) = \dim(\mathsf{W})$.
   (f) If T is linear, then T carries linearly independent subsets of V onto linearly independent subsets of W.
   (g) If $\mathsf{T}, \mathsf{U}: \mathsf{V} \to \mathsf{W}$ are both linear and agree on a basis for V, then $\mathsf{T} = \mathsf{U}$.
   (h) Given $x_1, x_2 \in \mathsf{V}$ and $y_1, y_2 \in \mathsf{W}$, there exists a linear transformation $\mathsf{T}: \mathsf{V} \to \mathsf{W}$ such that $\mathsf{T}(x_1) = y_1$ and $\mathsf{T}(x_2) = y_2$.

For Exercises 2 through 6, prove that T is a linear transformation, and find bases for both N(T) and R(T). Then compute the nullity and rank of T, and verify the dimension theorem. Finally, use the appropriate theorems in this section to determine whether T is one-to-one or onto.

2. $\mathsf{T}: \mathsf{R}^3 \to \mathsf{R}^2$ defined by $\mathsf{T}(a_1, a_2, a_3) = (a_1 - a_2, 2a_3)$.

3. $\mathsf{T}: \mathsf{R}^2 \to \mathsf{R}^3$ defined by $\mathsf{T}(a_1, a_2) = (a_1 + a_2, 0, 2a_1 - a_2)$.

4. $\mathsf{T}: \mathsf{M}_{2\times 3}(F) \to \mathsf{M}_{2\times 2}(F)$ defined by

$$\mathsf{T}\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} = \begin{pmatrix} 2a_{11} - a_{12} & a_{13} + 2a_{12} \\ 0 & 0 \end{pmatrix}.$$

5. $\mathsf{T}: \mathsf{P}_2(R) \to \mathsf{P}_3(R)$ defined by $\mathsf{T}(f(x)) = xf(x) + f'(x)$.

6. $\mathsf{T}\colon \mathsf{M}_{n\times n}(F) \to F$ defined by $\mathsf{T}(A) = \operatorname{tr}(A)$. Recall (Example 4, Section 1.3) that
$$\operatorname{tr}(A) = \sum_{i=1}^{n} A_{ii}.$$

7. Prove properties 1, 2, 3, and 4 on page 65.

8. Prove that the transformations in Examples 2 and 3 are linear.

9. In this exercise, $\mathsf{T}\colon \mathsf{R}^2 \to \mathsf{R}^2$ is a function. For each of the following parts, state why T is *not* linear.
   (a) $\mathsf{T}(a_1, a_2) = (1, a_2)$
   (b) $\mathsf{T}(a_1, a_2) = (a_1, a_1^2)$
   (c) $\mathsf{T}(a_1, a_2) = (\sin a_1, 0)$
   (d) $\mathsf{T}(a_1, a_2) = (|a_1|, a_2)$
   (e) $\mathsf{T}(a_1, a_2) = (a_1 + 1, a_2)$

10. Suppose that $\mathsf{T}\colon \mathsf{R}^2 \to \mathsf{R}^2$ is linear, $\mathsf{T}(1,0) = (1,4)$, and $\mathsf{T}(1,1) = (2,5)$. What is $\mathsf{T}(2,3)$? Is T one-to-one?

11. Prove that there exists a linear transformation $\mathsf{T}\colon \mathsf{R}^2 \to \mathsf{R}^3$ such that $\mathsf{T}(1,1) = (1,0,2)$ and $\mathsf{T}(2,3) = (1,-1,4)$. What is $\mathsf{T}(8,11)$?

12. Is there a linear transformation $\mathsf{T}\colon \mathsf{R}^3 \to \mathsf{R}^2$ such that $\mathsf{T}(1,0,3) = (1,1)$ and $\mathsf{T}(-2,0,-6) = (2,1)$?

$\boxed{13.}$ Let V and W be vector spaces, let $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$ be linear, and let $\{w_1, w_2, \ldots, w_k\}$ be a linearly independent set of k vectors from $\mathsf{R}(\mathsf{T})$. Prove that if $S = \{v_1, v_2, \ldots, v_k\}$ is chosen so that $\mathsf{T}(v_i) = w_i$ for $i = 1, 2, \ldots, k$, then S is linearly independent. Visit goo.gl/kmaQS2 for a solution.

14. Let V and W be vector spaces and $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$ be linear.
   (a) Prove that T is one-to-one if and only if T carries linearly independent subsets of V onto linearly independent subsets of W.
   (b) Suppose that T is one-to-one and that S is a subset of V. Prove that S is linearly independent if and only if $\mathsf{T}(S)$ is linearly independent.
   (c) Suppose $\beta = \{v_1, v_2, \ldots, v_n\}$ is a basis for V and T is one-to-one and onto. Prove that $\mathsf{T}(\beta) = \{\mathsf{T}(v_1), \mathsf{T}(v_2), \ldots, \mathsf{T}(v_n)\}$ is a basis for W.

15. Recall the definition of $\mathsf{P}(R)$ on page 11. Define
$$\mathsf{T}\colon \mathsf{P}(R) \to \mathsf{P}(R) \quad \text{by} \quad \mathsf{T}(f(x)) = \int_0^x f(t)\,dt.$$
Prove that T linear and one-to-one, but not onto.

16. Let $\mathsf{T}\colon \mathsf{P}(R) \to \mathsf{P}(R)$ be defined by $\mathsf{T}(f(x)) = f'(x)$. Recall that T is linear. Prove that T is onto, but not one-to-one.

17. Let V and W be finite-dimensional vector spaces and $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$ be linear.

(a) Prove that if $\dim(\mathsf{V}) < \dim(\mathsf{W})$, then T cannot be onto.
(b) Prove that if $\dim(\mathsf{V}) > \dim(\mathsf{W})$, then T cannot be one-to-one.

18. Give an example of a linear transformation $\mathsf{T}\colon \mathsf{R}^2 \to \mathsf{R}^2$ such that $\mathsf{N}(\mathsf{T}) = \mathsf{R}(\mathsf{T})$.

19. Give an example of vector spaces V and W and distinct linear transformations T and U from V to W such that $\mathsf{N}(\mathsf{T}) = \mathsf{N}(\mathsf{U})$ and $\mathsf{R}(\mathsf{T}) = \mathsf{R}(\mathsf{U})$.

20. Let V and W be vector spaces with subspaces V_1 and W_1, respectively. If $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$ is linear, prove that $\mathsf{T}(\mathsf{V}_1)$ is a subspace of W and that $\{x \in \mathsf{V}\colon \mathsf{T}(x) \in \mathsf{W}_1\}$ is a subspace of V.

21. Let V be the vector space of sequences described in Example 5 of Section 1.2. Define the functions $\mathsf{T}, \mathsf{U}\colon \mathsf{V} \to \mathsf{V}$ by

$$\mathsf{T}(a_1, a_2, \ldots) = (a_2, a_3, \ldots) \quad \text{and} \quad \mathsf{U}(a_1, a_2, \ldots) = (0, a_1, a_2, \ldots).$$

T and U are called the **left shift** and **right shift** operators on V, respectively.

(a) Prove that T and U are linear.
(b) Prove that T is onto, but not one-to-one.
(c) Prove that U is one-to-one, but not onto.

22. Let $\mathsf{T}\colon \mathsf{R}^3 \to R$ be linear. Show that there exist scalars a, b, and c such that $\mathsf{T}(x, y, z) = ax + by + cz$ for all $(x, y, z) \in \mathsf{R}^3$. Can you generalize this result for $\mathsf{T}\colon \mathsf{F}^n \to F$? State and prove an analogous result for $\mathsf{T}\colon \mathsf{F}^n \to \mathsf{F}^m$.

23. Let $\mathsf{T}\colon \mathsf{R}^3 \to R$ be linear. Describe geometrically the possibilities for the null space of T. *Hint:* Use Exercise 22.

24. Let $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$ be linear, $b \in \mathsf{W}$, and $K = \{x \in \mathsf{V}\colon \mathsf{T}(x) = b\}$ be nonempty. Prove that if $s \in K$, then $K = \{s\} + \mathsf{N}(\mathsf{T})$. (See page 22 for the definition of the sum of subsets.)

The following definition is used in Exercises 25–28 and in Exercise 31.

Definition. *Let V be a vector space and W_1 and W_2 be subspaces of V such that $\mathsf{V} = \mathsf{W}_1 \oplus \mathsf{W}_2$. (Recall the definition of direct sum given on page 22.) The function $\mathsf{T}\colon \mathsf{V} \to \mathsf{V}$ defined by $\mathsf{T}(x) = x_1$ where $x = x_1 + x_2$ with $x_1 \in \mathsf{W}_1$ and $x_2 \in \mathsf{W}_2$, is called the* ***projection of*** V ***on*** W_1 *or the* ***projection on*** W_1 ***along*** W_2.

25. Let $\mathsf{T}\colon \mathsf{R}^2 \to \mathsf{R}^2$. Include figures for each of the following parts.

(a) Find a formula for $\mathsf{T}(a,b)$, where T represents the projection on the y-axis along the x-axis.
(b) Find a formula for $\mathsf{T}(a,b)$, where T represents the projection on the y-axis along the line $L = \{(s,s)\colon s \in R\}$.

26. Let $\mathsf{T}\colon \mathsf{R}^3 \to \mathsf{R}^3$.

(a) If $\mathsf{T}(a,b,c) = (a,b,0)$, show that T is the projection on the xy-plane along the z-axis.
(b) Find a formula for $\mathsf{T}(a,b,c)$, where T represents the projection on the z-axis along the xy-plane.
(c) If $\mathsf{T}(a,b,c) = (a-c,b,0)$, show that T is the projection on the xy-plane along the line $L = \{(a,0,a)\colon a \in R\}$.

27. Using the notation in the definition above, assume that $\mathsf{T}\colon \mathsf{V} \to \mathsf{V}$ is the projection on W_1 along W_2.

(a) Prove that T is linear and $\mathsf{W}_1 = \{x \in \mathsf{V}\colon \mathsf{T}(x) = x\}$.
(b) Prove that $\mathsf{W}_1 = \mathsf{R}(\mathsf{T})$ and $\mathsf{W}_2 = \mathsf{N}(\mathsf{T})$.
(c) Describe T if $\mathsf{W}_1 = \mathsf{V}$.
(d) Describe T if W_1 is the zero subspace.

28. Suppose that W is a subspace of a finite-dimensional vector space V.

(a) Prove that there exists a subspace W' and a function $\mathsf{T}\colon \mathsf{V} \to \mathsf{V}$ such that T is a projection on W along W'.
(b) Give an example of a subspace W of a vector space V such that there are *two* projections on W along two (distinct) subspaces.

The following definitions are used in Exercises 29–33.

Definitions. *Let V be a vector space, and let $\mathsf{T}\colon \mathsf{V} \to \mathsf{V}$ be linear. A subspace W of V is said to be* **T-invariant** *if $\mathsf{T}(x) \in \mathsf{W}$ for every $x \in \mathsf{W}$, that is, $\mathsf{T}(\mathsf{W}) \subseteq \mathsf{W}$. If W is T-invariant, we define the* **restriction of T on W** *to be the function $\mathsf{T}_\mathsf{W}\colon \mathsf{W} \to \mathsf{W}$ defined by $\mathsf{T}_\mathsf{W}(x) = \mathsf{T}(x)$ for all $x \in \mathsf{W}$.*

Exercises 29–33 assume that W is a subspace of a vector space V and that $\mathsf{T}\colon \mathsf{V} \to \mathsf{V}$ is linear. *Warning:* Do not assume that W is T-invariant or that T is a projection unless explicitly stated.

29. Prove that the subspaces $\{0\}$, V, $\mathsf{R}(\mathsf{T})$, and $\mathsf{N}(\mathsf{T})$ are all T-invariant.

30. If W is T-invariant, prove that T_W is linear.

31. Suppose that T is the projection on W along some subspace W'. Prove that W is T-invariant and that $\mathsf{T}_\mathsf{W} = \mathsf{I}_\mathsf{W}$.

32. Suppose that $\mathsf{V} = \mathsf{R}(\mathsf{T}) \oplus \mathsf{W}$ and W is T-invariant. See page 22 for the definition of *direct sum.*

(a) Prove that $\mathsf{W} \subseteq \mathsf{N}(\mathsf{T})$.
(b) Show that if V is finite-dimensional, then $\mathsf{W} = \mathsf{N}(\mathsf{T})$.
(c) Show by example that the conclusion of (b) is not necessarily true if V is not finite-dimensional.

33. Suppose that W is T-invariant. Prove that $\mathsf{N}(\mathsf{T_W}) = \mathsf{N}(\mathsf{T}) \cap \mathsf{W}$ and $\mathsf{R}(\mathsf{T_W}) = \mathsf{T}(\mathsf{W})$.

34. Prove Theorem 2.2 for the case that β is infinite, that is, $\mathsf{R}(\mathsf{T}) = \operatorname{span}(\{\mathsf{T}(v) : v \in \beta\})$.

35. Prove the following generalization of Theorem 2.6: Let V and W be vector spaces over a common field, and let β be a basis for V. Then for any function $f : \beta \to \mathsf{W}$ there exists exactly one linear transformation $\mathsf{T} : \mathsf{V} \to \mathsf{W}$ such that $\mathsf{T}(x) = f(x)$ for all $x \in \beta$.

Exercises 36 and 37 require the definition of *direct sum* given on page 22.

36. Let V be a finite-dimensional vector space and $\mathsf{T} : \mathsf{V} \to \mathsf{V}$ be linear.

(a) Suppose that $\mathsf{V} = \mathsf{R}(\mathsf{T}) + \mathsf{N}(\mathsf{T})$. Prove that $\mathsf{V} = \mathsf{R}(\mathsf{T}) \oplus \mathsf{N}(\mathsf{T})$.
(b) Suppose that $\mathsf{R}(\mathsf{T}) \cap \mathsf{N}(\mathsf{T}) = \{0\}$. Prove that $\mathsf{V} = \mathsf{R}(\mathsf{T}) \oplus \mathsf{N}(\mathsf{T})$.

Be careful to say in each part where finite-dimensionality is used.

37. Let V and T be as defined in Exercise 21.

(a) Prove that $\mathsf{V} = \mathsf{R}(\mathsf{T}) + \mathsf{N}(\mathsf{T})$, but V is not a direct sum of these two spaces. Thus the result of Exercise 36(a) above cannot be proved without assuming that V is finite-dimensional.
(b) Find a linear operator T_1 on V such that $\mathsf{R}(\mathsf{T}_1) \cap \mathsf{N}(\mathsf{T}_1) = \{0\}$ but V is not a direct sum of $\mathsf{R}(\mathsf{T}_1)$ and $\mathsf{N}(\mathsf{T}_1)$. Conclude that V being finite-dimensional is also essential in Exercise 36(b).

38. A function $\mathsf{T} : \mathsf{V} \to \mathsf{W}$ between vector spaces V and W is called **additive** if $\mathsf{T}(x + y) = \mathsf{T}(x) + \mathsf{T}(y)$ for all $x, y \in \mathsf{V}$. Prove that if V and W are vector spaces over the field of rational numbers, then any additive function from V into W is a linear transformation.

39. Let $\mathsf{T} : C \to C$ be the function defined by $\mathsf{T}(z) = \overline{z}$. Prove that T is additive (as defined in Exercise 38) but not linear.

40. Prove that there is an additive function $\mathsf{T} : R \to R$ (as defined in Exercise 38) that is not linear. *Hint:* Let V be the set of real numbers regarded as a vector space over the field of rational numbers. By the corollary to Theorem 1.13 (p. 61), V has a basis β. Let x and y be two distinct vectors in β, and define $f\colon \beta \to \mathsf{V}$ by $f(x) = y$, $f(y) = x$, and $f(z) = z$ otherwise. By Exercise 35, there exists a linear transformation $\mathsf{T}\colon \mathsf{V} \to \mathsf{V}$ such that $\mathsf{T}(u) = f(u)$ for all $u \in \beta$. Then T is additive, but for $c = y/x$, $\mathsf{T}(cx) \neq c\mathsf{T}(x)$.

41. Prove that Theorem 2.6 and its corollary are true when V is infinite-dimensional.

The following exercise requires familiarity with the definition of *quotient space* given in Exercise 31 of Section 1.3.

42. Let V be a vector space and W be a subspace of V. Define the mapping $\eta\colon \mathsf{V} \to \mathsf{V}/\mathsf{W}$ by $\eta(v) = v + \mathsf{W}$ for $v \in \mathsf{V}$.

(a) Prove that η is a linear transformation from V onto V/W and that $\mathsf{N}(\eta) = \mathsf{W}$.

(b) Suppose that V is finite-dimensional. Use (a) and the dimension theorem to derive a formula relating $\dim(\mathsf{V})$, $\dim(\mathsf{W})$, and $\dim(\mathsf{V}/\mathsf{W})$.

(c) Read the proof of the dimension theorem. Compare the method of solving (b) with the method of deriving the same result as outlined in Exercise 35 of Section 1.6.

2.2 THE MATRIX REPRESENTATION OF A LINEAR TRANSFORMATION

Until now, we have studied linear transformations by examining their ranges and null spaces. In this section, we embark on one of the most useful approaches to the analysis of a linear transformation on a finite-dimensional vector space: the representation of a linear transformation by a matrix. In fact, we develop a one-to-one correspondence between matrices and linear transformations that allows us to utilize properties of one to study properties of the other.

We first need the concept of an *ordered basis* for a vector space.

Definition. *Let* V *be a finite-dimensional vector space. An* **ordered basis** *for* V *is a basis for* V *endowed with a specific order; that is, an ordered basis for* V *is a finite sequence of linearly independent vectors in* V *that generates* V.

Example 1

In F^3, $\beta = \{e_1, e_2, e_3\}$ can be considered an ordered basis. Also $\gamma = \{e_2, e_1, e_3\}$ is an ordered basis, but $\beta \neq \gamma$ as ordered bases. ◆

For the vector space F^n, we call $\{e_1, e_2, \ldots, e_n\}$ the **standard ordered basis** for F^n. Similarly, for the vector space $\mathsf{P}_n(F)$, we call $\{1, x, \ldots, x^n\}$ the **standard ordered basis** for $\mathsf{P}_n(F)$.

Thus

$$([\mathsf{T}+\mathsf{U}]_\beta^\gamma)_{ij} = a_{ij} + b_{ij} = ([\mathsf{T}]_\beta^\gamma + [\mathsf{U}]_\beta^\gamma)_{ij}.$$

So (a) is proved, and the proof of (b) is similar. ∎

Example 5

Let $\mathsf{T}\colon \mathsf{R}^2 \to \mathsf{R}^3$ and $\mathsf{U}\colon \mathsf{R}^2 \to \mathsf{R}^3$ be the linear transformations respectively defined by

$$\mathsf{T}(a_1, a_2) = (a_1 + 3a_2, 0, 2a_1 - 4a_2) \text{ and } \mathsf{U}(a_1, a_2) = (a_1 - a_2, 2a_1, 3a_1 + 2a_2).$$

Let β and γ be the standard ordered bases of R^2 and R^3, respectively. Then

$$[\mathsf{T}]_\beta^\gamma = \begin{pmatrix} 1 & 3 \\ 0 & 0 \\ 2 & -4 \end{pmatrix},$$

(as computed in Example 3), and

$$[\mathsf{U}]_\beta^\gamma = \begin{pmatrix} 1 & -1 \\ 2 & 0 \\ 3 & 2 \end{pmatrix}.$$

If we compute $\mathsf{T}+\mathsf{U}$ using the preceding definitions, we obtain

$$(\mathsf{T}+\mathsf{U})(a_1, a_2) = (2a_1 + 2a_2, 2a_1, 5a_1 - 2a_2).$$

So

$$[\mathsf{T}+\mathsf{U}]_\beta^\gamma = \begin{pmatrix} 2 & 2 \\ 2 & 0 \\ 5 & -2 \end{pmatrix},$$

which is simply $[\mathsf{T}]_\beta^\gamma + [\mathsf{U}]_\beta^\gamma$, illustrating Theorem 2.8. ♦

EXERCISES

1. Label the following statements as true or false. Assume that V and W are finite-dimensional vector spaces with ordered bases β and γ, respectively, and $\mathsf{T}, \mathsf{U}\colon \mathsf{V} \to \mathsf{W}$ are linear transformations.
   (a) For any scalar a, $a\mathsf{T} + \mathsf{U}$ is a linear transformation from V to W.
   (b) $[\mathsf{T}]_\beta^\gamma = [\mathsf{U}]_\beta^\gamma$ implies that $\mathsf{T} = \mathsf{U}$.
   (c) If $m = \dim(\mathsf{V})$ and $n = \dim(\mathsf{W})$, then $[\mathsf{T}]_\beta^\gamma$ is an $m \times n$ matrix.
   (d) $[\mathsf{T}+\mathsf{U}]_\beta^\gamma = [\mathsf{T}]_\beta^\gamma + [\mathsf{U}]_\beta^\gamma$.
   (e) $\mathcal{L}(\mathsf{V}, \mathsf{W})$ is a vector space.
   (f) $\mathcal{L}(\mathsf{V}, \mathsf{W}) = \mathcal{L}(\mathsf{W}, \mathsf{V})$.

2. Let β and γ be the standard ordered bases for R^n and R^m, respectively. For each linear transformation $\mathsf{T}\colon \mathsf{R}^n \to \mathsf{R}^m$, compute $[\mathsf{T}]_\beta^\gamma$.

   (a) $\mathsf{T}\colon \mathsf{R}^2 \to \mathsf{R}^3$ defined by $\mathsf{T}(a_1, a_2) = (2a_1 - a_2, 3a_1 + 4a_2, a_1)$.
   (b) $\mathsf{T}\colon \mathsf{R}^3 \to \mathsf{R}^2$ defined by $\mathsf{T}(a_1, a_2, a_3) = (2a_1 + 3a_2 - a_3, a_1 + a_3)$.
   (c) $\mathsf{T}\colon \mathsf{R}^3 \to R$ defined by $\mathsf{T}(a_1, a_2, a_3) = 2a_1 + a_2 - 3a_3$.
   (d) $\mathsf{T}\colon \mathsf{R}^3 \to \mathsf{R}^3$ defined by

$$\mathsf{T}(a_1, a_2, a_3) = (2a_2 + a_3, -a_1 + 4a_2 + 5a_3, a_1 + a_3).$$

   (e) $\mathsf{T}\colon \mathsf{R}^n \to \mathsf{R}^n$ defined by $\mathsf{T}(a_1, a_2, \ldots, a_n) = (a_1, a_1, \ldots, a_1)$.
   (f) $\mathsf{T}\colon \mathsf{R}^n \to \mathsf{R}^n$ defined by $\mathsf{T}(a_1, a_2, \ldots, a_n) = (a_n, a_{n-1}, \ldots, a_1)$.
   (g) $\mathsf{T}\colon \mathsf{R}^n \to R$ defined by $\mathsf{T}(a_1, a_2, \ldots, a_n) = a_1 + a_n$.

3. Let $\mathsf{T}\colon \mathsf{R}^2 \to \mathsf{R}^3$ be defined by $\mathsf{T}(a_1, a_2) = (a_1 - a_2, a_1, 2a_1 + a_2)$. Let β be the standard ordered basis for R^2 and $\gamma = \{(1,1,0), (0,1,1), (2,2,3)\}$. Compute $[\mathsf{T}]_\beta^\gamma$. If $\alpha = \{(1,2), (2,3)\}$, compute $[\mathsf{T}]_\alpha^\gamma$.

4. Define

$$\mathsf{T}\colon \mathsf{M}_{2\times 2}(R) \to \mathsf{P}_2(R) \quad \text{by} \quad \mathsf{T}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = (a+b) + (2d)x + bx^2.$$

   Let

$$\beta = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\} \quad \text{and} \quad \gamma = \{1, x, x^2\}.$$

   Compute $[\mathsf{T}]_\beta^\gamma$.

5. Let

$$\alpha = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\},$$
$$\beta = \{1, x, x^2\},$$

   and

$$\gamma = \{1\}.$$

   (a) Define $\mathsf{T}\colon \mathsf{M}_{2\times 2}(F) \to \mathsf{M}_{2\times 2}(F)$ by $\mathsf{T}(A) = A^t$. Compute $[\mathsf{T}]_\alpha$.
   (b) Define

$$\mathsf{T}\colon \mathsf{P}_2(R) \to \mathsf{M}_{2\times 2}(R) \quad \text{by} \quad \mathsf{T}(f(x)) = \begin{pmatrix} f'(0) & 2f(1) \\ 0 & f''(3) \end{pmatrix},$$

   where $'$ denotes differentiation. Compute $[\mathsf{T}]_\beta^\alpha$.
   (c) Define $\mathsf{T}\colon \mathsf{M}_{2\times 2}(F) \to F$ by $\mathsf{T}(A) = \operatorname{tr}(A)$. Compute $[\mathsf{T}]_\alpha^\gamma$.

(d) Define $\mathsf{T}\colon \mathsf{P}_2(R) \to R$ by $\mathsf{T}(f(x)) = f(2)$. Compute $[\mathsf{T}]_\beta^\gamma$.
(e) If

$$A = \begin{pmatrix} 1 & -2 \\ 0 & 4 \end{pmatrix},$$

compute $[A]_\alpha$.
(f) If $f(x) = 3 - 6x + x^2$, compute $[f(x)]_\beta$.
(g) For $a \in F$, compute $[a]_\gamma$.

6. Complete the proof of part (b) of Theorem 2.7.

7. Prove part (b) of Theorem 2.8.

8.† Let V be an n-dimensional vector space with an ordered basis β. Define $\mathsf{T}\colon \mathsf{V} \to \mathsf{F}^n$ by $\mathsf{T}(x) = [x]_\beta$. Prove that T is linear.

9. Let V be the vector space of complex numbers over the field R. Define $\mathsf{T}\colon \mathsf{V} \to \mathsf{V}$ by $\mathsf{T}(z) = \overline{z}$, where $\overline{z}$ is the complex conjugate of z. Prove that T is linear, and compute $[\mathsf{T}]_\beta$, where $\beta = \{1, i\}$. (Recall by Exercise 39 of Section 2.1 that T is not linear if V is regarded as a vector space over the field C.)

10. Let V be a vector space with the ordered basis $\beta = \{v_1, v_2, \ldots, v_n\}$. Define $v_0 = 0$. By Theorem 2.6 (p. 73), there exists a linear transformation $\mathsf{T}\colon \mathsf{V} \to \mathsf{V}$ such that $\mathsf{T}(v_j) = v_j + v_{j-1}$ for $j = 1, 2, \ldots, n$. Compute $[\mathsf{T}]_\beta$.

11. Let V be an n-dimensional vector space, and let $\mathsf{T}\colon \mathsf{V} \to \mathsf{V}$ be a linear transformation. Suppose that W is a T-invariant subspace of V (see the exercises of Section 2.1) having dimension k. Show that there is a basis β for V such that $[\mathsf{T}]_\beta$ has the form

$$\begin{pmatrix} A & B \\ O & C \end{pmatrix},$$

where A is a $k \times k$ matrix and O is the $(n-k) \times k$ zero matrix.

12.† Let $\beta = \{v_1, v_2, \ldots, v_n\}$ be a basis for a vector space V and $\mathsf{T}\colon \mathsf{V} \to \mathsf{V}$ be a linear transformation. Prove that $[\mathsf{T}]_\beta$ is upper triangular if and only if $\mathsf{T}(v_j) \in \operatorname{span}(\{v_1, v_2, \ldots, v_j\})$ for $j = 1, 2, \ldots, n$. Visit `goo.gl/k9ZrQb` for a solution.

13. Let V be a finite-dimensional vector space and T be the projection on W along W', where W and W' are subspaces of V. (See the definition in the exercises of Section 2.1 on page 76.) Find an ordered basis β for V such that $[\mathsf{T}]_\beta$ is a diagonal matrix.

14. Let V and W be vector spaces, and let T and U be nonzero linear transformations from V into W. If $\mathsf{R}(\mathsf{T}) \cap \mathsf{R}(\mathsf{U}) = \{0\}$, prove that $\{\mathsf{T}, \mathsf{U}\}$ is a linearly independent subset of $\mathcal{L}(\mathsf{V}, \mathsf{W})$.

15. Let $\mathsf{V} = \mathsf{P}(R)$, and for $j \geq 1$ define $\mathsf{T}_j(f(x)) = f^{(j)}(x)$, where $f^{(j)}(x)$ is the jth derivative of $f(x)$. Prove that the set $\{\mathsf{T}_1, \mathsf{T}_2, \ldots, \mathsf{T}_n\}$ is a linearly independent subset of $\mathcal{L}(\mathsf{V})$ for any positive integer n.

16. Let V and W be vector spaces, and let S be a subset of V. Define $S^0 = \{\mathsf{T} \in \mathcal{L}(\mathsf{V}, \mathsf{W}) \colon \mathsf{T}(x) = 0 \text{ for all } x \in S\}$. Prove the following statements.
    (a) S^0 is a subspace of $\mathcal{L}(\mathsf{V}, \mathsf{W})$.
    (b) If S_1 and S_2 are subsets of V and $S_1 \subseteq S_2$, then $S_2^0 \subseteq S_1^0$.
    (c) If V_1 and V_2 are subspaces of V, then $(\mathsf{V}_1 \cup \mathsf{V}_2)^0 = (\mathsf{V}_1 + \mathsf{V}_2)^0 = \mathsf{V}_1^0 \cap \mathsf{V}_2^0$.

17. Let V and W be vector spaces such that $\dim(\mathsf{V}) = \dim(\mathsf{W})$, and let $\mathsf{T} \colon \mathsf{V} \to \mathsf{W}$ be linear. Show that there exist ordered bases β and γ for V and W, respectively, such that $[\mathsf{T}]_\beta^\gamma$ is a diagonal matrix.

2.3 COMPOSITION OF LINEAR TRANSFORMATIONS AND MATRIX MULTIPLICATION

In Section 2.2, we learned how to associate a matrix with a linear transformation in such a way that both sums and scalar multiples of matrices are associated with the corresponding sums and scalar multiples of the transformations. The question now arises as to how the matrix representation of a composite of linear transformations is related to the matrix representation of each of the associated linear transformations. The attempt to answer this question leads to a definition of matrix multiplication. We use the more convenient notation of UT rather than $\mathsf{U} \circ \mathsf{T}$ for the composite of linear transformations U and T. (See Appendix B.)

Our first result shows that the composite of linear transformations is linear.

Theorem 2.9. *Let V, W, and Z be vector spaces over the same field F, and let $\mathsf{T} \colon \mathsf{V} \to \mathsf{W}$ and $\mathsf{U} \colon \mathsf{W} \to \mathsf{Z}$ be linear. Then $\mathsf{UT} \colon \mathsf{V} \to \mathsf{Z}$ is linear.*

Proof. Let $x, y \in \mathsf{V}$ and $a \in F$. Then

$$\begin{aligned}\mathsf{UT}(ax + y) &= \mathsf{U}(\mathsf{T}(ax + y)) = \mathsf{U}(a\mathsf{T}(x) + \mathsf{T}(y)) \\ &= a\mathsf{U}(\mathsf{T}(x)) + \mathsf{U}(\mathsf{T}(y)) = a(\mathsf{UT})(x) + \mathsf{UT}(y).\end{aligned}$$ ∎

The following theorem lists some of the properties of the composition of linear transformations.

earlier, and compute B^3. In this case,

$$B = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix} \quad \text{and} \quad B^3 = \begin{pmatrix} 0 & 4 & 0 & 4 \\ 4 & 0 & 4 & 0 \\ 0 & 4 & 0 & 4 \\ 4 & 0 & 4 & 0 \end{pmatrix}.$$

Since all the diagonal entries of B^3 are zero, we conclude that there are no cliques in this relationship.

Our final example of the use of incidence matrices is concerned with the concept of *dominance*. A relation among a group of people is called a **dominance relation** if the associated incidence matrix A has the property that for all distinct pairs i and j, $A_{ij} = 1$ if and only if $A_{ji} = 0$, that is, given any two people, exactly one of them *dominates* (or, using the terminology of our first example, can send a message to) the other. Since A is an incidence matrix, $A_{ii} = 0$ for all i. For such a relation, it can be shown (see Exercise 22) that the matrix $A + A^2$ has a row [column] in which each entry is positive except for the diagonal entry. In other words, there is at least one person who dominates [is dominated by] all others in one or two stages. In fact, it can be shown that any person who dominates [is dominated by] the greatest number of people in the first stage has this property. Consider, for example, the matrix

$$A = \begin{pmatrix} 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 \end{pmatrix}.$$

The reader should verify that this matrix corresponds to a dominance relation. Now

$$A + A^2 = \begin{pmatrix} 0 & 2 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 0 \\ 1 & 2 & 0 & 2 & 1 \\ 1 & 2 & 2 & 0 & 1 \\ 2 & 2 & 2 & 2 & 0 \end{pmatrix}.$$

Thus persons 1, 3, 4, and 5 dominate (can send messages to) all the others in at most two stages, while persons 1, 2, 3, and 4 are dominated by (can receive messages from) all the others in at most two stages.

EXERCISES

1. Label the following statements as true or false. In each part, V, W, and Z denote vector spaces with ordered (finite) bases α, β, and γ, respectively; $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$ and $\mathsf{U}\colon \mathsf{W} \to \mathsf{Z}$ denote linear transformations; and A and B denote matrices.

(a) $[\mathsf{UT}]_\alpha^\gamma = [\mathsf{T}]_\alpha^\beta [\mathsf{U}]_\beta^\gamma$.
(b) $[\mathsf{T}(v)]_\beta = [\mathsf{T}]_\alpha^\beta [v]_\alpha$ for all $v \in \mathsf{V}$.
(c) $[\mathsf{U}(w)]_\beta = [\mathsf{U}]_\alpha^\beta [w]_\beta$ for all $w \in \mathsf{W}$.
(d) $[\mathsf{I_V}]_\alpha = I$.
(e) $[\mathsf{T}^2]_\alpha^\beta = ([\mathsf{T}]_\alpha^\beta)^2$.
(f) $A^2 = I$ implies that $A = I$ or $A = -I$.
(g) $\mathsf{T} = \mathsf{L}_A$ for some matrix A.
(h) $A^2 = O$ implies that $A = O$, where O denotes the zero matrix.
(i) $\mathsf{L}_{A+B} = \mathsf{L}_A + \mathsf{L}_B$.
(j) If A is square and $A_{ij} = \delta_{ij}$ for all i and j, then $A = I$.

2. (a) Let

$$A = \begin{pmatrix} 1 & 3 \\ 2 & -1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 & -3 \\ 4 & 1 & 2 \end{pmatrix},$$
$$C = \begin{pmatrix} 1 & 1 & 4 \\ -1 & -2 & 0 \end{pmatrix}, \quad \text{and} \quad D = \begin{pmatrix} 2 \\ -2 \\ 3 \end{pmatrix}.$$

Compute $A(2B + 3C)$, $(AB)D$, and $A(BD)$.

(b) Let

$$A = \begin{pmatrix} 2 & 5 \\ -3 & 1 \\ 4 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} 3 & -2 & 0 \\ 1 & -1 & 4 \\ 5 & 5 & 3 \end{pmatrix}, \quad \text{and} \quad C = \begin{pmatrix} 4 & 0 & 3 \end{pmatrix}.$$

Compute A^t, A^tB, BC^t, CB, and CA.

3. Let $g(x) = 3 + x$. Let $\mathsf{T}\colon \mathsf{P}_2(R) \to \mathsf{P}_2(R)$ and $\mathsf{U}\colon \mathsf{P}_2(R) \to \mathsf{R}^3$ be the linear transformations respectively defined by

$$\mathsf{T}(f(x)) = f'(x)g(x) + 2f(x) \quad \text{and} \quad U(a + bx + cx^2) = (a + b, c, a - b).$$

Let β and γ be the standard ordered bases of $\mathsf{P}_2(R)$ and R^3, respectively.

(a) Compute $[\mathsf{U}]_\beta^\gamma$, $[\mathsf{T}]_\beta$, and $[\mathsf{UT}]_\beta^\gamma$ directly. Then use Theorem 2.11 to verify your result.
(b) Let $h(x) = 3 - 2x + x^2$. Compute $[h(x)]_\beta$ and $[\mathsf{U}(h(x))]_\gamma$. Then use $[\mathsf{U}]_\beta^\gamma$ from (a) and Theorem 2.14 to verify your result.

4. For each of the following parts, let T be the linear transformation defined in the corresponding part of Exercise 5 of Section 2.2. Use Theorem 2.14 to compute the following vectors:

(a) $[\mathsf{T}(A)]_\alpha$, where $A = \begin{pmatrix} 1 & 4 \\ -1 & 6 \end{pmatrix}$.

(b) $[\mathsf{T}(f(x))]_\alpha$, where $f(x) = 4 - 6x + 3x^2$.
(c) $[\mathsf{T}(A)]_\gamma$, where $A = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}$.
(d) $[\mathsf{T}(f(x))]_\gamma$, where $f(x) = 6 - x + 2x^2$.

5. Complete the proof of Theorem 2.12 and its corollary.

6. Prove (b) of Theorem 2.13.

7. Prove (c) and (f) of Theorem 2.15.

8. Prove Theorem 2.10. Now state and prove a more general result involving linear transformations with domains unequal to their codomains.

9. Find linear transformations $\mathsf{U}, \mathsf{T}\colon \mathsf{F}^2 \to \mathsf{F}^2$ such that $\mathsf{UT} = \mathsf{T}_0$ (the zero transformation) but $\mathsf{TU} \neq \mathsf{T}_0$. Use your answer to find matrices A and B such that $AB = O$ but $BA \neq O$.

10. Let A be an $n \times n$ matrix. Prove that A is a diagonal matrix if and only if $A_{ij} = \delta_{ij} A_{ij}$ for all i and j.

11. Let V be a vector space, and let $\mathsf{T}\colon \mathsf{V} \to \mathsf{V}$ be linear. Prove that $\mathsf{T}^2 = \mathsf{T}_0$ if and only if $\mathsf{R}(\mathsf{T}) \subseteq \mathsf{N}(\mathsf{T})$.

12. Let V, W, and Z be vector spaces, and let $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$ and $\mathsf{U}\colon \mathsf{W} \to \mathsf{Z}$ be linear.

(a) Prove that if UT is one-to-one, then T is one-to-one. Must U also be one-to-one?
(b) Prove that if UT is onto, then U is onto. Must T also be onto?
(c) Prove that if U and T are one-to-one and onto, then UT is also.

13. Let A and B be $n \times n$ matrices. Recall that the trace of A is defined by

$$\operatorname{tr}(A) = \sum_{i=1}^{n} A_{ii}.$$

Prove that $\operatorname{tr}(AB) = \operatorname{tr}(BA)$ and $\operatorname{tr}(A) = \operatorname{tr}(A^t)$.

14. Assume the notation in Theorem 2.13.

(a) Suppose that z is a (column) vector in F^p. Use Theorem 2.13(b) to prove that Bz is a linear combination of the columns of B. In particular, if $z = (a_1, a_2, \ldots, a_p)^t$, then show that

$$Bz = \sum_{j=1}^{p} a_j v_j.$$

(b) Extend (a) to prove that column j of AB is a linear combination of the columns of A with the coefficients in the linear combination being the entries of column j of B.

(c) For any row vector $w \in \mathsf{F}^m$, prove that wA is a linear combination of the rows of A with the coefficients in the linear combination being the coordinates of w. *Hint:* Use properties of the transpose operation applied to (a).

(d) Prove the analogous result to (b) about rows: Row i of AB is a linear combination of the rows of B with the coefficients in the linear combination being the entries of row i of A.

$\boxed{15.}$ Let A and B be matrices for which the product matrix AB is defined; and let u_j and v_j denote the jth columns of AB and B, respectively. If $v_p = c_1v_{j_1} + c_2v_{j_2} + \cdots + c_kv_{j_k}$ for some scalars $c_1, c_2, \ldots c_k$, prove that $u_p = c_1u_{j_1} + c_2u_{j_2} + \cdots + c_ku_{j_k}$. Visit goo.gl/sRpves for a solution.

16. Let V be a finite-dimensional vector space, and let $\mathsf{T}\colon \mathsf{V} \to \mathsf{V}$ be linear.

(a) If $\operatorname{rank}(\mathsf{T}) = \operatorname{rank}(\mathsf{T}^2)$, prove that $\mathsf{R}(\mathsf{T}) \cap \mathsf{N}(\mathsf{T}) = \{0\}$. Deduce that $\mathsf{V} = \mathsf{R}(\mathsf{T}) \oplus \mathsf{N}(\mathsf{T})$ (see the exercises of Section 1.3).

(b) Prove that $\mathsf{V} = \mathsf{R}(\mathsf{T}^k) \oplus \mathsf{N}(\mathsf{T}^k)$ for some positive integer k.

17. For the definition of *projection* and related facts, see pages 76–77. Let V be a vector space and $\mathsf{T}\colon \mathsf{V} \to \mathsf{V}$ be a linear transformation. Prove that $\mathsf{T} = \mathsf{T}^2$ if and only if T is a projection on $\mathsf{W}_1 = \{y : \mathsf{T}(y) = y\}$ along $\mathsf{N}(\mathsf{T})$.

18. Let β be an ordered basis for a finite-dimensional vector space V, and let $\mathsf{T}\colon \mathsf{V} \to \mathsf{V}$ be linear. Prove that, for any nonnegative integer k, $[\mathsf{T}^k]_\beta = ([\mathsf{T}]_\beta)^k$.

19. Using only the definition of matrix multiplication, prove that, multiplication of matrices is associative.

20. For an incidence matrix A with related matrix B defined by $B_{ij} = 1$ if i is related to j and j is related to i, and $B_{ij} = 0$ otherwise, prove that i belongs to a clique if and only if $(B^3)_{ii} > 0$.

21. Use Exercise 20 to determine the cliques in the relations corresponding to the following incidence matrices.

(a) $\begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix}$ (b) $\begin{pmatrix} 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix}$

22. Let A be an incidence matrix that is associated with a dominance relation. Prove that the matrix $A + A^2$ has a row [column] in which each entry is positive except for the diagonal entry.

23. Prove that the matrix

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

corresponds to a dominance relation. Use Exercise 22 to determine which persons dominate [are dominated by] each of the others within two stages.

24. Let A be an $n \times n$ incidence matrix that corresponds to a dominance relation. Determine the number of nonzero entries of A.

2.4 INVERTIBILITY AND ISOMORPHISMS

The concept of invertibility is introduced quite early in the study of functions. Fortunately, many of the intrinsic properties of functions are shared by their inverses. For example, in calculus we learn that the properties of being continuous or differentiable are generally retained by the inverse functions. We see in this section (Theorem 2.17) that the inverse of a linear transformation is also linear. This result greatly aids us in the study of *inverses* of matrices. As one might expect from Section 2.3, the inverse of the left-multiplication transformation L_A (when it exists) can be used to determine properties of the inverse of the matrix A.

In the remainder of this section, we apply many of the results about invertibility to the concept of *isomorphism.* We will see that finite-dimensional vector spaces (over F) of equal dimension may be identified. These ideas will be made precise shortly.

The facts about inverse functions presented in Appendix B are, of course, true for linear transformations. Nevertheless, we repeat some of the definitions for use in this section.

Definition. *Let* V *and* W *be vector spaces, and let* $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$ *be linear. A function* $\mathsf{U}\colon \mathsf{W} \to \mathsf{V}$ *is said to be an* **inverse** *of* T *if* $\mathsf{TU} = \mathsf{I_W}$ *and* $\mathsf{UT} = \mathsf{I_V}$. *If* T *has an inverse, then* T *is said to be* **invertible**. *As noted in Appendix B, if* T *is invertible, then the inverse of* T *is unique and is denoted by* T^{-1}.

The following facts hold for invertible functions T and U.

1. $(\mathsf{TU})^{-1} = \mathsf{U}^{-1}\mathsf{T}^{-1}$.
2. $(\mathsf{T}^{-1})^{-1} = \mathsf{T}$; in particular, T^{-1} is invertible.

We often use the fact that a function is invertible if and only if it is both one-to-one and onto. We can therefore restate Theorem 2.5 as follows.

So $\mathsf{L}_A\phi_\beta(p(x)) = \phi_\gamma\mathsf{T}(p(x))$. ♦

Try repeating Example 7 with different polynomials $p(x)$.

EXERCISES

1. Label the following statements as true or false. In each part, V and W are vector spaces with ordered (finite) bases α and β, respectively, $\mathsf{T}: \mathsf{V} \to \mathsf{W}$ is linear, and A and B are matrices.

   (a) $([\mathsf{T}]_\alpha^\beta)^{-1} = [\mathsf{T}^{-1}]_\alpha^\beta$.
   (b) T is invertible if and only if T is one-to-one and onto.
   (c) $\mathsf{T} = \mathsf{L}_A$, where $A = [\mathsf{T}]_\alpha^\beta$.
   (d) $\mathsf{M}_{2\times 3}(F)$ is isomorphic to F^5.
   (e) $\mathsf{P}_n(F)$ is isomorphic to $\mathsf{P}_m(F)$ if and only if $n = m$.
   (f) $AB = I$ implies that A and B are invertible.
   (g) If A is invertible, then $(A^{-1})^{-1} = A$.
   (h) A is invertible if and only if L_A is invertible.
   (i) A must be square in order to possess an inverse.

2. For each of the following linear transformations T, determine whether T is invertible and justify your answer.

   (a) $\mathsf{T}: \mathsf{R}^2 \to \mathsf{R}^3$ defined by $\mathsf{T}(a_1, a_2) = (a_1 - 2a_2, a_2, 3a_1 + 4a_2)$.
   (b) $\mathsf{T}: \mathsf{R}^2 \to \mathsf{R}^3$ defined by $\mathsf{T}(a_1, a_2) = (3a_1 - a_2, a_2, 4a_1)$.
   (c) $\mathsf{T}: \mathsf{R}^3 \to \mathsf{R}^3$ defined by $\mathsf{T}(a_1, a_2, a_3) = (3a_1 - 2a_3, a_2, 3a_1 + 4a_2)$.
   (d) $\mathsf{T}: \mathsf{P}_3(R) \to \mathsf{P}_2(R)$ defined by $\mathsf{T}(p(x)) = p'(x)$.
   (e) $\mathsf{T}: \mathsf{M}_{2\times 2}(R) \to \mathsf{P}_2(R)$ defined by $\mathsf{T}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = a + 2bx + (c+d)x^2$.
   (f) $\mathsf{T}: \mathsf{M}_{2\times 2}(R) \to \mathsf{M}_{2\times 2}(R)$ defined by $\mathsf{T}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a+b & a \\ c & c+d \end{pmatrix}$.

3. Which of the following pairs of vector spaces are isomorphic? Justify your answers.

   (a) F^3 and $\mathsf{P}_3(F)$.
   (b) F^4 and $\mathsf{P}_3(F)$.
   (c) $\mathsf{M}_{2\times 2}(R)$ and $\mathsf{P}_3(R)$.
   (d) $\mathsf{V} = \{A \in \mathsf{M}_{2\times 2}(R) \colon \operatorname{tr}(A) = 0\}$ and R^4.

4.† Let A and B be $n \times n$ invertible matrices. Prove that AB is invertible and $(AB)^{-1} = B^{-1}A^{-1}$.

$\boxed{5.}$ Let A be invertible. Prove that A^t is invertible and $(A^t)^{-1} = (A^{-1})^t$. Visit goo.gl/suFm6V for a solution.

6. Prove that if A is invertible and $AB = O$, then $B = O$.

7. Let A be an $n \times n$ matrix.
   (a) Suppose that $A^2 = O$. Prove that A is not invertible.
   (b) Suppose that $AB = O$ for some nonzero $n \times n$ matrix B. Could A be invertible? Explain.

8. Prove Corollaries 1 and 2 of Theorem 2.18.

9.† Let A and B be $n \times n$ matrices such that AB is invertible.
   (a) Prove that A and B are invertible. *Hint:* See Exercise 12 of Section 2.3.
   (b) Give an example to show that a product of nonsquare matrices can be invertible even though the factors, by definition, are not.

10.† Let A and B be $n \times n$ matrices such that $AB = I_n$.
   (a) Use Exercise 9 to conclude that A and B are invertible.
   (b) Prove $A = B^{-1}$ (and hence $B = A^{-1}$). (We are, in effect, saying that for square matrices, a "one-sided" inverse is a "two-sided" inverse.)
   (c) State and prove analogous results for linear transformations defined on finite-dimensional vector spaces.

11. Verify that the transformation in Example 5 is one-to-one.

12. Prove Theorem 2.21.

13. Let $\sim$ mean "is isomorphic to." Prove that $\sim$ is an equivalence relation on the class of vector spaces over F.

14. Let
$$\mathsf{V} = \left\{ \begin{pmatrix} a & a+b \\ 0 & c \end{pmatrix} : a, b, c \in F \right\}.$$
Construct an isomorphism from V to F^3.

15. Let V and W be n-dimensional vector spaces, and let $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$ be a linear transformation. Suppose that β is a basis for V. Prove that T is an isomorphism if and only if $\mathsf{T}(\beta)$ is a basis for W.

16. Let B be an $n \times n$ invertible matrix. Define $\Phi\colon \mathsf{M}_{n\times n}(F) \to \mathsf{M}_{n\times n}(F)$ by $\Phi(A) = B^{-1}AB$. Prove that Φ is an isomorphism.

17.† Let V and W be finite-dimensional vector spaces and $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$ be an isomorphism. Let V_0 be a subspace of V.
   (a) Prove that $\mathsf{T}(\mathsf{V}_0)$ is a subspace of W.
   (b) Prove that $\dim(\mathsf{V}_0) = \dim(\mathsf{T}(\mathsf{V}_0))$.

18. Repeat Example 7 with the polynomial $p(x) = 1 + x + 2x^2 + x^3$.

19. In Example 5 of Section 2.1, the mapping $\mathsf{T}\colon \mathsf{M}_{2\times 2}(R) \to \mathsf{M}_{2\times 2}(R)$ defined by $\mathsf{T}(M) = M^t$ for each $M \in \mathsf{M}_{2\times 2}(R)$ is a linear transformation. Let $\beta = \{E^{11}, E^{12}, E^{21}, E^{22}\}$, which is a basis for $\mathsf{M}_{2\times 2}(R)$, as noted in Example 3 of Section 1.6.

(a) Compute $[\mathsf{T}]_\beta$.
(b) Verify that $\mathsf{L}_A\phi_\beta(M) = \phi_\beta\mathsf{T}(M)$ for $A = [\mathsf{T}]_\beta$ and

$$M = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}.$$

20.† Let $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$ be a linear transformation from an n-dimensional vector space V to an m-dimensional vector space W. Let β and γ be ordered bases for V and W, respectively. Prove that $\text{rank}(\mathsf{T}) = \text{rank}(\mathsf{L}_A)$ and that $\text{nullity}(\mathsf{T}) = \text{nullity}(\mathsf{L}_A)$, where $A = [\mathsf{T}]_\beta^\gamma$. *Hint:* Apply Exercise 17 to Figure 2.2.

21. Let V and W be finite-dimensional vector spaces with ordered bases $\beta = \{v_1, v_2, \ldots, v_n\}$ and $\gamma = \{w_1, w_2, \ldots, w_m\}$, respectively. By Theorem 2.6 (p. 73), there exist linear transformations $\mathsf{T}_{ij}\colon \mathsf{V} \to \mathsf{W}$ such that

$$\mathsf{T}_{ij}(v_k) = \begin{cases} w_i & \text{if } k = j \\ 0 & \text{if } k \neq j. \end{cases}$$

First prove that $\{\mathsf{T}_{ij}\colon 1 \le i \le m,\ 1 \le j \le n\}$ is a basis for $\mathcal{L}(\mathsf{V}, \mathsf{W})$. Then let M^{ij} be the $m \times n$ matrix with 1 in the ith row and jth column and 0 elsewhere, and prove that $[\mathsf{T}_{ij}]_\beta^\gamma = M^{ij}$. Again by Theorem 2.6, there exists a linear transformation $\Phi_\beta^\gamma\colon \mathcal{L}(\mathsf{V}, \mathsf{W}) \to \mathsf{M}_{m\times n}(F)$ such that $\Phi_\beta^\gamma(\mathsf{T}_{ij}) = M^{ij}$. Prove that Φ_β^γ is an isomorphism.

22. Let $c_0, c_1, \ldots, c_n$ be distinct scalars from an infinite field F. Define $\mathsf{T}\colon \mathsf{P}_n(F) \to \mathsf{F}^{n+1}$ by $\mathsf{T}(f) = (f(c_0), f(c_1), \ldots, f(c_n))$. Prove that T is an isomorphism. *Hint:* Use the Lagrange polynomials associated with $c_0, c_1, \ldots, c_n$.

23. Let W denote the vector space of all sequences in F that have only a finite number of nonzero terms (defined in Exercise 18 of Section 1.6), and let $\mathsf{Z} = \mathsf{P}(F)$. Define

$$\mathsf{T}\colon \mathsf{W} \to \mathsf{Z} \quad \text{by} \quad \mathsf{T}(\sigma) = \sum_{i=0}^{n} \sigma(i) x^i,$$

where n is the largest integer such that $\sigma(n) \neq 0$. Prove that T is an isomorphism.

The following exercise requires familiarity with the concept of *quotient space* defined in Exercise 31 of Section 1.3 and with Exercise 42 of Section 2.1.

24. Let V and Z be vector spaces and $\mathsf{T}\colon \mathsf{V} \to \mathsf{Z}$ be a linear transformation that is onto. Define the mapping

$$\overline{\mathsf{T}}\colon \mathsf{V}/\mathsf{N}(\mathsf{T}) \to \mathsf{Z} \quad \text{by} \quad \overline{\mathsf{T}}(v+\mathsf{N}(\mathsf{T})) = \mathsf{T}(v)$$

for any coset $v + \mathsf{N}(\mathsf{T})$ in $\mathsf{V}/\mathsf{N}(\mathsf{T})$.

(a) Prove that $\overline{\mathsf{T}}$ is well-defined; that is, prove that if $v + \mathsf{N}(\mathsf{T}) = v' + \mathsf{N}(\mathsf{T})$, then $\mathsf{T}(v) = \mathsf{T}(v')$.
(b) Prove that $\overline{\mathsf{T}}$ is linear.
(c) Prove that $\overline{\mathsf{T}}$ is an isomorphism.
(d) Prove that the diagram shown in Figure 2.3 commutes; that is, prove that $\mathsf{T} = \overline{\mathsf{T}}\eta$.

$$\begin{tikzcd}
\mathsf{V} \arrow[rr, dashed, "\mathsf{T}"] \arrow[dr, dashed, "\eta"'] & & \mathsf{Z} \\
& \mathsf{V}/\mathsf{N}(\mathsf{T}) \arrow[ur, dashed, "\overline{\mathsf{T}}"'] &
\end{tikzcd}$$

Figure 2.3

25. Let V be a nonzero vector space over a field F, and suppose that S is a basis for V. (By the corollary to Theorem 1.13 (p. 61) in Section 1.7, every vector space has a basis.) Let $\mathcal{C}(S,F)$ denote the vector space of all functions $f \in \mathcal{F}(S,F)$ such that $f(s) = 0$ for all but a finite number of vectors in S. (See Exercise 14 of Section 1.3.) Let $\Psi\colon \mathcal{C}(S,F) \to \mathsf{V}$ be defined by $\Psi(f) = \mathit{0}$ if f is the zero function, and

$$\Psi(f) = \sum_{s\in S,\, f(s)\neq 0} f(s)s,$$

otherwise. Prove that Ψ is an isomorphism. Thus every nonzero vector space can be viewed as a space of functions.

2.5 THE CHANGE OF COORDINATE MATRIX

In many areas of mathematics, a change of variable is used to simplify the appearance of an expression. For example, in calculus an antiderivative of $2xe^{x^2}$ can be found by making the change of variable $u = x^2$. The resulting expression is of such a simple form that an antiderivative is easily recognized:

$$\int 2xe^{x^2}\,dx = \int e^u\,du = e^u + c = e^{x^2} + c.$$

which is an ordered basis for R^3. Let Q be the 3×3 matrix whose jth column is the jth vector of γ. Then

$$Q = \begin{pmatrix} -1 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad Q^{-1} = \begin{pmatrix} -1 & 2 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}.$$

So by the preceding corollary,

$$[\mathsf{L}_A]_\gamma = Q^{-1}AQ = \begin{pmatrix} 0 & 2 & 8 \\ -1 & 4 & 6 \\ 0 & -1 & -1 \end{pmatrix}. \quad \blacklozenge$$

The relationship between the matrices $[\mathsf{T}]_{\beta'}$ and $[\mathsf{T}]_\beta$ in Theorem 2.23 will be the subject of further study in Chapters 5, 6, and 7. At this time, however, we introduce the name for this relationship.

Definition. *Let A and B be matrices in $\mathsf{M}_{n\times n}(F)$. We say that B is* ***similar*** *to A if there exists an invertible matrix Q such that $B = Q^{-1}AQ$.*

Observe that the relation of similarity is an equivalence relation (see Exercise 9). So we need only say that A and B are similar.

Notice also that in this terminology Theorem 2.23 can be stated as follows: If T is a linear operator on a finite-dimensional vector space V, and if β and β' are any ordered bases for V, then $[\mathsf{T}]_{\beta'}$ is similar to $[\mathsf{T}]_\beta$.

Theorem 2.23 can be generalized to allow $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$, where V is distinct from W. In this case, we can change bases in V as well as in W (see Exercise 8).

EXERCISES

1. Label the following statements as true or false.
   (a) Suppose that $\beta = \{x_1, x_2, \ldots, x_n\}$ and $\beta' = \{x'_1, x'_2, \ldots, x'_n\}$ are ordered bases for a vector space and Q is the change of coordinate matrix that changes β'-coordinates into β-coordinates. Then the jth column of Q is $[x_j]_{\beta'}$.
   (b) Every change of coordinate matrix is invertible.
   (c) Let T be a linear operator on a finite-dimensional vector space V, let β and β' be ordered bases for V, and let Q be the change of coordinate matrix that changes β'-coordinates into β-coordinates. Then $[\mathsf{T}]_\beta = Q[\mathsf{T}]_{\beta'}Q^{-1}$.
   (d) The matrices $A, B \in \mathsf{M}_{n\times n}(F)$ are called similar if $B = Q^tAQ$ for some $Q \in \mathsf{M}_{n\times n}(F)$.
   (e) Let T be a linear operator on a finite-dimensional vector space V. Then for any ordered bases β and γ for V, $[\mathsf{T}]_\beta$ is similar to $[\mathsf{T}]_\gamma$.

2. For each of the following pairs of ordered bases β and β' for R^2, find the change of coordinate matrix that changes β'-coordinates into β-coordinates.

(a) $\beta = \{e_1, e_2\}$ and $\beta' = \{(a_1, a_2), (b_1, b_2)\}$
(b) $\beta = \{(-1,3), (2,-1)\}$ and $\beta' = \{(0,10), (5,0)\}$
(c) $\beta = \{(2,5), (-1,-3)\}$ and $\beta' = \{e_1, e_2\}$
(d) $\beta = \{(-4,3), (2,-1)\}$ and $\beta' = \{(2,1), (-4,1)\}$

3. For each of the following pairs of ordered bases β and β' for $\mathsf{P}_2(R)$, find the change of coordinate matrix that changes β'-coordinates into β-coordinates.

(a) $\beta = \{x^2, x, 1\}$ and
$\beta' = \{a_2x^2 + a_1x + a_0, b_2x^2 + b_1x + b_0, c_2x^2 + c_1x + c_0\}$
(b) $\beta = \{1, x, x^2\}$ and
$\beta' = \{a_2x^2 + a_1x + a_0, b_2x^2 + b_1x + b_0, c_2x^2 + c_1x + c_0\}$
(c) $\beta = \{2x^2 - x, 3x^2 + 1, x^2\}$ and $\beta' = \{1, x, x^2\}$
(d) $\beta = \{x^2 - x + 1, x + 1, x^2 + 1\}$ and
$\beta' = \{x^2 + x + 4, 4x^2 - 3x + 2, 2x^2 + 3\}$
(e) $\beta = \{x^2 - x, x^2 + 1, x - 1\}$ and
$\beta' = \{5x^2 - 2x - 3, -2x^2 + 5x + 5, 2x^2 - x - 3\}$
(f) $\beta = \{2x^2 - x + 1, x^2 + 3x - 2, -x^2 + 2x + 1\}$ and
$\beta' = \{9x - 9, x^2 + 21x - 2, 3x^2 + 5x + 2\}$

4. Let T be the linear operator on R^2 defined by

$$\mathsf{T}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 2a + b \\ a - 3b \end{pmatrix},$$

let β be the standard ordered basis for R^2, and let

$$\beta' = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix} \right\}.$$

Use Theorem 2.23 and the fact that

$$\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}^{-1} = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}$$

to find $[\mathsf{T}]_{\beta'}$.

5. Let T be the linear operator on $\mathsf{P}_1(R)$ defined by $\mathsf{T}(p(x)) = p'(x)$, the derivative of $p(x)$. Let $\beta = \{1, x\}$ and $\beta' = \{1 + x, 1 - x\}$. Use Theorem 2.23 and the fact that

$$\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}^{-1} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix}$$

to find $[\mathsf{T}]_{\beta'}$.

6. For each matrix A and ordered basis β, find $[\mathsf{L}_A]_\beta$. Also, find an invertible matrix Q such that $[\mathsf{L}_A]_\beta = Q^{-1}AQ$.

   (a) $A = \begin{pmatrix} 1 & 3 \\ 1 & 1 \end{pmatrix}$ and $\beta = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix} \right\}$

   (b) $A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ and $\beta = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}$

   (c) $A = \begin{pmatrix} 1 & 1 & -1 \\ 2 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$ and $\beta = \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} \right\}$

   (d) $A = \begin{pmatrix} 13 & 1 & 4 \\ 1 & 13 & 4 \\ 4 & 4 & 10 \end{pmatrix}$ and $\beta = \left\{ \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\}$

7. In R^2, let L be the line $y = mx$, where $m \neq 0$. Find an expression for $\mathsf{T}(x, y)$, where

   (a) T is the reflection of R^2 about L.
   (b) T is the projection on L along the line perpendicular to L. (See the definition of *projection* in the exercises of Section 2.1.)

8. Prove the following generalization of Theorem 2.23. Let $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$ be a linear transformation from a finite-dimensional vector space V to a finite-dimensional vector space W. Let β and β' be ordered bases for V, and let γ and γ' be ordered bases for W. Then $[\mathsf{T}]_{\beta'}^{\gamma'} = P^{-1}[\mathsf{T}]_\beta^\gamma Q$, where Q is the matrix that changes β'-coordinates into β-coordinates and P is the matrix that changes γ'-coordinates into γ-coordinates.

9. Prove that "is similar to" is an equivalence relation on $\mathsf{M}_{n\times n}(F)$.

10. (a) Prove that if A and B are similar $n \times n$ matrices, then $\operatorname{tr}(A) = \operatorname{tr}(B)$. *Hint:* Use Exercise 13 of Section 2.3.
    (b) How would you define the trace of a linear operator on a finite-dimensional vector space? Justify that your definition is well-defined.

11. Let V be a finite-dimensional vector space with ordered bases α, β, and γ.

    (a) Prove that if Q and R are the change of coordinate matrices that change α-coordinates into β-coordinates and β-coordinates into γ-coordinates, respectively, then RQ is the change of coordinate matrix that changes α-coordinates into γ-coordinates.
    (b) Prove that if Q changes α-coordinates into β-coordinates, then Q^{-1} changes β-coordinates into α-coordinates.

12. Prove the corollary to Theorem 2.23.

13.[†] Let V be a finite-dimensional vector space over a field F, and let $\beta = \{x_1, x_2, \ldots, x_n\}$ be an ordered basis for V. Let Q be an $n \times n$ invertible matrix with entries from F. Define

$$x'_j = \sum_{i=1}^{n} Q_{ij} x_i \quad \text{for } 1 \leq j \leq n,$$

and set $\beta' = \{x'_1, x'_2, \ldots, x'_n\}$. Prove that β' is a basis for V and hence that Q is the change of coordinate matrix changing β'-coordinates into β-coordinates. Visit `goo.gl/vsxsGH` for a solution.

14. Prove the converse of Exercise 8: If A and B are each $m \times n$ matrices with entries from a field F, and if there exist invertible $m \times m$ and $n \times n$ matrices P and Q, respectively, such that $B = P^{-1}AQ$, then there exist an n-dimensional vector space V and an m-dimensional vector space W (both over F), ordered bases β and β' for V and γ and γ' for W, and a linear transformation $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$ such that

$$A = [\mathsf{T}]_\beta^\gamma \quad \text{and} \quad B = [\mathsf{T}]_{\beta'}^{\gamma'}.$$

Hints: Let $\mathsf{V} = \mathsf{F}^n$, $\mathsf{W} = \mathsf{F}^m$, $\mathsf{T} = \mathsf{L}_A$, and β and γ be the standard ordered bases for F^n and F^m, respectively. Now apply the results of Exercise 13 to obtain ordered bases β' and γ' from β and γ via Q and P, respectively.

2.6* DUAL SPACES

In this section, we are concerned exclusively with linear transformations from a vector space V into its field of scalars F, which is itself a vector space of dimension 1 over F. Such a linear transformation is called a **linear functional** on V. We generally use the letters $\mathsf{f}, \mathsf{g}, \mathsf{h}, \ldots$ to denote linear functionals. As we see in Example 1, the definite integral provides us with one of the most important examples of a linear functional in mathematics.

Example 1

Let V be the vector space of continuous real-valued functions on the interval $[0, 2\pi]$. Fix a function $g \in \mathsf{V}$. The function $\mathsf{h}\colon \mathsf{V} \to R$ defined by

$$\mathsf{h}(x) = \frac{1}{2\pi} \int_0^{2\pi} x(t)g(t)\,dt$$

is a linear functional on V. In the cases that $g(t)$ equals $\sin nt$ or $\cos nt$, $h(x)$ is often called the **nth Fourier coefficient of x.** ◆

Although many of the ideas of this section (e.g., the existence of a dual space) can be extended to the case where V is not finite-dimensional, only a finite-dimensional vector space is isomorphic to its double dual via the map $x \to \widehat{x}$. In fact, for infinite-dimensional vector spaces, no two of V, V*, and V** are isomorphic.

EXERCISES

1. Label the following statements as true or false. Assume that all vector spaces are finite-dimensional.
   (a) Every linear transformation is a linear functional.
   (b) A linear functional defined on a field may be represented as a 1×1 matrix.
   (c) Every vector space is isomorphic to its dual space.
   (d) Every vector space is isomorphic to the dual of some vector space.
   (e) If T is an isomorphism from V onto V* and β is a finite ordered basis for V, then $\mathsf{T}(\beta) = \beta^*$.
   (f) If T is a linear transformation from V to W, then the domain of $(\mathsf{T}^t)^t$ is V^{**}.
   (g) If V is isomorphic to W, then V* is isomorphic to W*.
   (h) The derivative of a function may be considered as a linear functional on the vector space of differentiable functions.

2. For the following functions f on a vector space V, determine which are linear functionals.
   (a) $\mathsf{V} = \mathsf{P}(R)$; $\mathsf{f}(p(x)) = 2p'(0) + p''(1)$, where $'$ denotes differentiation
   (b) $\mathsf{V} = \mathsf{R}^2$; $\mathsf{f}(x, y) = (2x, 4y)$
   (c) $\mathsf{V} = \mathsf{M}_{2\times 2}(F)$; $\mathsf{f}(A) = \operatorname{tr}(A)$
   (d) $\mathsf{V} = \mathsf{R}^3$; $\mathsf{f}(x, y, z) = x^2 + y^2 + z^2$
   (e) $\mathsf{V} = \mathsf{P}(R)$; $\mathsf{f}(p(x)) = \int_0^1 p(t)\,dt$
   (f) $\mathsf{V} = \mathsf{M}_{2\times 2}(F)$; $\mathsf{f}(A) = A_{11}$

3. For each of the following vector spaces V and bases β, find explicit formulas for vectors of the dual basis β^* for V*, as in Example 4.
   (a) $\mathsf{V} = \mathsf{R}^3$; $\beta = \{(1,0,1), (1,2,1), (0,0,1)\}$
   (b) $\mathsf{V} = \mathsf{P}_2(R)$; $\beta = \{1, x, x^2\}$

4. Let $\mathsf{V} = \mathsf{R}^3$, and define $\mathsf{f}_1, \mathsf{f}_2, \mathsf{f}_3 \in \mathsf{V}^*$ as follows:

$$\mathsf{f}_1(x,y,z) = x - 2y, \quad \mathsf{f}_2(x,y,z) = x + y + z, \quad \mathsf{f}_3(x,y,z) = y - 3z.$$

   Prove that $\{\mathsf{f}_1, \mathsf{f}_2, \mathsf{f}_3\}$ is a basis for V*, and then find a basis for V for which it is the dual basis.

5. Let $V = P_1(R)$, and, for $p(x) \in V$, define $f_1, f_2 \in V^*$ by

$$f_1(p(x)) = \int_0^1 p(t)\,dt \quad \text{and} \quad f_2(p(x)) = \int_0^2 p(t)\,dt.$$

Prove that $\{f_1, f_2\}$ is a basis for V^*, and find a basis for V for which it is the dual basis.

6. Define $f \in (R^2)^*$ by $f(x, y) = 2x + y$ and $T\colon R^2 \to R^2$ by $T(x, y) = (3x + 2y, x)$.

(a) Compute $T^t(f)$.
(b) Compute $[T^t]_{\beta^*}$, where β is the standard ordered basis for R^2 and $\beta^* = \{f_1, f_2\}$ is the dual basis, by finding scalars a, b, c, and d such that $T^t(f_1) = af_1 + cf_2$ and $T^t(f_2) = bf_1 + df_2$.
(c) Compute $[T]_\beta$ and $([T]_\beta)^t$, and compare your results with (b).

7. Let $V = P_1(R)$ and $W = R^2$ with respective standard ordered bases β and γ. Define $T\colon V \to W$ by

$$T(p(x)) = (p(0) - 2p(1), p(0) + p'(0)),$$

where $p'(x)$ is the derivative of $p(x)$.

(a) For $f \in W^*$ defined by $f(a, b) = a - 2b$, compute $T^t(f)$.
(b) Compute $[T^t]_{\gamma^*}^{\beta^*}$ without appealing to Theorem 2.25.
(c) Compute $[T]_\beta^\gamma$ and its transpose, and compare your results with (b).

8. Let $\{u, v\}$ be a linearly independent set in R^3. Show that the plane $\{su + tv\colon s, t \in R\}$ through the origin in R^3 may be identified with the null space of a vector in $(R^3)^*$.

9. Prove that a function $T\colon F^n \to F^m$ is linear if and only if there exist $f_1, f_2, \ldots, f_m \in (F^n)^*$ such that $T(x) = (f_1(x), f_2(x), \ldots, f_m(x))$ for all $x \in F^n$. *Hint:* If T is linear, define $f_i(x) = (g_i T)(x)$ for $x \in F^n$; that is, $f_i = T^t(g_i)$ for $1 \le i \le m$, where $\{g_1, g_2, \ldots, g_m\}$ is the dual basis of the standard ordered basis for F^m.

10. Let $V = P_n(F)$, and let $c_0, c_1, \ldots, c_n$ be distinct scalars in F.

(a) For $0 \le i \le n$, define $f_i \in V^*$ by $f_i(p(x)) = p(c_i)$. Prove that $\{f_0, f_1, \ldots, f_n\}$ is a basis for V^*. *Hint:* Apply any linear combination of this set that equals the zero transformation to $p(x) = (x - c_1)(x - c_2)\cdots(x - c_n)$, and deduce that the first coefficient is zero.

(b) Use the corollary to Theorem 2.26 and (a) to show that there exist unique polynomials $p_0(x), p_1(x), \ldots, p_n(x)$ such that $p_i(c_j) = \delta_{ij}$ for $0 \leq i \leq n$. These polynomials are the Lagrange polynomials defined in Section 1.6.

(c) For any scalars $a_0, a_1, \ldots, a_n$ (not necessarily distinct), deduce that there exists a unique polynomial $q(x)$ of degree at most n such that $q(c_i) = a_i$ for $0 \leq i \leq n$. In fact,

$$q(x) = \sum_{i=0}^{n} a_i p_i(x).$$

(d) Deduce the Lagrange interpolation formula:

$$p(x) = \sum_{i=0}^{n} p(c_i) p_i(x)$$

for any $p(x) \in \mathsf{V}$.

(e) Prove that

$$\int_a^b p(t)\,dt = \sum_{i=0}^{n} p(c_i) d_i,$$

where

$$d_i = \int_a^b p_i(t)\,dt.$$

Suppose now that

$$c_i = a + \frac{i(b-a)}{n} \quad \text{for } i = 0, 1, \ldots, n.$$

For $n = 1$, the preceding result yields the trapezoidal rule for evaluating the definite integral of a polynomial. For $n = 2$, this result yields Simpson's rule for evaluating the definite integral of a polynomial.

$$\begin{tikzcd}
\mathsf{V} \arrow[r, "\mathsf{T}"] \arrow[d, "\psi_1"'] & \mathsf{W} \arrow[d, "\psi_2"] \\
\mathsf{V}^{**} \arrow[r, "\mathsf{T}^{tt}"] & \mathsf{W}^{**}
\end{tikzcd}$$

Figure 2.6

$\boxed{11.}$ Let V and W be finite-dimensional vector spaces over F, and let ψ_1 and ψ_2 be the isomorphisms between V and V^{**} and W and W^{**}, respectively, as defined in Theorem 2.26. Let $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$ be linear, and define

$T^{tt} = (T^t)^t$. Prove that the diagram depicted in Figure 2.6 commutes (i.e., prove that $\psi_2 T = T^{tt}\psi_1$). Visit goo.gl/Lkd6XZ for a solution.

12. Let V be a finite-dimensional vector space with the ordered basis β. Prove that $\psi(\beta) = \beta^{**}$, where ψ is defined in Theorem 2.26.

In Exercises 13 through 17, V denotes a finite-dimensional vector space over F. For every subset S of V, define the **annihilator** S^0 of S as

$$S^0 = \{f \in V^*: f(x) = 0 \text{ for all } x \in S\}.$$

13. (a) Prove that S^0 is a subspace of V^*.
    (b) If W is a subspace of V and $x \notin W$, prove that there exists $f \in W^0$ such that $f(x) \neq 0$.
    (c) Prove that $(S^0)^0 = \text{span}(\psi(S))$, where ψ is defined as in Theorem 2.26.
    (d) For subspaces W_1 and W_2, prove that $W_1 = W_2$ if and only if $W_1^0 = W_2^0$.
    (e) For subspaces W_1 and W_2, show that $(W_1 + W_2)^0 = W_1^0 \cap W_2^0$.

14. Prove that if W is a subspace of V, then $\dim(W) + \dim(W^0) = \dim(V)$. *Hint:* Extend an ordered basis $\{x_1, x_2, \ldots, x_k\}$ of W to an ordered basis $\beta = \{x_1, x_2, \ldots, x_n\}$ of V. Let $\beta^* = \{f_1, f_2, \ldots, f_n\}$. Prove that $\{f_{k+1}, f_{k+2}, \ldots, f_n\}$ is a basis for W^0.

15. Suppose that W is a finite-dimensional vector space and that $T: V \to W$ is linear. Prove that $N(T^t) = (R(T))^0$.

16. Use Exercises 14 and 15 to deduce that $\text{rank}(L_{A^t}) = \text{rank}(L_A)$ for any $A \in M_{m\times n}(F)$.

In Exercises 17 through 20, assume that V and W are finite-dimensional vector spaces. (It can be shown, however, that these exercises are true for all vector spaces V and W.)

17. Let T be a linear operator on V, and let W be a subspace of V. Prove that W is T-invariant (as defined in the exercises of Section 2.1) if and only if W^0 is T^t-invariant.

18. Let V be a nonzero vector space over a field F, and let S be a basis for V. (By the corollary to Theorem 1.13 (p. 61) in Section 1.7, every vector space has a basis.) Let $\Phi: V^* \to \mathcal{F}(S, F)$ be the mapping defined by $\Phi(f) = f_S$, the restriction of f to S. Prove that Φ is an isomorphism. *Hint:* Apply Exercise 35 of Section 2.1.

19. Let V be a nonzero vector space, and let W be a proper subspace of V (i.e., $W \neq V$).

(a) Let $g \in W^*$ and $v \in V$ with $v \notin W$. Prove that for any scalar a there exists a function $f \in V^*$ such that $f(v) = a$ and $f(x) = g(x)$ for all x in W. *Hint:* For the infinite-dimensional case, use Exercise 4 of Section 1.7 and Exercise 35 of Section 2.1.

(b) Use (a) to prove there exists a nonzero linear functional $f \in V^*$ such that $f(x) = 0$ for all $x \in W$.

20. Let V and W be nonzero vector spaces over the same field, and let $T: V \to W$ be a linear transformation.

(a) Prove that T is onto if and only if T^t is one-to-one.

(b) Prove that T^t is onto if and only if T is one-to-one.

Hint: In the infinite-dimensional case, use Exercise 19 for parts of the proof.

2.7* HOMOGENEOUS LINEAR DIFFERENTIAL EQUATIONS WITH CONSTANT COEFFICIENTS

As an introduction to this section, consider the following physical problem. A weight of mass m is attached to a vertically suspended spring that is allowed to stretch until the forces acting on the weight are in equilibrium. Suppose that the weight is now motionless and impose an xy-coordinate system with the weight at the origin and the spring lying on the positive y-axis (see Figure 2.7).

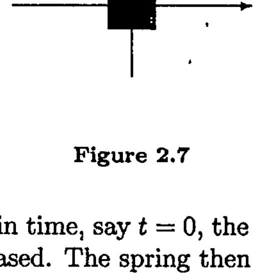


Figure 2.7

Suppose that at a certain time, say $t = 0$, the weight is lowered a distance s along the y-axis and released. The spring then begins to oscillate.

We describe the motion of the spring. At any time $t \geq 0$, let $F(t)$ denote the force acting on the weight and $y(t)$ denote the position of the weight along the y-axis. For example, $y(0) = -s$. The second derivative of y with respect

The most general situation is stated in the following theorem.

Theorem 2.34. *Given a homogeneous linear differential equation with constant coefficients and auxiliary polynomial*

$$(t-c_1)^{n_1}(t-c_2)^{n_2}\cdots(t-c_k)^{n_k},$$

where $n_1, n_2, \ldots, n_k$ are positive integers and $c_1, c_2, \ldots, c_k$ are distinct complex numbers, the following set is a basis for the solution space of the equation:

$$\{e^{c_1t}, te^{c_1t}, \ldots, t^{n_1-1}e^{c_1t}, \ldots, e^{c_kt}, te^{c_kt}, \ldots, t^{n_k-1}e^{c_kt}\}.$$

Proof. Exercise. ∎

Example 8

The differential equation

$$y^{(3)} - 4y^{(2)} + 5y^{(1)} - 2y = 0$$

has the auxiliary polynomial

$$t^3 - 4t^2 + 5t - 2 = (t-1)^2(t-2).$$

By Theorem 2.34, $\{e^t, te^t, e^{2t}\}$ is a basis for the solution space of the differential equation. Thus any solution y has the form

$$y(t) = b_1e^t + b_2te^t + b_3e^{2t}$$

for unique scalars b_1, b_2, and b_3. ◆

EXERCISES

1. Label the following statements as true or false.
   (a) The set of solutions to an nth-order homogeneous linear differential equation with constant coefficients is an n-dimensional subspace of C^∞.
   (b) The solution space of a homogeneous linear differential equation with constant coefficients is the null space of a differential operator.
   (c) The auxiliary polynomial of a homogeneous linear differential equation with constant coefficients is a solution to the differential equation.
   (d) Any solution to a homogeneous linear differential equation with constant coefficients is of the form ae^{ct} or at^ke^{ct}, where a and c are complex numbers and k is a positive integer.

(e) Any linear combination of solutions to a given homogeneous linear differential equation with constant coefficients is also a solution to the given equation.
(f) For any homogeneous linear differential equation with constant coefficients having auxiliary polynomial $p(t)$, if $c_1, c_2, \ldots, c_k$ are the distinct zeros of $p(t)$, then $\{e^{c_1 t}, e^{c_2 t}, \ldots, e^{c_k t}\}$ is a basis for the solution space of the given differential equation.
(g) Given any polynomial $p(t) \in \mathsf{P}(C)$, there exists a homogeneous linear differential equation with constant coefficients whose auxiliary polynomial is $p(t)$.

2. For each of the following parts, determine whether the statement is true or false. Justify your claim with either a proof or a counterexample, whichever is appropriate.

(a) Any finite-dimensional subspace of C^∞ is the solution space of a homogeneous linear differential equation with constant coefficients.
(b) There exists a homogeneous linear differential equation with constant coefficients whose solution space has the basis $\{t, t^2\}$.
(c) For any homogeneous linear differential equation with constant coefficients, if x is a solution to the equation, so is its derivative x'.

Given two polynomials $p(t)$ and $q(t)$ in $\mathsf{P}(C)$, if $x \in \mathsf{N}(p(\mathsf{D}))$ and $y \in \mathsf{N}(q(\mathsf{D}))$, then

(d) $x + y \in \mathsf{N}(p(\mathsf{D})q(\mathsf{D}))$.
(e) $xy \in \mathsf{N}(p(\mathsf{D})q(\mathsf{D}))$.

3. Find a basis for the solution space of each of the following differential equations.

(a) $y'' + 2y' + y = 0$
(b) $y''' = y'$
(c) $y^{(4)} - 2y^{(2)} + y = 0$
(d) $y'' + 2y' + y = 0$
(e) $y^{(3)} - y^{(2)} + 3y^{(1)} + 5y = 0$

4. Find a basis for each of the following subspaces of C^∞.

(a) $\mathsf{N}(\mathsf{D}^2 - \mathsf{D} - \mathsf{I})$
(b) $\mathsf{N}(\mathsf{D}^3 - 3\mathsf{D}^2 + 3\mathsf{D} - \mathsf{I})$
(c) $\mathsf{N}(\mathsf{D}^3 + 6\mathsf{D}^2 + 8\mathsf{D})$

5. Show that C^∞ is a subspace of $\mathcal{F}(R, C)$.

6. (a) Show that $\mathsf{D}\colon \mathsf{C}^\infty \to \mathsf{C}^\infty$ is a linear operator.
(b) Show that any differential operator is a linear operator on C^∞.

7. Prove that if $\{x, y\}$ is a basis for a vector space over C, then so is

$$\left\{\frac{1}{2}(x+y), \frac{1}{2i}(x-y)\right\}.$$

8. Consider a second-order homogeneous linear differential equation with constant coefficients in which the auxiliary polynomial has distinct conjugate complex roots $a+ib$ and $a-ib$, where $a, b \in R$. Show that $\{e^{at}\cos bt, e^{at}\sin bt\}$ is a basis for the solution space.

9. Suppose that $\{\mathsf{U}_1, \mathsf{U}_2, \ldots, \mathsf{U}_n\}$ is a collection of pairwise commutative linear operators on a vector space V (i.e., operators such that $\mathsf{U}_i\mathsf{U}_j = \mathsf{U}_j\mathsf{U}_i$ for all i, j). Prove that, for any i $(1 \le i \le n)$,

$$\mathsf{N}(\mathsf{U}_i) \subseteq \mathsf{N}(\mathsf{U}_1\mathsf{U}_2\cdots\mathsf{U}_n).$$

10. Prove Theorem 2.33 and its corollary. *Hints:* For Theorem 2.33, use mathematical induction on n. In the inductive step, let $a_1, a_2, \ldots, a_n$ be scalars such that $\sum_{i=1}^{n} a_i e^{c_i t} = 0$. Multiply both sides of this equation by $e^{-c_n t}$, and differentiate the resulting equation with respect to t. For the corollary, use Theorems 2.31, 2.33, and 2.32. Visit goo.gl/oKTEbV for a solution.

11. Prove Theorem 2.34. *Hint:* First verify that the alleged basis lies in the solution space. Then verify that this set is linearly independent by mathematical induction on k as follows. The case $k = 1$ is the lemma to Theorem 2.34. Assuming that the theorem holds for $k-1$ distinct c_i's, apply the operator $(\mathsf{D} - c_k\mathsf{I})^{n_k}$ to any linear combination of the alleged basis that equals 0.

12. Let V be the solution space of an nth-order homogeneous linear differential equation with constant coefficients having auxiliary polynomial $p(t)$. Prove that if $p(t) = g(t)h(t)$, where $g(t)$ and $h(t)$ are polynomials of positive degree, then

$$\mathsf{N}(h(\mathsf{D})) = \mathsf{R}(g(\mathsf{D}_\mathsf{V})) = g(\mathsf{D})(\mathsf{V}),$$

where $\mathsf{D}_\mathsf{V}\colon \mathsf{V} \to \mathsf{V}$ is defined by $\mathsf{D}_\mathsf{V}(x) = x'$ for $x \in \mathsf{V}$. *Hint:* First prove $g(\mathsf{D})(\mathsf{V}) \subseteq \mathsf{N}(h(\mathsf{D}))$. Then prove that the two spaces have the same finite dimension.

13. A differential equation

$$y^{(n)} + a_{n-1}y^{(n-1)} + \cdots + a_1 y^{(1)} + a_0 y = x$$

is called a **nonhomogeneous** linear differential equation with constant coefficients if the a_i's are constant and x is a function that is not identically zero.

(a) Prove that for any $x \in \mathsf{C}^\infty$ there exists $y \in \mathsf{C}^\infty$ such that y is a solution to the differential equation. *Hint:* Use Lemma 1 to Theorem 2.32 to show that for any polynomial $p(t)$, the linear operator $p(\mathsf{D})\colon \mathsf{C}^\infty \to \mathsf{C}^\infty$ is onto.

(b) Let V be the solution space for the homogeneous linear equation

$$y^{(n)} + a_{n-1}y^{(n-1)} + \cdots + a_1 y^{(1)} + a_0 y = 0.$$

Prove that if z is any solution to the associated nonhomogeneous linear differential equation, then the set of all solutions to the nonhomogeneous linear differential equation is

$$\{z + y \colon y \in \mathsf{V}\}.$$

14. Given any nth-order homogeneous linear differential equation with constant coefficients, prove that, for any solution x and any $t_0 \in R$, if $x(t_0) = x'(t_0) = \cdots = x^{(n-1)}(t_0) = 0$, then $x = 0$ (the zero function). *Hint:* Use mathematical induction on n as follows. First prove the conclusion for the case $n = 1$. Next suppose that it is true for equations of order $n - 1$, and consider an nth-order differential equation with auxiliary polynomial $p(t)$. Factor $p(t) = q(t)(t - c)$, and let $z = q((D))x$. Show that $z(t_0) = 0$ and $z' - cz = 0$ to conclude that $z = 0$. Now apply the induction hypothesis.

15. Let V be the solution space of an nth-order homogeneous linear differential equation with constant coefficients. Fix $t_0 \in R$, and define a mapping $\Phi\colon \mathsf{V} \to \mathsf{C}^n$ by

$$\Phi(x) = \begin{pmatrix} x(t_0) \\ x'(t_0) \\ \vdots \\ x^{(n-1)}(t_0) \end{pmatrix} \quad \text{for each } x \in \mathsf{V}.$$

(a) Prove that Φ is linear and its null space is the zero subspace of V. Deduce that Φ is an isomorphism. *Hint:* Use Exercise 14.

(b) Prove the following: For any nth-order homogeneous linear differential equation with constant coefficients, any $t_0 \in R$, and any complex numbers $c_0, c_1, \ldots, c_{n-1}$ (not necessarily distinct), there exists exactly one solution, x, to the given differential equation such that $x(t_0) = c_0$ and $x^{(k)}(t_0) = c_k$ for $k = 1, 2, \ldots, n-1$.

16. *Pendular Motion.* It is well known that the motion of a pendulum is approximated by the differential equation

$$\theta'' + \frac{g}{l}\theta = 0,$$

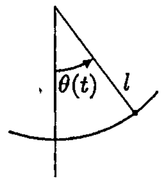


Figure 2.8

where $\theta(t)$ is the angle in radians that the pendulum makes with a vertical line at time t (see Figure 2.8), interpreted so that θ is positive if the pendulum is to the right and negative if the pendulum is to the left of the vertical line as viewed by the reader. Here l is the length of the pendulum and g is the magnitude of acceleration due to gravity. The variable t and constants l and g must be in compatible units (e.g., t in seconds, l in meters, and g in meters per second per second).

(a) Express an arbitrary solution to this equation as a linear combination of two real-valued solutions.

(b) Find the unique solution to the equation that satisfies the conditions

$$\theta(0) = \theta_0 > 0 \quad \text{and} \quad \theta'(0) = 0.$$

(The significance of these conditions is that at time $t = 0$ the pendulum is released from a position displaced from the vertical by θ_0.)

(c) Prove that it takes $2\pi\sqrt{l/g}$ units of time for the pendulum to make one circuit back and forth. (This time is called the **period** of the pendulum.)

17. *Periodic Motion of a Spring without Damping.* Find the general solution to (3), which describes the periodic motion of a spring, ignoring frictional forces.

18. *Periodic Motion of a Spring with Damping.* The ideal periodic motion described by solutions to (3) is due to the ignoring of frictional forces. In reality, however, there is a frictional force acting on the motion that is proportional to the speed of motion, but that acts in the opposite direction. The modification of (3) to account for the frictional force, called the *damping force*, is given by

$$my'' + ry' + ky = 0,$$

where $r > 0$ is the proportionality constant.

(a) Find the general solution to this equation.
(b) Find the unique solution in (a) that satisfies the initial conditions $y(0) = 0$ and $y'(0) = v_0$, the initial velocity.
(c) For $y(t)$ as in (b), show that the amplitude of the oscillation decreases to zero; that is, prove that $\lim_{t\to\infty} y(t) = 0$.

19. In our study of differential equations, we have regarded solutions as complex-valued functions even though functions that are useful in describing physical motion are real-valued. Justify this approach.

20. The following parts, which do not involve linear algebra, are included for the sake of completeness.

(a) Prove Theorem 2.27. *Hint:* Use mathematical induction on the number of derivatives possessed by a solution.
(b) For any $c, d \in C$, prove that

$$e^{c+d} = c^c e^d \quad \text{and} \quad e^{-c} = \frac{1}{e^c}.$$

(c) Prove Theorem 2.28.
(d) Prove Theorem 2.29.
(e) Prove the product rule for differentiating complex-valued functions of a real variable: For any differentiable functions x and y in $\mathcal{F}(R, C)$, the product xy is differentiable and

$$(xy)' = x'y + xy'.$$

Hint: Apply the rules of differentiation to the real and imaginary parts of xy.
(f) Prove that if $x \in \mathcal{F}(R, C)$ and $x' = 0$, then x is a constant function.

INDEX OF DEFINITIONS FOR CHAPTER 2

Example 2

Consider the matrices A and B in Example 1. In this case, B is obtained from A by interchanging the first two rows of A. Performing this same operation on I_3, we obtain the elementary matrix

$$E = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Note that $EA = B$.

In the second part of Example 1, C is obtained from A by multiplying the second column of A by 3. Performing this same operation on I_4, we obtain the elementary matrix

$$E = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Observe that $AE = C$. ◆

It is a useful fact that the inverse of an elementary matrix is also an elementary matrix.

Theorem 3.2. *Elementary matrices are invertible, and the inverse of an elementary matrix is an elementary matrix of the same type.*

Proof. Let E be an elementary $n \times n$ matrix. Then E can be obtained by an elementary row operation on I_n. By reversing the steps used to transform I_n into E, we can transform E back into I_n. The result is that I_n can be obtained from E by an elementary row operation of the same type. By Theorem 3.1, there is an elementary matrix $\overline{E}$ such that $\overline{E}E = I_n$. Therefore, by Exercise 10 of Section 2.4, E is invertible and $E^{-1} = \overline{E}$. ∎

EXERCISES

1. Label the following statements as true or false.
   (a) An elementary matrix is always square.
   (b) The only entries of an elementary matrix are zeros and ones.
   (c) The $n \times n$ identity matrix is an elementary matrix.
   (d) The product of two $n \times n$ elementary matrices is an elementary matrix.
   (e) The inverse of an elementary matrix is an elementary matrix.
   (f) The sum of two $n \times n$ elementary matrices is an elementary matrix.
   (g) The transpose of an elementary matrix is an elementary matrix.

(h) If B is a matrix that can be obtained by performing an elementary row operation on a matrix A, then B can also be obtained by performing an elementary column operation on A.

(i) If B is a matrix that can be obtained by performing an elementary row operation on a matrix A, then A can be obtained by performing an elementary row operation on B.

2. Let

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 0 & 1 \\ 1 & -1 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 & 3 \\ 1 & -2 & 1 \\ 1 & -3 & 1 \end{pmatrix}, \quad \text{and } C = \begin{pmatrix} 1 & 0 & 3 \\ 0 & -2 & -2 \\ 1 & -3 & 1 \end{pmatrix}.$$

Find an elementary operation that transforms A into B and an elementary operation that transforms B into C. By means of several additional operations, transform C into I_3.

3. Use the proof of Theorem 3.2 to obtain the inverse of each of the following elementary matrices.

(a) $\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$ (b) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ (c) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix}$

4. Prove the assertion made on page 149: Any elementary $n \times n$ matrix can be obtained in at least two ways—either by performing an elementary row operation on I_n or by performing an elementary column operation on I_n.

5. Prove that E is an elementary matrix if and only if E^t is.

6. Let A be an $m \times n$ matrix. Prove that if B can be obtained from A by an elementary row [column] operation, then B^t can be obtained from A^t by the corresponding elementary column [row] operation.

7. Prove Theorem 3.1.

8. Prove that if a matrix Q can be obtained from a matrix P by an elementary row operation, then P can be obtained from Q by an elementary row operation of the same type. *Hint:* Treat each type of elementary row operation separately.

$\boxed{9.}$ Prove that any elementary row [column] operation of type 1 can be obtained by a succession of three elementary row [column] operations of type 3 followed by one elementary row [column] operation of type 2. Visit `goo.gl/oNJBFz` for a solution.

10. Prove that any elementary row [column] operation of type 2 can be obtained by *dividing* some row [column] by a nonzero scalar.

11. Prove that any elementary row [column] operation of type 3 can be obtained by *subtracting* a multiple of some row [column] from another row [column].

12. Let A be an $m \times n$ matrix. Prove that there exists a sequence of elementary row operations of types 1 and 3 that transforms A into an upper triangular matrix.

3.2 THE RANK OF A MATRIX AND MATRIX INVERSES

In this section, we define the *rank* of a matrix. We then use elementary operations to compute the rank of a matrix and a linear transformation. The section concludes with a procedure for computing the inverse of an invertible matrix.

Definition. *If $A \in \mathsf{M}_{m\times n}(F)$, we define the* **rank** *of A, denoted* $\text{rank}(A)$, *to be the rank of the linear transformation $\mathsf{L}_A \colon \mathsf{F}^n \to \mathsf{F}^m$.*

Many results about the rank of a matrix follow immediately from the corresponding facts about a linear transformation. An important result of this type, which follows from Fact 3 (p. 101) and Corollary 2 to Theorem 2.18 (p. 103), is that *an $n \times n$ matrix is invertible if and only if its rank is n.*

Every matrix A is the matrix representation of the linear transformation L_A with respect to the appropriate standard ordered bases. Thus the rank of the linear transformation L_A is the same as the rank of one of its matrix representations, namely, A. The next theorem extends this fact to any matrix representation of any linear transformation defined on finite-dimensional vector spaces.

Theorem 3.3. *Let $\mathsf{T} \colon \mathsf{V} \to \mathsf{W}$ be a linear transformation between finite-dimensional vector spaces, and let β and γ be ordered bases for V and W, respectively. Then* $\text{rank}(\mathsf{T}) = \text{rank}([\mathsf{T}]_\beta^\gamma)$.

Proof. This is a restatement of Exercise 20 of Section 2.4. ∎

Now that the problem of finding the rank of a linear transformation has been reduced to the problem of finding the rank of a matrix, we need a result that allows us to perform rank-preserving operations on matrices. The next theorem and its corollary tell us how to do this.

Theorem 3.4. *Let A be an $m \times n$ matrix. If P and Q are invertible $m \times m$ and $n \times n$ matrices, respectively, then*

(a) $\text{rank}(AQ) = \text{rank}(A)$,

(b) $\text{rank}(PA) = \text{rank}(A)$,

and therefore,

is a matrix with a row whose first 3 entries are zeros. Therefore A is not invertible. ♦

Being able to test for invertibility and compute the inverse of a matrix allows us, with the help of Theorem 2.18 (p. 102) and its corollaries, to test for invertibility and compute the inverse of a linear transformation. The next example demonstrates this technique.

Example 7

Let $\mathsf{T}\colon \mathsf{P}_2(R) \to \mathsf{P}_2(R)$ be defined by $\mathsf{T}(f(x)) = f(x) + f'(x) + f''(x)$, where $f'(x)$ and $f''(x)$ denote the first and second derivatives of $f(x)$. We use Corollary 1 of Theorem 2.18 (p. 103) to test T for invertibility and compute the inverse if T is invertible. Taking β to be the standard ordered basis of $\mathsf{P}_2(R)$, we have

$$[\mathsf{T}]_\beta = \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}.$$

Using the method of Examples 5 and 6, we can show that $[\mathsf{T}]_\beta$ is invertible with inverse

$$([\mathsf{T}]_\beta)^{-1} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix}.$$

Thus T is invertible, and $([\mathsf{T}]_\beta)^{-1} = [\mathsf{T}^{-1}]_\beta$. Hence by Theorem 2.14 (p. 92), we have

$$\begin{aligned}[\mathsf{T}^{-1}(a_0 + a_1x + a_2x^2)]_\beta &= \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} \\ &= \begin{pmatrix} a_0 - a_1 \\ a_1 - 2a_2 \\ a_2 \end{pmatrix}.\end{aligned}$$

Therefore

$$\mathsf{T}^{-1}(a_0 + a_1x + a_2x^2) = (a_0 - a_1) + (a_1 - 2a_2)x + a_2x^2. \quad ♦$$

EXERCISES

1. Label the following statements as true or false.
   (a) The rank of a matrix is equal to the number of its nonzero columns.
   (b) The product of two matrices always has rank equal to the lesser of the ranks of the two matrices.

(c) The $m \times n$ zero matrix is the only $m \times n$ matrix having rank 0.
(d) Elementary row operations preserve rank.
(e) Elementary column operations do not necessarily preserve rank.
(f) The rank of a matrix is equal to the maximum number of linearly independent rows in the matrix.
(g) The inverse of a matrix can be computed exclusively by means of elementary row operations.
(h) The rank of an $n \times n$ matrix is at most n.
(i) An $n \times n$ matrix having rank n is invertible.

2. Find the rank of the following matrices.

(a) $\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix}$ (b) $\begin{pmatrix} 1 & 1 & 0 \\ 2 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$ (c) $\begin{pmatrix} 1 & 0 & 2 \\ 1 & 1 & 4 \end{pmatrix}$

(d) $\begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \end{pmatrix}$ (e) $\begin{pmatrix} 1 & 2 & 3 & 1 & 1 \\ 1 & 4 & 0 & 1 & 2 \\ 0 & 2 & -3 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}$

(f) $\begin{pmatrix} 1 & 2 & 0 & 1 & 1 \\ 2 & 4 & 1 & 3 & 0 \\ 3 & 6 & 2 & 5 & 1 \\ -4 & -8 & 1 & -3 & 1 \end{pmatrix}$ (g) $\begin{pmatrix} 1 & 1 & 0 & 1 \\ 2 & 2 & 0 & 2 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 \end{pmatrix}$

3. Prove that for any $m \times n$ matrix A, $\operatorname{rank}(A) = 0$ if and only if A is the zero matrix.

4. Use elementary row and column operations to transform each of the following matrices into a matrix D satisfying the conditions of Theorem 3.6, and then determine the rank of each matrix.

(a) $\begin{pmatrix} 1 & 1 & 1 & 2 \\ 2 & 0 & -1 & 2 \\ 1 & 1 & 1 & 2 \end{pmatrix}$ (b) $\begin{pmatrix} 2 & 1 \\ -1 & 2 \\ 2 & 1 \end{pmatrix}$

5. For each of the following matrices, compute the rank and the inverse if it exists.

(a) $\begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}$ (b) $\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$ (c) $\begin{pmatrix} 1 & 2 & 1 \\ 1 & 3 & 4 \\ 2 & 3 & -1 \end{pmatrix}$

(d) $\begin{pmatrix} 0 & -2 & 4 \\ 1 & 1 & -1 \\ 2 & 4 & -5 \end{pmatrix}$ (e) $\begin{pmatrix} 1 & 2 & 1 \\ -1 & 1 & 2 \\ 1 & 0 & 1 \end{pmatrix}$ (f) $\begin{pmatrix} 1 & 2 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}$

(g) $\begin{pmatrix} 1 & 2 & 1 & 0 \\ 2 & 5 & 5 & 1 \\ -2 & -3 & 0 & 3 \\ 3 & 4 & -2 & -3 \end{pmatrix}$ (h) $\begin{pmatrix} 1 & 0 & 1 & 1 \\ 1 & 1 & -1 & 2 \\ 2 & 0 & 1 & 0 \\ 0 & -1 & 1 & -3 \end{pmatrix}$

6. For each of the following linear transformations T, determine whether T is invertible, and compute T^{-1} if it exists.
   (a) $\mathsf{T}\colon \mathsf{P}_2(R) \to \mathsf{P}_2(R)$ defined by $\mathsf{T}(f(x)) = f''(x) + 2f'(x) - f(x)$.
   (b) $\mathsf{T}\colon \mathsf{P}_2(R) \to \mathsf{P}_2(R)$ defined by $\mathsf{T}(f(x)) = (x+1)f'(x)$.
   (c) $\mathsf{T}\colon \mathsf{R}^3 \to \mathsf{R}^3$ defined by
   $$\mathsf{T}(a_1, a_2, a_3) = (a_1 + 2a_2 + a_3, -a_1 + a_2 + 2a_3, a_1 + a_3).$$
   (d) $\mathsf{T}\colon \mathsf{R}^3 \to \mathsf{P}_2(R)$ defined by
   $$\mathsf{T}(a_1, a_2, a_3) = (a_1 + a_2 + a_3) + (a_1 - a_2 + a_3)x + a_1x^2.$$
   (e) $\mathsf{T}\colon \mathsf{P}_2(R) \to \mathsf{R}^3$ defined by $\mathsf{T}(f(x)) = (f(-1), f(0), f(1))$.
   (f) $\mathsf{T}\colon \mathsf{M}_{2\times 2}(R) \to \mathsf{R}^4$ defined by
   $$\mathsf{T}(A) = (\mathrm{tr}(A), \mathrm{tr}(A^t), \mathrm{tr}(EA), \mathrm{tr}(AE)),$$
   where
   $$E = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

7. Express the invertible matrix
$$\begin{pmatrix} 1 & 2 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 2 \end{pmatrix}$$
as a product of elementary matrices.

8. Let A be an $m \times n$ matrix. Prove that if c is any nonzero scalar, then $\mathrm{rank}(cA) = \mathrm{rank}(A)$.

9. Complete the proof of the corollary to Theorem 3.4 by showing that elementary column operations preserve rank. Visit goo.gl/7KgM6F for a solution.

10. Prove Theorem 3.6 for the case that A is an $m \times 1$ matrix.

11. Let
$$B = \left(\begin{array}{c|c} 1 & 0 \ \cdots \ 0 \\ \hline 0 & \\ \vdots & B' \\ 0 & \end{array}\right),$$
where B' is an $m \times n$ submatrix of B. Prove that if $\mathrm{rank}(B) = r$, then $\mathrm{rank}(B') = r - 1$.

12. Let B' and D' be $m \times n$ matrices, and let B and D be $(m+1) \times (n+1)$ matrices respectively defined by

$$B = \left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & B' & \\ 0 & & & \end{array}\right) \quad \text{and} \quad D = \left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & D' & \\ 0 & & & \end{array}\right).$$

Prove that if B' can be transformed into D' by an elementary row [column] operation, then B can be transformed into D by an elementary row [column] operation.

13. Prove (b) and (c) of Corollary 2 to Theorem 3.6.

14. Let $\mathsf{T}, \mathsf{U}\colon \mathsf{V} \to \mathsf{W}$ be linear transformations.

(a) Prove that $\mathsf{R}(\mathsf{T}+\mathsf{U}) \subseteq \mathsf{R}(\mathsf{T})+\mathsf{R}(\mathsf{U})$. (See the definition of the sum of subsets of a vector space on page 22.)
(b) Prove that if W is finite-dimensional, then $\operatorname{rank}(\mathsf{T}+\mathsf{U}) \leq \operatorname{rank}(\mathsf{T})+\operatorname{rank}(\mathsf{U})$.
(c) Deduce from (b) that $\operatorname{rank}(A+B) \leq \operatorname{rank}(A)+\operatorname{rank}(B)$ for any $m \times n$ matrices A and B.

15. Suppose that A and B are matrices having n rows. Prove that $M(A|B) = (MA|MB)$ for any $m \times n$ matrix M.

16. Supply the details to the proof of (b) of Theorem 3.4.

17. Prove that if B is a 3×1 matrix and C is a 1×3 matrix, then the 3×3 matrix BC has rank at most 1. Conversely, show that if A is any 3×3 matrix having rank 1, then there exist a 3×1 matrix B and a 1×3 matrix C such that $A = BC$.

18. Let A be an $m \times n$ matrix and B be an $n \times p$ matrix. Prove that AB can be written as a sum of n matrices of rank at most one.

19. Let A be an $m \times n$ matrix with rank m and B be an $n \times p$ matrix with rank n. Determine the rank of AB. Justify your answer.

20. Let

$$A = \begin{pmatrix} 1 & 0 & -1 & 2 & 1 \\ -1 & 1 & 3 & -1 & 0 \\ -2 & 1 & 4 & -1 & 3 \\ 3 & -1 & -5 & 1 & -6 \end{pmatrix}.$$

(a) Find a 5×5 matrix M with rank 2 such that $AM = O$, where O is the 4×5 zero matrix.

(b) Suppose that B is a 5×5 matrix such that $AB = O$. Prove that $\operatorname{rank}(B) \leq 2$.

21. Let A be an $m \times n$ matrix with rank m. Prove that there exists an $n \times m$ matrix B such that $AB = I_m$.

22. Let B be an $n \times m$ matrix with rank m. Prove that there exists an $m \times n$ matrix A such that $AB = I_m$.

3.3 SYSTEMS OF LINEAR EQUATIONS—THEORETICAL ASPECTS

This section and the next are devoted to the study of systems of linear equations, which arise naturally in both the physical and social sciences. In this section, we apply results from Chapter 2 to describe the solution sets of systems of linear equations as subsets of a vector space. In Section 3.4, we will use elementary row operations to provide a computational method for finding all solutions to such systems.

The system of equations

$$(S) \qquad \begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m, \end{aligned}$$

where a_{ij} and b_i ($1 \leq i \leq m$ and $1 \leq j \leq n$) are scalars in a field F and $x_1, x_2, \ldots, x_n$ are n variables taking values in F, is called a **system of m linear equations in n unknowns over the field F.**

The $m \times n$ matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

is called the **coefficient matrix** of the system (S).

If we let

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \quad \text{and} \quad b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix},$$

then the system (S) may be rewritten as a single matrix equation

$$Ax = b.$$

To exploit the results that we have developed, we often consider a system of linear equations as a single matrix equation.

Recall that for a real number a, the series $1 + a + a^2 + \cdots$ converges to $(1-a)^{-1}$ if $|a| < 1$. Similarly, it can be shown (using the concept of convergence of matrices developed in Section 5.3) that the series $I + A + A^2 + \cdots$ converges to $(I-A)^{-1}$ if $\{A^n\}$ converges to the zero matrix. In this case, $(I-A)^{-1}$ is nonnegative since the matrices $I, A, A^2, \ldots$ are nonnegative.

To illustrate the open model, suppose that 30 cents worth of food, 10 cents worth of clothing, and 30 cents worth of housing are required for the production of \$1 worth of food. Similarly, suppose that 20 cents worth of food, 40 cents worth of clothing, and 20 cents worth of housing are required for the production of \$1 of clothing. Finally, suppose that 30 cents worth of food, 10 cents worth of clothing, and 30 cents worth of housing are required for the production of \$1 worth of housing. Then the input–output matrix is

$$A = \begin{pmatrix} 0.30 & 0.20 & 0.30 \\ 0.10 & 0.40 & 0.10 \\ 0.30 & 0.20 & 0.30 \end{pmatrix};$$

so

$$I - A = \begin{pmatrix} 0.70 & -0.20 & -0.30 \\ -0.10 & 0.60 & -0.10 \\ -0.30 & -0.20 & 0.70 \end{pmatrix} \quad \text{and} \quad (I-A)^{-1} = \begin{pmatrix} 2.0 & 1.0 & 1.0 \\ 0.5 & 2.0 & 0.5 \\ 1.0 & 1.0 & 2.0 \end{pmatrix}.$$

Since $(I-A)^{-1}$ is nonnegative, we can find a (unique) nonnegative solution to $(I-A)x = d$ for any demand d. For example, suppose that there are outside demands for \$30 billion in food, \$20 billion in clothing, and \$10 billion in housing. If we set

$$d = \begin{pmatrix} 30 \\ 20 \\ 10 \end{pmatrix},$$

then

$$x = (I-A)^{-1}d = \begin{pmatrix} 90 \\ 60 \\ 70 \end{pmatrix}.$$

So a gross production of \$90 billion of food, \$60 billion of clothing, and \$70 billion of housing is necessary to meet the required demands.

EXERCISES

1. Label the following statements as true or false.
   (a) Any system of linear equations has at least one solution.
   (b) Any system of linear equations has at most one solution.
   (c) Any homogeneous system of linear equations has at least one solution.

(d) Any system of n linear equations in n unknowns has at most one solution.
(e) Any system of n linear equations in n unknowns has at least one solution.
(f) If the homogeneous system corresponding to a given system of linear equations has a solution, then the given system has a solution.
(g) If the coefficient matrix of a homogeneous system of n linear equations in n unknowns is invertible, then the system has no nonzero solutions.
(h) The solution set of any system of m linear equations in n unknowns is a subspace of F^n.

2. For each of the following homogeneous systems of linear equations, find the dimension of and a basis for the solution space.

(a) $\begin{aligned} x_1 + 3x_2 &= 0 \\ 2x_1 + 6x_2 &= 0 \end{aligned}$ (b) $\begin{aligned} x_1 + x_2 - x_3 &= 0 \\ 4x_1 + x_2 - 2x_3 &= 0 \end{aligned}$

(c) $\begin{aligned} x_1 + 2x_2 - x_3 &= 0 \\ 2x_1 + x_2 + x_3 &= 0 \end{aligned}$ (d) $\begin{aligned} 2x_1 + x_2 - x_3 &= 0 \\ x_1 - x_2 + x_3 &= 0 \\ x_1 + 2x_2 - 2x_3 &= 0 \end{aligned}$

(e) $x_1 + 2x_2 - 3x_3 + x_4 = 0$ (f) $\begin{aligned} x_1 + 2x_2 &= 0 \\ x_1 - x_2 &= 0 \end{aligned}$

(g) $\begin{aligned} x_1 + 2x_2 + x_3 + x_4 &= 0 \\ x_2 - x_3 + x_4 &= 0 \end{aligned}$

3. Using the results of Exercise 2, find all solutions to the following systems.

(a) $\begin{aligned} x_1 + 3x_2 &= 5 \\ 2x_1 + 6x_2 &= 10 \end{aligned}$ (b) $\begin{aligned} x_1 + x_2 - x_3 &= 1 \\ 4x_1 + x_2 - 2x_3 &= 3 \end{aligned}$

(c) $\begin{aligned} x_1 + 2x_2 - x_3 &= 3 \\ 2x_1 + x_2 + x_3 &= 6 \end{aligned}$ (d) $\begin{aligned} 2x_1 + x_2 - x_3 &= 5 \\ x_1 - x_2 + x_3 &= 1 \\ x_1 + 2x_2 - 2x_3 &= 4 \end{aligned}$

(e) $x_1 + 2x_2 - 3x_3 + x_4 = 1$ (f) $\begin{aligned} x_1 + 2x_2 &= 5 \\ x_1 - x_2 &= -1 \end{aligned}$

(g) $\begin{aligned} x_1 + 2x_2 + x_3 + x_4 &= 1 \\ x_2 - x_3 + x_4 &= 1 \end{aligned}$

4. For each system of linear equations with the invertible coefficient matrix A,

(1) Compute A^{-1}.

(2) Use A^{-1} to solve the system.

(a) $\begin{aligned} x_1 + 3x_2 &= 4 \\ 2x_1 + 5x_2 &= 3 \end{aligned}$ (b) $\begin{aligned} x_1 + 2x_2 - x_3 &= 5 \\ x_1 + x_2 + x_3 &= 1 \\ 2x_1 - 2x_2 + x_3 &= 4 \end{aligned}$

5. Give an example of a system of n linear equations in n unknowns with infinitely many solutions.

6. Let $\mathsf{T}\colon \mathsf{R}^3 \to \mathsf{R}^2$ be defined by $\mathsf{T}(a,b,c) = (a+b, 2a-c)$. Determine $\mathsf{T}^{-1}(1,11)$.

7. Determine which of the following systems of linear equations has a solution.

(a) $\begin{aligned} x_1 + x_2 - x_3 + 2x_4 &= 2 \\ x_1 + x_2 + 2x_3 &= 1 \\ 2x_1 + 2x_2 + x_3 + 2x_4 &= 4 \end{aligned}$ (b) $\begin{aligned} x_1 + x_2 - x_3 &= 1 \\ 2x_1 + x_2 + 3x_3 &= 2 \end{aligned}$

(c) $\begin{aligned} x_1 + 2x_2 + 3x_3 &= 1 \\ x_1 + x_2 - x_3 &= 0 \\ x_1 + 2x_2 + x_3 &= 3 \end{aligned}$ (d) $\begin{aligned} x_1 + x_2 + 3x_3 - x_4 &= 0 \\ x_1 + x_2 + x_3 + x_4 &= 1 \\ x_1 - 2x_2 + x_3 - x_4 &= 1 \\ 4x_1 + x_2 + 8x_3 - x_4 &= 0 \end{aligned}$

(e) $\begin{aligned} x_1 + 2x_2 - x_3 &= 1 \\ 2x_1 + x_2 + 2x_3 &= 3 \\ x_1 - 4x_2 + 7x_3 &= 4 \end{aligned}$

8. Let $\mathsf{T}\colon \mathsf{R}^3 \to \mathsf{R}^3$ be defined by $\mathsf{T}(a,b,c) = (a+b, b-2c, a+2c)$. For each vector v in R^3, determine whether $v \in \mathsf{R}(\mathsf{T})$.

(a) $v = (1,3,-2)$ (b) $v = (2,1,1)$

$\boxed{9.}$ Prove that the system of linear equations $Ax = b$ has a solution if and only if $b \in \mathsf{R}(\mathsf{L}_A)$. Visit `goo.gl/JfwjBa` for a solution.

10. Prove or give a counterexample to the following statement: If the coefficient matrix of a system of m linear equations in n unknowns has rank m, then the system has a solution.

11. In the closed model of Leontief with food, clothing, and housing as the basic industries, suppose that the input–output matrix is

$$A = \begin{pmatrix} \frac{7}{16} & \frac{1}{2} & \frac{3}{16} \\ \frac{5}{16} & \frac{1}{6} & \frac{5}{16} \\ \frac{1}{4} & \frac{1}{3} & \frac{1}{2} \end{pmatrix}.$$

At what ratio must the farmer, tailor, and carpenter produce in order for equilibrium to be attained?

12. A certain economy consists of two sectors: goods and services. Suppose that 60% of all goods and 30% of all services are used in the production of goods. What proportion of the total economic output is used in the production of goods?

13. In the notation of the open model of Leontief, suppose that

$$A = \begin{pmatrix} \frac{1}{2} & \frac{1}{5} \\ \frac{1}{3} & \frac{1}{5} \end{pmatrix} \quad \text{and} \quad d = \begin{pmatrix} 2 \\ 5 \end{pmatrix}$$

are the input–output matrix and the demand vector, respectively. How much of each commodity must be produced to satisfy this demand?

14. A certain economy consisting of the two sectors of goods and services supports a defense system that consumes \$90 billion worth of goods and \$20 billion worth of services from the economy but does not contribute to economic production. Suppose that 50 cents worth of goods and 20 cents worth of services are required to produce \$1 worth of goods and that 30 cents worth of goods and 60 cents worth of services are required to produce \$1 worth of services. What must the total output of the economic system be to support this defense system?

3.4 SYSTEMS OF LINEAR EQUATIONS—COMPUTATIONAL ASPECTS

In Section 3.3, we obtained a necessary and sufficient condition for a system of linear equations to have solutions (Theorem 3.11 p. 174) and learned how to express the solutions to a nonhomogeneous system in terms of solutions to the corresponding homogeneous system (Theorem 3.9 p. 172). The latter result enables us to determine all the solutions to a given system if we can find one solution to the given system and a basis for the solution set of the corresponding homogeneous system. In this section, we use elementary row operations to accomplish these two objectives simultaneously. The essence of this technique is to transform a given system of linear equations into a system having the same solutions, but which is easier to solve (as in Section 1.4).

Definition. *Two systems of linear equations are called* **equivalent** *if they have the same solution set.*

The following theorem and corollary give a useful method for obtaining equivalent systems.

EXERCISES

1. Label the following statements as true or false.
   (a) If $(A'|b')$ is obtained from $(A|b)$ by a finite sequence of elementary column operations, then the systems $Ax = b$ and $A'x = b'$ are equivalent.
   (b) If $(A'|b')$ is obtained from $(A|b)$ by a finite sequence of elementary row operations, then the systems $Ax = b$ and $A'x = b'$ are equivalent.
   (c) If A is an $n \times n$ matrix with rank n, then the reduced row echelon form of A is I_n.
   (d) Any matrix can be put in reduced row echelon form by means of a finite sequence of elementary row operations.
   (e) If $(A|b)$ is in reduced row echelon form, then the system $Ax = b$ is consistent.
   (f) Let $Ax = b$ be a system of m linear equations in n unknowns for which the augmented matrix is in reduced row echelon form. If this system is consistent, then the dimension of the solution set of $Ax = 0$ is $n - r$, where r equals the number of nonzero rows in A.
   (g) If a matrix A is transformed by elementary row operations into a matrix A' in reduced row echelon form, then the number of nonzero rows in A' equals the rank of A.

2. Use Gaussian elimination to solve the following systems of linear equations.

(a)
$$\begin{aligned} x_1 + 2x_2 - x_3 &= -1 \\ 2x_1 + 2x_2 + x_3 &= 1 \\ 3x_1 + 5x_2 - 2x_3 &= -1 \end{aligned}$$

(b)
$$\begin{aligned} x_1 - 2x_2 - x_3 &= 1 \\ 2x_1 - 3x_2 + x_3 &= 6 \\ 3x_1 - 5x_2 &= 7 \\ x_1 + 5x_3 &= 9 \end{aligned}$$

(c)
$$\begin{aligned} x_1 + 2x_2 + 2x_4 &= 6 \\ 3x_1 + 5x_2 - x_3 + 6x_4 &= 17 \\ 2x_1 + 4x_2 + x_3 + 2x_4 &= 12 \\ 2x_1 - 7x_3 + 11x_4 &= 7 \end{aligned}$$

(d)
$$\begin{aligned} x_1 - x_2 - 2x_3 + 3x_4 &= -63 \\ 2x_1 - x_2 + 6x_3 + 6x_4 &= -2 \\ -2x_1 + x_2 - 4x_3 - 3x_4 &= 0 \\ 3x_1 - 2x_2 + 9x_3 + 10x_4 &= -5 \end{aligned}$$

(e)
$$\begin{aligned} x_1 - 4x_2 - x_3 + x_4 &= 3 \\ 2x_1 - 8x_2 + x_3 - 4x_4 &= 9 \\ -x_1 + 4x_2 - 2x_3 + 5x_4 &= -6 \end{aligned}$$

(f)
$$\begin{aligned} x_1 + 2x_2 - x_3 + 3x_4 &= 2 \\ 2x_1 + 4x_2 - x_3 + 6x_4 &= 5 \\ x_2 + 2x_4 &= 3 \end{aligned}$$

(g)
$$\begin{aligned} 2x_1 - 2x_2 - x_3 + 6x_4 - 2x_5 &= 1 \\ x_1 - x_2 + x_3 + 2x_4 - x_5 &= 2 \\ 4x_1 - 4x_2 + 5x_3 + 7x_4 - x_5 &= 6 \end{aligned}$$

(h)
$$\begin{aligned} 3x_1 - x_2 + x_3 - x_4 + 2x_5 &= 5 \\ x_1 - x_2 - x_3 - 2x_4 - x_5 &= 2 \\ 5x_1 - 2x_2 + x_3 - 3x_4 + 3x_5 &= 10 \\ 2x_1 - x_2 \qquad - 2x_4 + x_5 &= 5 \end{aligned}$$

(i)
$$\begin{aligned} 3x_1 - x_2 + 2x_3 + 4x_4 + x_5 &= 2 \\ x_1 - x_2 + 2x_3 + 3x_4 + x_5 &= -1 \\ 2x_1 - 3x_2 + 6x_3 + 9x_4 + 4x_5 &= -5 \\ 7x_1 - 2x_2 + 4x_3 + 8x_4 + x_5 &= 6 \end{aligned}$$

(j)
$$\begin{aligned} 2x_1 \qquad + 3x_3 \qquad - 4x_5 &= 5 \\ 3x_1 - 4x_2 + 8x_3 + 3x_4 \qquad &= 8 \\ x_1 - x_2 + 2x_3 + x_4 - x_5 &= 2 \\ -2x_1 + 5x_2 - 9x_3 - 3x_4 - 5x_5 &= -8 \end{aligned}$$

3. Suppose that the augmented matrix of a system $Ax = b$ is transformed into a matrix $(A'|b')$ in reduced row echelon form by a finite sequence of elementary row operations.

(a) Prove that $\text{rank}(A') \neq \text{rank}(A'|b')$ if and only if $(A'|b')$ contains a row in which the only nonzero entry lies in the last column.

(b) Deduce that $Ax = b$ is consistent if and only if $(A'|b')$ contains no row in which the only nonzero entry lies in the last column.

4. For each of the systems that follow, apply Exercise 3 to determine whether the system is consistent. If the system is consistent, find all solutions. Finally, find a basis for the solution set of the corresponding homogeneous system.

(a)
$$\begin{aligned} x_1 + 2x_2 - x_3 + x_4 &= 2 \\ 2x_1 + x_2 + x_3 - x_4 &= 3 \\ x_1 + 2x_2 - 3x_3 + 2x_4 &= 2 \end{aligned}$$

(b)
$$\begin{aligned} x_1 + x_2 - 3x_3 + x_4 &= -2 \\ x_1 + x_2 + x_3 - x_4 &= 2 \\ x_1 + x_2 - x_3 \qquad &= 0 \end{aligned}$$

(c)
$$\begin{aligned} x_1 + x_2 - 3x_3 + x_4 &= 1 \\ x_1 + x_2 + x_3 - x_4 &= 2 \\ x_1 + x_2 - x_3 \qquad &= 0 \end{aligned}$$

5. Let the reduced row echelon form of A be

$$\begin{pmatrix} 1 & 0 & 2 & 0 & -2 \\ 0 & 1 & -5 & 0 & -3 \\ 0 & 0 & 0 & 1 & 6 \end{pmatrix}.$$

Determine A if the first, second, and fourth columns of A are

$$\begin{pmatrix} 1 \\ -1 \\ 3 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}, \quad \text{and} \quad \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix},$$

respectively.

6. Let the reduced row echelon form of A be

$$\begin{pmatrix} 1 & -3 & 0 & 4 & 0 & 5 \\ 0 & 0 & 1 & 3 & 0 & 2 \\ 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Determine A if the first, third, and sixth columns of A are

$$\begin{pmatrix} 1 \\ -2 \\ -1 \\ 3 \end{pmatrix}, \quad \begin{pmatrix} -1 \\ 1 \\ 2 \\ -4 \end{pmatrix}, \quad \text{and} \quad \begin{pmatrix} 3 \\ -9 \\ 2 \\ 5 \end{pmatrix},$$

respectively.

7. It can be shown that the vectors $u_1 = (2, -3, 1)$, $u_2 = (1, 4, -2)$, $u_3 = (-8, 12, -4)$, $u_4 = (1, 37, -17)$, and $u_5 = (-3, -5, 8)$ generate R^3. Find a subset of $\{u_1, u_2, u_3, u_4, u_5\}$ that is a basis for R^3.

8. Let W denote the subspace of R^5 consisting of all vectors having coordinates that sum to zero. The vectors

$$\begin{aligned} u_1 &= (2, -3, 4, -5, 2), & u_2 &= (-6, 9, -12, 15, -6), \\ u_3 &= (3, -2, 7, -9, 1), & u_4 &= (2, -8, 2, -2, 6), \\ u_5 &= (-1, 1, 2, 1, -3), & u_6 &= (0, -3, -18, 9, 12), \\ u_7 &= (1, 0, -2, 3, -2), \quad \text{and} & u_8 &= (2, -1, 1, -9, 7) \end{aligned}$$

generate W. Find a subset of $\{u_1, u_2, \ldots, u_8\}$ that is a basis for W.

9. Let W be the subspace of $\mathsf{M}_{2\times 2}(R)$ consisting of the symmetric 2×2 matrices. The set

$$S = \left\{ \begin{pmatrix} 0 & -1 \\ -1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}, \begin{pmatrix} 2 & 1 \\ 1 & 9 \end{pmatrix}, \begin{pmatrix} 1 & -2 \\ -2 & 4 \end{pmatrix}, \begin{pmatrix} -1 & 2 \\ 2 & -1 \end{pmatrix} \right\}$$

generates W. Find a subset of S that is a basis for W.

10. Let

$$\mathsf{V} = \{(x_1, x_2, x_3, x_4, x_5) \in \mathsf{R}^5 \colon x_1 - 2x_2 + 3x_3 - x_4 + 2x_5 = 0\}.$$

(a) Show that $S = \{(0, 1, 1, 1, 0)\}$ is a linearly independent subset of V.

(b) Extend S to a basis for V.

11. Let V be as in Exercise 10.

(a) Show that $S = \{(1,2,1,0,0)\}$ is a linearly independent subset of V.
(b) Extend S to a basis for V.

12. Let V denote the set of all solutions to the system of linear equations

$$\begin{aligned} x_1 - x_2 + 2x_4 - 3x_5 + x_6 &= 0 \\ 2x_1 - x_2 - x_3 + 3x_4 - 4x_5 + 4x_6 &= 0. \end{aligned}$$

(a) Show that $S = \{(0,-1,0,1,1,0),(1,0,1,1,1,0)\}$ is a linearly independent subset of V.
(b) Extend S to a basis for V.

13. Let V be as in Exercise 12.

(a) Show that $S = \{(1,0,1,1,1,0),(0,2,1,1,0,0)\}$ is a linearly independent subset of V.
(b) Extend S to a basis for V.

14. If $(A|b)$ is in reduced row echelon form, prove that A is also in reduced row echelon form.

15. Prove the corollary to Theorem 3.16: The reduced row echelon form of a matrix is unique. Visit `goo.gl/cZVzxM` for a solution.

INDEX OF DEFINITIONS FOR CHAPTER 3

for any $u \in \mathsf{R}^2$.

(c) Because the parallelogram determined by e_1 and e_2 is the unit square,

$$\delta\begin{pmatrix} e_1 \\ e_2 \end{pmatrix} = \mathcal{O}\begin{pmatrix} e_1 \\ e_2 \end{pmatrix} \cdot \mathcal{A}\begin{pmatrix} e_1 \\ e_2 \end{pmatrix} = 1 \cdot 1 = 1.$$

Therefore δ satisfies the three conditions of Exercise 11, and hence $\delta = \det$. So the area of the parallelogram determined by u and v equals

$$\mathcal{O}\begin{pmatrix} u \\ v \end{pmatrix} \cdot \det\begin{pmatrix} u \\ v \end{pmatrix}.$$

Thus we see, for example, that the area of the parallelogram determined by $u = (-1, 5)$ and $v = (4, -2)$ is

$$\left|\det\begin{pmatrix} u \\ v \end{pmatrix}\right| = \left|\det\begin{pmatrix} -1 & 5 \\ 4 & -2 \end{pmatrix}\right| = 18.$$

EXERCISES

1. Label the following statements as true or false.
   (a) The function $\det\colon \mathsf{M}_{2\times 2}(F) \to F$ is a linear transformation.
   (b) The determinant of a 2×2 matrix is a linear function of each row of the matrix when the other row is held fixed.
   (c) If $A \in \mathsf{M}_{2\times 2}(F)$ and $\det(A) = 0$, then A is invertible.
   (d) If u and v are vectors in R^2 emanating from the origin, then the area of the parallelogram having u and v as adjacent sides is
   $$\det\begin{pmatrix} u \\ v \end{pmatrix}.$$
   (e) A coordinate system is right-handed if and only if its orientation equals 1.

2. Compute the determinants of the following matrices in $\mathsf{M}_{2\times 2}(R)$.
   (a) $\begin{pmatrix} 6 & -3 \\ 2 & 4 \end{pmatrix}$ (b) $\begin{pmatrix} -5 & 2 \\ 6 & 1 \end{pmatrix}$ (c) $\begin{pmatrix} 8 & 0 \\ 3 & -1 \end{pmatrix}$

3. Compute the determinants of the following matrices in $\mathsf{M}_{2\times 2}(C)$.
   (a) $\begin{pmatrix} -1+i & 1-4i \\ 3+2i & 2-3i \end{pmatrix}$ (b) $\begin{pmatrix} 5-2i & 6+4i \\ -3+i & 7i \end{pmatrix}$ (c) $\begin{pmatrix} 2i & 3 \\ 4 & 6i \end{pmatrix}$

4. For each of the following pairs of vectors u and v in R^2, compute the area of the parallelogram determined by u and v.
   (a) $u = (3, -2)$ and $v = (2, 5)$

(b) $u = (1, 3)$ and $v = (-3, 1)$
(c) $u = (4, -1)$ and $v = (-6, -2)$
(d) $u = (3, 4)$ and $v = (2, -6)$

5. Prove that if B is the matrix obtained by interchanging the rows of a 2×2 matrix A, then $\det(B) = -\det(A)$.

6. Prove that if the two columns of $A \in \mathsf{M}_{2\times 2}(F)$ are identical, then $\det(A) = 0$.

7. Prove that $\det(A^t) = \det(A)$ for any $A \in \mathsf{M}_{2\times 2}(F)$.

8. Prove that if $A \in \mathsf{M}_{2\times 2}(F)$ is upper triangular, then $\det(A)$ equals the product of the diagonal entries of A.

9. Prove that $\det(AB) = \det(A) \cdot \det(B)$ for any $A, B \in \mathsf{M}_{2\times 2}(F)$.

10. The **classical adjoint** of a 2×2 matrix $A \in \mathsf{M}_{2\times 2}(F)$ is the matrix

$$C = \begin{pmatrix} A_{22} & -A_{12} \\ -A_{21} & A_{11} \end{pmatrix}.$$

Prove that

(a) $CA = AC = [\det(A)]I$.
(b) $\det(C) = \det(A)$.
(c) The classical adjoint of A^t is C^t.
(d) If A is invertible, then $A^{-1} = [\det(A)]^{-1}C$.

11. Let $\delta\colon \mathsf{M}_{2\times 2}(F) \to F$ be a function with the following three properties.

(i) δ is a linear function of each row of the matrix when the other row is held fixed.

(ii) If the two rows of $A \in \mathsf{M}_{2\times 2}(F)$ are identical, then $\delta(A) = 0$.

(iii) If I is the 2×2 identity matrix, then $\delta(I) = 1$.

(a) Prove that $\delta(E) = \det(E)$ for all elementary matrices $E \in \mathsf{M}_{2\times 2}(F)$.
(b) Prove that $\delta(EA) = \delta(E)\delta(A)$ for all $A \in \mathsf{M}_{2\times 2}(F)$ and all elementary matrices $E \in \mathsf{M}_{2\times 2}(F)$.

$\boxed{\textbf{12.}}$ Let $\delta\colon \mathsf{M}_{2\times 2}(F) \to F$ be a function with properties (i), (ii), and (iii) in Exercise 11. Use Exercise 11 to prove that $\delta(A) = \det(A)$ for all $A \in \mathsf{M}_{2\times 2}(F)$. (This result is generalized in Section 4.5.) Visit `goo.gl/ztwxWA` for a solution.

13. Let $\{u, v\}$ be an ordered basis for R^2. Prove that

$$\mathcal{O}\begin{pmatrix} u \\ v \end{pmatrix} = 1$$

if and only if $\{u, v\}$ forms a right-handed coordinate system. *Hint:* Recall the definition of a rotation given in Example 2 of Section 2.1.

4.2 DETERMINANTS OF ORDER n

In this section, we extend the definition of the determinant to $n \times n$ matrices for $n \geq 3$. For this definition, it is convenient to introduce the following notation: Given $A \in \mathsf{M}_{n\times n}(F)$, for $n \geq 2$, denote the $(n-1) \times (n-1)$ matrix obtained from A by deleting row i and column j by $\tilde{A}_{ij}$. Thus for

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \in \mathsf{M}_{3\times 3}(R),$$

we have

$$\tilde{A}_{11} = \begin{pmatrix} 5 & 6 \\ 8 & 9 \end{pmatrix}, \qquad \tilde{A}_{13} = \begin{pmatrix} 4 & 5 \\ 7 & 8 \end{pmatrix}, \quad \text{and} \quad \tilde{A}_{32} = \begin{pmatrix} 1 & 3 \\ 4 & 6 \end{pmatrix},$$

and for

$$B = \begin{pmatrix} 1 & -1 & 2 & -1 \\ -3 & 4 & 1 & -1 \\ 2 & -5 & -3 & 8 \\ -2 & 6 & -4 & 1 \end{pmatrix} \in \mathsf{M}_{4\times 4}(R),$$

we have

$$\tilde{B}_{23} = \begin{pmatrix} 1 & -1 & -1 \\ 2 & -5 & 8 \\ -2 & 6 & 1 \end{pmatrix} \quad \text{and} \quad \tilde{B}_{42} = \begin{pmatrix} 1 & 2 & -1 \\ -3 & 1 & -1 \\ 2 & -3 & 8 \end{pmatrix}.$$

Definitions. *Let $A \in \mathsf{M}_{n\times n}(F)$. If $n = 1$, so that $A = (A_{11})$, we define $\det(A) = A_{11}$. For $n \geq 2$, we define $\det(A)$ recursively as*

$$\det(A) = \sum_{j=1}^{n} (-1)^{1+j} A_{1j} \cdot \det(\tilde{A}_{1j}).$$

The scalar $\det(A)$ is called the **determinant** *of A and is also denoted by $|A|$. The scalar*

$$(-1)^{i+j} \det(\tilde{A}_{ij})$$

is called the **cofactor** *of the entry of A in row i, column j.*

the evaluation of the determinant of a 2×2 matrix requires 2 multiplications (and 1 subtraction). For $n \geq 3$, evaluating the determinant of an $n \times n$ matrix by cofactor expansion along any row expresses the determinant as a sum of n products involving determinants of $(n-1) \times (n-1)$ matrices. Thus in all, the evaluation of the determinant of an $n \times n$ matrix by cofactor expansion along any row requires over $n!$ multiplications, whereas evaluating the determinant of an $n \times n$ matrix by elementary row operations as in Examples 5 and 6 can be shown to require only $(n^3 + 2n - 3)/3$ multiplications. To evaluate the determinant of a 20×20 matrix, which is not large by present standards, cofactor expansion along a row requires over $20! \approx 2.4 \times 10^{18}$ multiplications. Thus it would take a computer performing one billion multiplications per second over 77 years to evaluate the determinant of a 20×20 matrix by this method. By contrast, the method using elementary row operations requires only 2679 multiplications for this calculation and would take the same computer less than three-millionths of a second! It is easy to see why most computer programs for evaluating the determinant of an arbitrary matrix do not use cofactor expansion.

In this section, we have defined the determinant of a square matrix in terms of cofactor expansion along the first row. We then showed that the determinant of a square matrix can be evaluated using cofactor expansion along *any* row. In addition, we showed that the determinant possesses a number of special properties, including properties that enable us to calculate $\det(B)$ from $\det(A)$ whenever B is a matrix obtained from A by means of an elementary row operation. These properties enable us to evaluate determinants much more efficiently. In the next section, we continue this approach to discover additional properties of determinants.

EXERCISES

1. Label the following statements as true or false.
   (a) The function $\det \colon \mathsf{M}_{n\times n}(F) \to F$ is a linear transformation.
   (b) The determinant of a square matrix can be evaluated by cofactor expansion along any row.
   (c) If two rows of a square matrix A are identical, then $\det(A) = 0$.
   (d) If B is a matrix obtained from a square matrix A by interchanging any two rows, then $\det(B) = -\det(A)$.
   (e) If B is a matrix obtained from a square matrix A by multiplying a row of A by a scalar, then $\det(B) = \det(A)$.
   (f) If B is a matrix obtained from a square matrix A by adding k times row i to row j, then $\det(B) = k\det(A)$.
   (g) If $A \in \mathsf{M}_{n\times n}(F)$ has rank n, then $\det(A) = 0$.
   (h) The determinant of an upper triangular matrix equals the product of its diagonal entries.

2. Find the value of k that satisfies the following equation:

$$\det\begin{pmatrix} 3a_1 & 3a_2 & 3a_3 \\ 3b_1 & 3b_2 & 3b_3 \\ 3c_1 & 3c_2 & 3c_3 \end{pmatrix} = k\det\begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix}.$$

3. Find the value of k that satisfies the following equation:

$$\det\begin{pmatrix} 2a_1 & 2a_2 & 2a_3 \\ 3b_1+5c_1 & 3b_2+5c_2 & 3b_3+5c_3 \\ 7c_1 & 7c_2 & 7c_3 \end{pmatrix} = k\det\begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix}.$$

4. Find the value of k that satisfies the following equation:

$$\det\begin{pmatrix} b_1+c_1 & b_2+c_2 & b_3+c_3 \\ a_1+c_1 & a_2+c_2 & a_3+c_3 \\ a_1+b_1 & a_2+b_2 & a_3+b_3 \end{pmatrix} = k\det\begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix}.$$

In Exercises 5–12, evaluate the determinant of the given matrix by cofactor expansion along the indicated row.

5. $\begin{pmatrix} 0 & 1 & 2 \\ -1 & 0 & -3 \\ 2 & 3 & 0 \end{pmatrix}$
along the first row

6. $\begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 5 \\ -1 & 3 & 0 \end{pmatrix}$
along the first row

7. $\begin{pmatrix} 0 & 1 & 2 \\ -1 & 0 & -3 \\ 2 & 3 & 0 \end{pmatrix}$
along the second row

8. $\begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 5 \\ -1 & 3 & 0 \end{pmatrix}$
along the third row

9. $\begin{pmatrix} 0 & 1+i & 2 \\ -2i & 0 & 1-i \\ 3 & 4i & 0 \end{pmatrix}$
along the third row

10. $\begin{pmatrix} i & 2+i & 0 \\ -1 & 3 & 2i \\ 0 & -1 & 1-i \end{pmatrix}$
along the second row

11. $\begin{pmatrix} 0 & 2 & 1 & 3 \\ 1 & 0 & -2 & 2 \\ 3 & -1 & 0 & 1 \\ -1 & 1 & 2 & 0 \end{pmatrix}$
along the fourth row

12. $\begin{pmatrix} 1 & -1 & 2 & -1 \\ -3 & 4 & 1 & -1 \\ 2 & -5 & -3 & 8 \\ -2 & 6 & -4 & 1 \end{pmatrix}$
along the fourth row

n Exercises 13–22, evaluate the determinant of the given matrix by any le-
itimate method.

13. $\begin{pmatrix} 0 & 0 & 1 \\ 0 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$ **14.** $\begin{pmatrix} 2 & 3 & 4 \\ 5 & 6 & 0 \\ 7 & 0 & 0 \end{pmatrix}$

15. $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$ **16.** $\begin{pmatrix} -1 & 3 & 2 \\ 4 & -8 & 1 \\ 2 & 2 & 5 \end{pmatrix}$

17. $\begin{pmatrix} 0 & 1 & 1 \\ 1 & 2 & -5 \\ 6 & -4 & 3 \end{pmatrix}$ **18.** $\begin{pmatrix} 1 & -2 & 3 \\ -1 & 2 & -5 \\ 3 & -1 & 2 \end{pmatrix}$

19. $\begin{pmatrix} i & 2 & -1 \\ 3 & 1+i & 2 \\ -2i & 1 & 4-i \end{pmatrix}$ **20.** $\begin{pmatrix} -1 & 2+i & 3 \\ 1-i & i & 1 \\ 3i & 2 & -1+i \end{pmatrix}$

21. $\begin{pmatrix} 1 & 0 & -2 & 3 \\ -3 & 1 & 1 & 2 \\ 0 & 4 & -1 & 1 \\ 2 & 3 & 0 & 1 \end{pmatrix}$ **22.** $\begin{pmatrix} 1 & -2 & 3 & -12 \\ -5 & 12 & -14 & 19 \\ -9 & 22 & -20 & 31 \\ -4 & 9 & -14 & 15 \end{pmatrix}$

23. Prove that the determinant of an upper triangular matrix is the product of its diagonal entries.

24. Prove the corollary to Theorem 4.3.

25. Prove that $\det(kA) = k^n \det(A)$ for any $A \in \mathsf{M}_{n\times n}(F)$.

26. Let $A \in \mathsf{M}_{n\times n}(F)$. Under what conditions is $\det(-A) = \det(A)$?

27. Prove that if $A \in \mathsf{M}_{n\times n}(F)$ has two identical columns, then $\det(A) = 0$.

28. Compute $\det(E_i)$ if E_i is an elementary matrix of type i.

29.† Prove that if E is an elementary matrix, then $\det(E^t) = \det(E)$. Visit `goo.gl/6ZoU5Z` for a solution.

30. Let the rows of $A \in \mathsf{M}_{n\times n}(F)$ be $a_1, a_2, \ldots, a_n$, and let B be the matrix in which the rows are $a_n, a_{n-1}, \ldots, a_1$. Calculate $\det(B)$ in terms of $\det(A)$.

4.3 PROPERTIES OF DETERMINANTS

In Theorem 3.1, we saw that performing an elementary row operation on a matrix can be accomplished by multiplying the matrix by an elementary matrix. This result is very useful in studying the effects on the determinant of

More generally, if β and γ are two ordered bases for R^n, then the coordinate systems induced by β and γ have the same **orientation** (either both are right-handed or both are left-handed) if and only if $\det(Q) > 0$, where Q is the change of coordinate matrix changing γ-coordinates into β-coordinates.

EXERCISES

1. Label the following statements as true or false.
   (a) If E is an elementary matrix, then $\det(E) = \pm 1$.
   (b) For any $A, B \in \mathsf{M}_{n\times n}(F)$, $\det(AB) = \det(A)\cdot\det(B)$.
   (c) A matrix $M \in \mathsf{M}_{n\times n}(F)$ is invertible if and only if $\det(M) = 0$.
   (d) A matrix $M \in \mathsf{M}_{n\times n}(F)$ has rank n if and only if $\det(M) \neq 0$.
   (e) For any $A \in \mathsf{M}_{n\times n}(F)$, $\det(A^t) = -\det(A)$.
   (f) The determinant of a square matrix can be evaluated by cofactor expansion along any column.
   (g) Every system of n linear equations in n unknowns can be solved by Cramer's rule.
   (h) Let $Ax = b$ be the matrix form of a system of n linear equations in n unknowns, where $x = (x_1, x_2, \ldots, x_n)^t$. If $\det(A) \neq 0$ and if M_k is the $n \times n$ matrix obtained from A by replacing row k of A by b^t, then the unique solution of $Ax = b$ is

$$x_k = \frac{\det(M_k)}{\det(A)} \quad \text{for } k = 1, 2, \ldots, n.$$

In Exercises 2–7, use Cramer's rule to solve the given system of linear equations.

2. $$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1 \\ a_{21}x_1 + a_{22}x_2 &= b_2 \\ \text{where } a_{11}a_{22} - a_{12}a_{21} &\neq 0 \end{aligned}$$

3. $$\begin{aligned} 2x_1 + x_2 - 3x_3 &= 5 \\ x_1 - 2x_2 + x_3 &= 10 \\ 3x_1 + 4x_2 - 2x_3 &= 0 \end{aligned}$$

4. $$\begin{aligned} 2x_1 + x_2 - 3x_3 &= 1 \\ x_1 - 2x_2 + x_3 &= 0 \\ 3x_1 + 4x_2 - 2x_3 &= -5 \end{aligned}$$

5. $$\begin{aligned} x_1 - x_2 + 4x_3 &= -4 \\ -8x_1 + 3x_2 + x_3 &= 8 \\ 2x_1 - x_2 + x_3 &= 0 \end{aligned}$$

6. $$\begin{aligned} x_1 - x_2 + 4x_3 &= -2 \\ -8x_1 + 3x_2 + x_3 &= 0 \\ 2x_1 - x_2 + x_3 &= 6 \end{aligned}$$

7. $$\begin{aligned} 3x_1 + x_2 + x_3 &= 4 \\ -2x_1 - x_2 &= 12 \\ x_1 + 2x_2 + x_3 &= -8 \end{aligned}$$

8. Use Theorem 4.8 to prove a result analogous to Theorem 4.3 (p. 212), but for columns.

9. Prove that an upper triangular $n \times n$ matrix is invertible if and only if all its diagonal entries are nonzero.

10. A matrix $M \in \mathsf{M}_{n\times n}(F)$ is called **nilpotent** if, for some positive integer k, $M^k = O$, where O is the $n \times n$ zero matrix. Prove that if M is nilpotent, then $\det(M) = 0$.

11. A matrix $M \in \mathsf{M}_{n\times n}(C)$ is called **skew-symmetric** if $M^t = -M$. Prove that if M is skew-symmetric and n is odd, then M is not invertible. What happens if n is even?

12. A matrix $Q \in \mathsf{M}_{n\times n}(R)$ is called **orthogonal** if $QQ^t = I$. Prove that if Q is orthogonal, then $\det(Q) = \pm 1$.

13. For $M \in \mathsf{M}_{n\times n}(C)$, let $\overline{M}$ be the matrix such that $(\overline{M})_{ij} = \overline{M_{ij}}$ for all i, j, where $\overline{M_{ij}}$ is the complex conjugate of M_{ij}.

(a) Prove that $\det(\overline{M}) = \overline{\det(M)}$.

(b) A matrix $Q \in \mathsf{M}_{n\times n}(C)$ is called **unitary** if $QQ^* = I$, where $Q^* = \overline{Q^t}$. Prove that if Q is a unitary matrix, then $|\det(Q)| = 1$.

14. Let $\beta = \{u_1, u_2, \ldots, u_n\}$ be a subset of F^n containing n distinct vectors, and let B be the matrix in $\mathsf{M}_{n\times n}(F)$ having u_j as column j. Prove that β is a basis for F^n if and only if $\det(B) \neq 0$.

15.† Prove that if $A, B \in \mathsf{M}_{n\times n}(F)$ are similar, then $\det(A) = \det(B)$.

16. Use determinants to prove that if $A, B \in \mathsf{M}_{n\times n}(F)$ are such that $AB = I$, then A is invertible (and hence $B = A^{-1}$).

17. Let $A, B \in \mathsf{M}_{n\times n}(F)$ be such that $AB = -BA$. Prove that if n is odd and F is not a field of characteristic two, then A or B is not invertible.

18. Complete the proof of Theorem 4.7 by showing that if A is an elementary matrix of type 2 or type 3, then $\det(AB) = \det(A) \cdot \det(B)$.

19. A matrix $A \in \mathsf{M}_{n\times n}(F)$ is called **lower triangular** if $A_{ij} = 0$ for $1 \leq i < j \leq n$. Suppose that A is a lower triangular matrix. Describe $\det(A)$ in terms of the entries of A.

20. Suppose that $M \in \mathsf{M}_{n\times n}(F)$ can be written in the form

$$M = \begin{pmatrix} A & B \\ O & I \end{pmatrix}$$

where A is a square matrix. Prove that $\det(M) = \det(A)$.

21.† Prove that if $M \in \mathsf{M}_{n\times n}(F)$ can be written in the form

$$M = \begin{pmatrix} A & B \\ O & C \end{pmatrix},$$

where A and C are square matrices, then $\det(M) = \det(A) \cdot \det(C)$. Visit `goo.gl/4sG3iv` for a solution.

22. Let $\mathsf{T}\colon \mathsf{P}_n(F) \to \mathsf{F}^{n+1}$ be the linear transformation defined in Exercise 22 of Section 2.4 by $\mathsf{T}(f) = (f(c_0), f(c_1), \ldots, f(c_n))$, where $c_0, c_1, \ldots, c_n$ are distinct scalars in an infinite field F. Let β be the standard ordered basis for $\mathsf{P}_n(F)$ and γ be the standard ordered basis for F^{n+1}.

(a) Show that $M = [\mathsf{T}]_\beta^\gamma$ has the form

$$\begin{pmatrix} 1 & c_0 & c_0^2 & \cdots & c_0^n \\ 1 & c_1 & c_1^2 & \cdots & c_1^n \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & c_n & c_n^2 & \cdots & c_n^n \end{pmatrix}.$$

A matrix with this form is called a **Vandermonde matrix.**

(b) Use Exercise 22 of Section 2.4 to prove that $\det(M) \neq 0$.

(c) Prove that

$$\det(M) = \prod_{0 \le i < j \le n} (c_j - c_i),$$

the product of all terms of the form $c_j - c_i$ for $0 \le i < j \le n$.

23. Let $A \in \mathsf{M}_{n\times n}(F)$ be nonzero. For any m $(1 \le m \le n)$, an $m \times m$ **submatrix** is obtained by deleting any $n - m$ rows and any $n - m$ columns of A.

(a) Let k $(1 \le k \le n)$ denote the largest integer such that some $k \times k$ submatrix has a nonzero determinant. Prove that $\operatorname{rank}(A) = k$.

(b) Conversely, suppose that $\operatorname{rank}(A) = k$. Prove that there exists a $k \times k$ submatrix with a nonzero determinant.

24. Let $A \in \mathsf{M}_{n\times n}(F)$ have the form

$$A = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & a_0 \\ -1 & 0 & 0 & \cdots & 0 & a_1 \\ 0 & -1 & 0 & \cdots & 0 & a_2 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -1 & a_{n-1} \end{pmatrix}.$$

Compute $\det(A + tI)$, where I is the $n \times n$ identity matrix.

25. Let c_{jk} denote the cofactor of the row j, column k entry of the matrix $A \in \mathsf{M}_{n\times n}(F)$.

(a) Prove that if B is the matrix obtained from A by replacing column k by e_j, then $\det(B) = c_{jk}$.

(b) Show that for $1 \le j \le n$, we have

$$A\begin{pmatrix} c_{j1} \\ c_{j2} \\ \vdots \\ c_{jn} \end{pmatrix} = \det(A) \cdot e_j.$$

Hint: Apply Cramer's rule to $Ax = e_j$.

(c) Deduce that if C is the $n \times n$ matrix such that $C_{ij} = c_{ji}$, then $AC = [\det(A)]I$.

(d) Show that if $\det(A) \neq 0$, then $A^{-1} = [\det(A)]^{-1}C$.

The following definition is used in Exercises 26–27.

Definition. *The* **classical adjoint** *of a square matrix A is the transpose of the matrix whose ij-entry is the ij-cofactor of A.*

26. Find the classical adjoint of each of the following matrices.

(a) $\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$ (b) $\begin{pmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{pmatrix}$

(c) $\begin{pmatrix} -4 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 5 \end{pmatrix}$ (d) $\begin{pmatrix} 3 & 6 & 7 \\ 0 & 4 & 8 \\ 0 & 0 & 5 \end{pmatrix}$

(e) $\begin{pmatrix} 1-i & 0 & 0 \\ 4 & 3i & 0 \\ 2i & 1+4i & -1 \end{pmatrix}$ (f) $\begin{pmatrix} 7 & 1 & 4 \\ 6 & -3 & 0 \\ -3 & 5 & -2 \end{pmatrix}$

(g) $\begin{pmatrix} -1 & 2 & 5 \\ 8 & 0 & -3 \\ 4 & 6 & 1 \end{pmatrix}$ (h) $\begin{pmatrix} 3 & 2+i & 0 \\ -1+i & 0 & i \\ 0 & 1 & 3-2i \end{pmatrix}$

27. Let C be the classical adjoint of $A \in \mathsf{M}_{n \times n}(F)$. Prove the following statements.

(a) $\det(C) = [\det(A)]^{n-1}$.

(b) C^t is the classical adjoint of A^t.

(c) If A is an invertible upper triangular matrix, then C and A^{-1} are both upper triangular matrices.

28. Let $y_1, y_2, \ldots, y_n$ be linearly independent functions in C^∞. For each $y \in \mathsf{C}^\infty$, define $\mathsf{T}(y) \in \mathsf{C}^\infty$ by

$$[\mathsf{T}(y)](t) = \det\begin{pmatrix} y(t) & y_1(t) & y_2(t) & \cdots & y_n(t) \\ y'(t) & y_1'(t) & y_2'(t) & \cdots & y_n'(t) \\ \vdots & \vdots & \vdots & & \vdots \\ y^{(n)}(t) & y_1^{(n)}(t) & y_2^{(n)}(t) & \cdots & y_n^{(n)}(t) \end{pmatrix}.$$

The preceding determinant is called the **Wronskian** of $y, y_1, \ldots, y_n$.

(a) Prove that $\mathsf{T}\colon \mathsf{C}^\infty \to \mathsf{C}^\infty$ is a linear transformation.
(b) Prove that $\mathsf{N}(\mathsf{T})$ contains $\mathrm{span}(\{y_1, y_2, \ldots, y_n\})$.

4.4 SUMMARY—IMPORTANT FACTS ABOUT DETERMINANTS

In this section, we summarize the important properties of the determinant needed for the remainder of the text. The results contained in this section have been derived in Sections 4.2 and 4.3; consequently, the facts presented here are stated without proofs.

The **determinant** of an $n \times n$ matrix A having entries from a field F is a scalar in F, denoted by $\det(A)$ or $|A|$, and can be computed as follows.

1. If A is 1×1, then $\det(A) = A_{11}$, the single entry of A.

2. If A is 2×2, then $\det(A) = A_{11}A_{22} - A_{12}A_{21}$. For example,

$$\det\begin{pmatrix} -1 & 2 \\ 5 & 3 \end{pmatrix} = (-1)(3) - (2)(5) = -13.$$

3. If A is $n \times n$ for $n > 2$, then, for each i, we can evaluate the determinant by *cofactor expansion along row i* as

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} A_{ij} \cdot \det(\tilde{A}_{ij}),$$

or, for each j, we can evaluate the determinant *by cofactor expansion along column j* as

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} A_{ij} \cdot \det(\tilde{A}_{ij}),$$

where $\tilde{A}_{ij}$ is the $(n-1) \times (n-1)$ matrix obtained by deleting row i and column j from A.

In the formulas above, the scalar $(-1)^{i+j}\det(\tilde{A}_{ij})$ is called the **cofactor** of the row i column j entry of A. In this language, the determinant of A is evaluated as the sum of terms obtained by multiplying each entry of some row or column of A by the cofactor of that entry. Thus $\det(A)$ is expressed in terms of n determinants of $(n-1) \times (n-1)$ matrices. These determinants are then evaluated in terms of determinants of $(n-2) \times (n-2)$ matrices, and so forth, until 2×2 matrices are obtained. The determinants of the 2×2 matrices are then evaluated as in item 2.

7. An $n \times n$ matrix A is invertible if and only if $\det(A) \neq 0$. Furthermore, if A is invertible, then $\det(A^{-1}) = \dfrac{1}{\det(A)}$.

8. For any $n \times n$ matrix A, the determinants of A and A^t are equal.

For example, property 7 guarantees that the matrix A on page 233 is invertible because $\det(A) = 102$.

The final property, stated as Exercise 15 of Section 4.3, is used in Chapter 5. It is a simple consequence of properties 6 and 7.

9. If A and B are similar matrices, then $\det(A) = \det(B)$.

EXERCISES

1. Label the following statements as true or false.
   (a) The determinant of a square matrix may be computed by expanding the matrix along any row or column.
   (b) In evaluating the determinant of a matrix, it is wise to expand along a row or column containing the largest number of zero entries.
   (c) If two rows or columns of A are identical, then $\det(A) = 0$.
   (d) If B is a matrix obtained by interchanging two rows or two columns of A, then $\det(B) = \det(A)$.
   (e) If B is a matrix obtained by multiplying each entry of some row or column of A by a scalar, then $\det(B) = \det(A)$.
   (f) If B is a matrix obtained from A by adding a multiple of some row to a different row, then $\det(B) = \det(A)$.
   (g) The determinant of an upper triangular $n \times n$ matrix is the product of its diagonal entries.
   (h) For every $A \in \mathsf{M}_{n\times n}(F)$, $\det(A^t) = -\det(A)$.
   (i) If $A, B \in \mathsf{M}_{n\times n}(F)$, then $\det(AB) = \det(A) \cdot \det(B)$.
   (j) If Q is an invertible matrix, then $\det(Q^{-1}) = [\det(Q)]^{-1}$.
   (k) A matrix Q is invertible if and only if $\det(Q) \neq 0$.

2. Evaluate the determinant of the following 2×2 matrices.

   (a) $\begin{pmatrix} 4 & -5 \\ 2 & 3 \end{pmatrix}$ (b) $\begin{pmatrix} -1 & 7 \\ 3 & 8 \end{pmatrix}$

   (c) $\begin{pmatrix} 2+i & -1+3i \\ 1-2i & 3-i \end{pmatrix}$ (d) $\begin{pmatrix} 3 & 4i \\ -6i & 2i \end{pmatrix}$

3. Evaluate the determinant of the following matrices in the manner indicated.

(a) $\begin{pmatrix} 0 & 1 & 2 \\ -1 & 0 & -3 \\ 2 & 3 & 0 \end{pmatrix}$ along the first row

(b) $\begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 5 \\ -1 & 3 & 0 \end{pmatrix}$ along the first column

(c) $\begin{pmatrix} 0 & 1 & 2 \\ -1 & 0 & -3 \\ 2 & 3 & 0 \end{pmatrix}$ along the second column

(d) $\begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 5 \\ -1 & 3 & 0 \end{pmatrix}$ along the third row

(e) $\begin{pmatrix} 0 & 1+i & 2 \\ -2i & 0 & 1-i \\ 3 & 4i & 0 \end{pmatrix}$ along the third row

(f) $\begin{pmatrix} i & 2+i & 0 \\ -1 & 3 & 2i \\ 0 & -1 & 1-i \end{pmatrix}$ along the third column

(g) $\begin{pmatrix} 0 & 2 & 1 & 3 \\ 1 & 0 & -2 & 2 \\ 3 & -1 & 0 & 1 \\ -1 & 1 & 2 & 0 \end{pmatrix}$ along the fourth column

(h) $\begin{pmatrix} 1 & -1 & 2 & -1 \\ -3 & 4 & 1 & -1 \\ 2 & -5 & -3 & 8 \\ -2 & 6 & -4 & 1 \end{pmatrix}$ along the fourth row

4. Evaluate the determinant of the following matrices by any legitimate method.

(a) $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$

(b) $\begin{pmatrix} -1 & 3 & 2 \\ 4 & -8 & 1 \\ 2 & 2 & 5 \end{pmatrix}$

(c) $\begin{pmatrix} 0 & 1 & 1 \\ 1 & 2 & -5 \\ 6 & -4 & 3 \end{pmatrix}$

(d) $\begin{pmatrix} 1 & -2 & 3 \\ -1 & 2 & -5 \\ 3 & -1 & 2 \end{pmatrix}$

(e) $\begin{pmatrix} i & 2 & -1 \\ 3 & 1+i & 2 \\ -2i & 1 & 4-i \end{pmatrix}$

(f) $\begin{pmatrix} -1 & 2+i & 3 \\ 1-i & i & 1 \\ 3i & 2 & -1+i \end{pmatrix}$

(g) $\begin{pmatrix} 1 & 0 & -2 & 3 \\ -3 & 1 & 1 & 2 \\ 0 & 4 & -1 & 1 \\ 2 & 3 & 0 & 1 \end{pmatrix}$

(h) $\begin{pmatrix} 1 & -2 & 3 & -12 \\ -5 & 12 & -14 & 19 \\ -9 & 22 & -20 & 31 \\ -4 & 9 & -14 & 15 \end{pmatrix}$

5. Suppose that $M \in \mathsf{M}_{n\times n}(F)$ can be written in the form

$$M = \begin{pmatrix} A & B \\ O & I \end{pmatrix},$$

where A is a square matrix. Prove that $\det(M) = \det(A)$.

6. Prove that if $M \in \mathsf{M}_{n\times n}(F)$ can be written in the form

$$M = \begin{pmatrix} A & B \\ O & C \end{pmatrix},$$

where A and C are square matrices, then $\det(M) = \det(A) \cdot \det(C)$. Visit `goo.gl/pGMdpX` for a solution.

4.5* A CHARACTERIZATION OF THE DETERMINANT

In Sections 4.2 and 4.3, we showed that the determinant possesses a number of properties. In this section, we show that three of these properties completely characterize the determinant; that is, the only function $\delta\colon \mathsf{M}_{n\times n}(F) \to F$ having these three properties is the determinant. This characterization of the determinant is the one used in Section 4.1 to establish the relationship between $\det\begin{pmatrix} u \\ v \end{pmatrix}$ and the area of the parallelogram determined by u and v. The first of these properties that characterize the determinant is the one described in Theorem 4.3 (p. 212).

Definition. *A function $\delta\colon \mathsf{M}_{n\times n}(F) \to F$ is called an **n-linear function** if it is a linear function of each row of an $n \times n$ matrix when the remaining $n-1$ rows are held fixed, that is, δ is n-linear if, for every $r = 1, 2, \ldots, n$, we have*

$$\delta\begin{pmatrix} a_1 \\ \vdots \\ a_{r-1} \\ u+kv \\ a_{r+1} \\ \vdots \\ a_n \end{pmatrix} = \delta\begin{pmatrix} a_1 \\ \vdots \\ a_{r-1} \\ u \\ a_{r+1} \\ \vdots \\ a_n \end{pmatrix} + k\delta\begin{pmatrix} a_1 \\ \vdots \\ a_{r-1} \\ v \\ a_{r+1} \\ \vdots \\ a_n \end{pmatrix}$$

whenever k is a scalar and u, v, and each a_i are vectors in F^n.

Example 1

The function $\delta\colon \mathsf{M}_{n\times n}(F) \to F$ defined by $\delta(A) = 0$ for each $A \in \mathsf{M}_{n\times n}(F)$ is an n-linear function. ◆

Example 2

For $1 \le j \le n$, define $\delta_j\colon \mathsf{M}_{n\times n}(F) \to F$ by $\delta_j(A) = A_{1j}A_{2j}\cdots A_{nj}$ for each $A \in \mathsf{M}_{n\times n}(F)$; that is, $\delta_j(A)$ equals the product of the entries of column j of

Proof. Exercise. ∎

Theorem 4.12. *If $\delta\colon \mathsf{M}_{n\times n}(F) \to F$ is an alternating n-linear function such that $\delta(I) = 1$, then $\delta(A) = \det(A)$ for every $A \in \mathsf{M}_{n\times n}(F)$.*

Proof. Let $\delta\colon \mathsf{M}_{n\times n}(F) \to F$ be an alternating n-linear function such that $\delta(I) = 1$, and let $A \in \mathsf{M}_{n\times n}(F)$. If A has rank less than n, then by Corollary 2 to Theorem 4.10, $\delta(A) = 0$. Since the corollary to Theorem 4.6 (p. 217) gives $\det(A) = 0$, we have $\delta(A) = \det(A)$ in this case. If, on the other hand, A has rank n, then A is invertible and hence is the product of elementary matrices (Corollary 3 to Theorem 3.6 p. 158), say $A = E_m \cdots E_2E_1$. Since $\delta(I) = 1$, it follows from Corollary 3 to Theorem 4.10 and the facts on page 223 that $\delta(E) = \det(E)$ for every elementary matrix E. Hence by Theorems 4.11 and 4.7 (p. 223), we have

$$\begin{aligned}
\delta(A) &= \delta(E_m \cdots E_2E_1)\\
&= \det(E_m)\delta(E_{m-1}\cdot \cdots \cdot E_2\cdot E_1)\\
&= \cdots\\
&= \det(E_m)\cdot \cdots \cdot \det(E_2)\cdot \det(E_1)\\
&= \det(E_m \cdots E_2E_1)\\
&= \det(A).
\end{aligned}$$ ∎

Theorem 4.12 provides the desired characterization of the determinant: It is the unique function $\delta\colon \mathsf{M}_{n\times n}(F) \to F$ that is n-linear, is alternating, and has the property that $\delta(I) = 1$.

EXERCISES

1. Label the following statements as true or false.

(a) Any n-linear function $\delta\colon \mathsf{M}_{n\times n}(F) \to F$ is a linear transformation.
(b) Any n-linear function $\delta\colon \mathsf{M}_{n\times n}(F) \to F$ is a linear function of each row of an $n \times n$ matrix when the other $n-1$ rows are held fixed.
(c) If $\delta\colon \mathsf{M}_{n\times n}(F) \to F$ is an alternating n-linear function and the matrix $A \in \mathsf{M}_{n\times n}(F)$ has two identical rows, then $\delta(A) = 0$.
(d) If $\delta\colon \mathsf{M}_{n\times n}(F) \to F$ is an alternating n-linear function and B is obtained from $A \in \mathsf{M}_{n\times n}(F)$ by interchanging two rows of A, then $\delta(B) = \delta(A)$.
(e) There is a unique alternating n-linear function $\delta\colon \mathsf{M}_{n\times n}(F) \to F$.
(f) The function $\delta\colon \mathsf{M}_{n\times n}(F) \to F$ defined by $\delta(A) = 0$ for every $A \in \mathsf{M}_{n\times n}(F)$ is an alternating n-linear function.

2. Determine all the 1-linear functions $\delta\colon \mathsf{M}_{1\times 1}(F) \to F$.

Determine which of the functions $\delta\colon \mathsf{M}_{3\times 3}(F) \to F$ in Exercises 3–10 are 3-linear functions. Justify each answer.

3. $\delta(A) = k$, where k is any nonzero scalar

4. $\delta(A) = A_{22}$

5. $\delta(A) = A_{11}A_{23}A_{32}$

6. $\delta(A) = A_{11} + A_{23} + A_{32}$

7. $\delta(A) = A_{11}A_{21}A_{32}$

8. $\delta(A) = A_{11}A_{31}A_{32}$

9. $\delta(A) = A_{11}^2A_{22}^2A_{33}^2$

10. $\delta(A) = A_{11}A_{22}A_{33} - A_{11}A_{21}A_{32}$

11. Prove Corollaries 2 and 3 of Theorem 4.10. Visit goo.gl/FKcuqu for a solution.

12. Prove Theorem 4.11.

13. Prove that $\det\colon \mathsf{M}_{2\times 2}(F) \to F$ is a 2-linear function of the *columns* of a matrix.

14. Let $a, b, c, d \in F$. Prove that the function $\delta\colon \mathsf{M}_{2\times 2}(F) \to F$ defined by $\delta(A) = A_{11}A_{22}a + A_{11}A_{21}b + A_{12}A_{22}c + A_{12}A_{21}d$ is a 2-linear function.

15. Prove that $\delta\colon \mathsf{M}_{2\times 2}(F) \to F$ is a 2-linear function if and only if it has the form

$$\delta(A) = A_{11}A_{22}a + A_{11}A_{21}b + A_{12}A_{22}c + A_{12}A_{21}d$$

for some scalars $a, b, c, d \in F$.

16. Prove that if $\delta\colon \mathsf{M}_{n\times n}(F) \to F$ is an alternating n-linear function, then there exists a scalar k such that $\delta(A) = k\det(A)$ for all $A \in \mathsf{M}_{n\times n}(F)$.

17. Prove that a linear combination of two n-linear functions is an n-linear function, where the sum and scalar product of n-linear functions are as defined in Example 3 of Section 1.2 (p. 9).

18. Prove that the set of all n-linear functions over a field F is a vector space over F under the operations of function addition and scalar multiplication as defined in Example 3 of Section 1.2 (p. 9).

19. Let $\delta\colon \mathsf{M}_{n\times n}(F) \to F$ be an n-linear function and F a field that does not have characteristic two. Prove that if $\delta(B) = -\delta(A)$ whenever B is obtained from $A \in \mathsf{M}_{n\times n}(F)$ by interchanging any two rows of A, then $\delta(M) = 0$ whenever $M \in \mathsf{M}_{n\times n}(F)$ has two identical rows.

20. Give an example to show that the implication in Exercise 19 need not hold if F has characteristic two.

INDEX OF DEFINITIONS FOR CHAPTER 4

Suppose that $0 < \theta < \pi$. Then for any nonzero vector v, the vectors v and $\mathsf{T}_\theta(v)$ are not collinear, and hence T_θ maps no one-dimensional subspace of R^2 into itself. But this implies that T_θ has no eigenvectors and therefore no eigenvalues. To confirm this conclusion, let β be the standard ordered basis for R^2, and note that the characteristic polynomial of T_θ is

$$\det([\mathsf{T}_\theta]_\beta - tI_2) = \det\begin{pmatrix}\cos\theta - t & -\sin\theta\\ \sin\theta & \cos\theta - t\end{pmatrix} = t^2 - (2\cos\theta)t + 1,$$

which has no real zeros because, for $0 < \theta < \pi$, the discriminant $4\cos^2\theta - 4$ is negative.

EXERCISES

1. Label the following statements as true or false.
   (a) Every linear operator on an n-dimensional vector space has n distinct eigenvalues.
   (b) If a real matrix has one eigenvector, then it has an infinite number of eigenvectors.
   (c) There exists a square matrix with no eigenvectors.
   (d) Eigenvalues must be nonzero scalars.
   (e) Any two eigenvectors are linearly independent.
   (f) The sum of two eigenvalues of a linear operator T is also an eigenvalue of T.
   (g) Linear operators on infinite-dimensional vector spaces never have eigenvalues.
   (h) An $n \times n$ matrix A with entries from a field F is similar to a diagonal matrix if and only if there is a basis for F^n consisting of eigenvectors of A.
   (i) Similar matrices always have the same eigenvalues.
   (j) Similar matrices always have the same eigenvectors.
   (k) The sum of two eigenvectors of an operator T is always an eigenvector of T.

2. For each of the following linear operators T on a vector space V, compute the determinant of T and the characteristic polynomial of T.
   (a) $\mathsf{V} = \mathsf{R}^2$, $\mathsf{T}\begin{pmatrix}a\\ b\end{pmatrix} = \begin{pmatrix}2a - b\\ 5a + 3b\end{pmatrix}$
   (b) $\mathsf{V} = \mathsf{R}^3$, $\mathsf{T}\begin{pmatrix}a\\ b\\ c\end{pmatrix} = \begin{pmatrix}a - 3b + 2c\\ -2a + b + c\\ 4a - c\end{pmatrix}$

(c) $\mathsf{V} = \mathsf{P}_3(R)$,
$\mathsf{T}(a+bx+cx^2+dx^3) = (a-c)+(-a+b+d)x+(a+b-d)x^2-cx^3$

(d) $\mathsf{V} = \mathsf{M}_{2\times 2}(R)$, $\mathsf{T}(A) = 2A^t - A$

3. For each of the following linear operators T on a vector space V and ordered bases β, compute $[\mathsf{T}]_\beta$, and determine whether β is a basis consisting of eigenvectors of T.

(a) $\mathsf{V} = \mathsf{R}^2$, $\mathsf{T}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 10a - 6b \\ 17a - 10b \end{pmatrix}$, and $\beta = \left\{ \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \end{pmatrix} \right\}$

(b) $\mathsf{V} = \mathsf{P}_1(R)$, $\mathsf{T}(a+bx) = (6a-6b)+(12a-11b)x$, and $\beta = \{3+4x, 2+3x\}$

(c) $\mathsf{V} = \mathsf{R}^3$, $\mathsf{T}\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 3a+2b-2c \\ -4a-3b+2c \\ -c \end{pmatrix}$, and

$$\beta = \left\{ \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} \right\}$$

(d) $\mathsf{V} = \mathsf{P}_2(R)$, $\mathsf{T}(a+bx+cx^2) =$

$$(-4a+2b-2c)-(7a+3b+7c)x+(7a+b+5c)x^2,$$

and $\beta = \{x - x^2, -1+x^2, -1-x+x^2\}$

(e) $\mathsf{V} = \mathsf{P}_3(R)$, $\mathsf{T}(a+bx+cx^2+dx^3) =$

$$-d+(-c+d)x+(a+b-2c)x^2+(-b+c-2d)x^3,$$

and $\beta = \{1-x+x^3, 1+x^2, 1, x+x^2\}$

(f) $\mathsf{V} = \mathsf{M}_{2\times 2}(R)$, $\mathsf{T}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} -7a-4b+4c-4d & b \\ -8a-4b+5c-4d & d \end{pmatrix}$, and

$$\beta = \left\{ \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} -1 & 2 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 2 & 0 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & 2 \end{pmatrix} \right\}$$

4. For each of the following matrices $A \in \mathsf{M}_{n\times n}(F)$,

(i) Determine all the eigenvalues of A.

(ii) For each eigenvalue λ of A, find the set of eigenvectors corresponding to λ.

(iii) If possible, find a basis for F^n consisting of eigenvectors of A.

(iv) If successful in finding such a basis, determine an invertible matrix Q and a diagonal matrix D such that $Q^{-1}AQ = D$.

(a) $A = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}$ for $F = R$

(b) $A = \begin{pmatrix} 0 & -2 & -3 \\ -1 & 1 & -1 \\ 2 & 2 & 5 \end{pmatrix}$ for $F = R$

(c) $A = \begin{pmatrix} i & 1 \\ 2 & -i \end{pmatrix}$ for $F = C$

(d) $A = \begin{pmatrix} 2 & 0 & -1 \\ 4 & 1 & -4 \\ 2 & 0 & -1 \end{pmatrix}$ for $F = R$

5. For each linear operator T on V, find the eigenvalues of T and an ordered basis β for V such that $[\mathsf{T}]_\beta$ is a diagonal matrix.
   (a) $\mathsf{V} = \mathsf{R}^2$ and $\mathsf{T}(a, b) = (-2a + 3b, -10a + 9b)$
   (b) $\mathsf{V} = \mathsf{R}^3$ and $\mathsf{T}(a, b, c) = (7a - 4b + 10c, 4a - 3b + 8c, -2a + b - 2c)$
   (c) $\mathsf{V} = \mathsf{R}^3$ and $\mathsf{T}(a, b, c) = (-4a + 3b - 6c, 6a - 7b + 12c, 6a - 6b + 11c)$
   (d) $\mathsf{V} = \mathsf{P}_1(R)$ and $\mathsf{T}(ax + b) = (-6a + 2b)x + (-6a + b)$
   (e) $\mathsf{V} = \mathsf{P}_2(R)$ and $\mathsf{T}(f(x)) = xf'(x) + f(2)x + f(3)$
   (f) $\mathsf{V} = \mathsf{P}_3(R)$ and $\mathsf{T}(f(x)) = f(x) + f(2)x$
   (g) $\mathsf{V} = \mathsf{P}_3(R)$ and $\mathsf{T}(f(x)) = xf'(x) + f''(x) - f(2)$
   (h) $\mathsf{V} = \mathsf{M}_{2\times 2}(R)$ and $\mathsf{T}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} d & b \\ c & a \end{pmatrix}$
   (i) $\mathsf{V} = \mathsf{M}_{2\times 2}(R)$ and $\mathsf{T}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} c & d \\ a & b \end{pmatrix}$
   (j) $\mathsf{V} = \mathsf{M}_{2\times 2}(R)$ and $\mathsf{T}(A) = A^t + 2 \cdot \mathrm{tr}(A) \cdot I_2$

6. Prove Theorem 5.4.

7. Let T be a linear operator on a finite-dimensional vector space V, and let β be an ordered basis for V. Prove that λ is an eigenvalue of T if and only if λ is an eigenvalue of $[\mathsf{T}]_\beta$.

8. Let T be a linear operator on a finite-dimensional vector space V. Refer to the definition of the determinant of T on page 249 to prove the following results.
   (a) Prove that this definition is independent of the choice of an ordered basis for V. That is, prove that if β and γ are two ordered bases for V, then $\det([\mathsf{T}]_\beta) = \det([\mathsf{T}]_\gamma)$.
   (b) Prove that T is invertible if and only if $\det(\mathsf{T}) \neq 0$.
   (c) Prove that if T is invertible, then $\det(\mathsf{T}^{-1}) = [\det(\mathsf{T})]^{-1}$.
   (d) Prove that if U is also a linear operator on V, then $\det(\mathsf{TU}) = \det(\mathsf{T}) \cdot \det(\mathsf{U})$.

(e) Prove that $\det(\mathsf{T} - \lambda\mathsf{I}_\mathsf{V}) = \det([\mathsf{T}]_\beta - \lambda I)$ for any scalar λ and any ordered basis β for V.

9. (a) Prove that a linear operator T on a finite-dimensional vector space is invertible if and only if zero is not an eigenvalue of T.
(b) Let T be an invertible linear operator. Prove that a scalar λ is an eigenvalue of T if and only if λ^{-1} is an eigenvalue of T^{-1}.
(c) State and prove results analogous to (a) and (b) for matrices.

10. Prove that the eigenvalues of an upper triangular matrix M are the diagonal entries of M.

11. Let V be a finite-dimensional vector space, and let λ be any scalar.
(a) For any ordered basis β for V, prove that $[\lambda\mathsf{I}_\mathsf{V}]_\beta = \lambda I$.
(b) Compute the characteristic polynomial of $\lambda\mathsf{I}_\mathsf{V}$.
(c) Show that $\lambda\mathsf{I}_\mathsf{V}$ is diagonalizable and has only one eigenvalue.

12. A **scalar matrix** is a square matrix of the form λI for some scalar λ; that is, a scalar matrix is a diagonal matrix in which all the diagonal entries are equal.
(a) Prove that if a square matrix A is similar to a scalar matrix λI, then $A = \lambda I$.
(b) Show that a diagonalizable matrix having only one eigenvalue is a scalar matrix.
(c) Prove that $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ is not diagonalizable.

13. (a) Prove that similar matrices have the same characteristic polynomial.
(b) Show that the definition of the characteristic polynomial of a linear operator on a finite-dimensional vector space V is independent of the choice of basis for V.

14. Let T be a linear operator on a finite-dimensional vector space V over a field F, let β be an ordered basis for V, and let $A = [\mathsf{T}]_\beta$. In reference to Figure 5.1, prove the following.
(a) If $v \in \mathsf{V}$ and $\phi_\beta(v)$ is an eigenvector of A corresponding to the eigenvalue λ, then v is an eigenvector of T corresponding to λ.
(b) If λ is an eigenvalue of A (and hence of T), then a vector $y \in \mathsf{F}^n$ is an eigenvector of A corresponding to λ if and only if $\phi_\beta^{-1}(y)$ is an eigenvector of T corresponding to λ.

15.† For any square matrix A, prove that A and A^t have the same characteristic polynomial (and hence the same eigenvalues). Visit goo.gl/7Qss2u for a solution.

16.† (a) Let T be a linear operator on a vector space V, and let x be an eigenvector of T corresponding to the eigenvalue λ. For any positive integer m, prove that x is an eigenvector of T^m corresponding to the eigenvalue λ^m.

(b) State and prove the analogous result for matrices.

17. Let T be a linear operator on a finite-dimensional vector space V, and let c be any scalar.

(a) Determine the relationship between the eigenvalues and eigenvectors of T (if any) and the eigenvalues and eigenvectors of $\mathsf{U} = \mathsf{T} - c\mathsf{I}$. Justify your answers.

(b) Prove that T is diagonalizable if and only if U is diagonalizable.

18. Let T be the linear operator on $\mathsf{M}_{n\times n}(R)$ defined by $\mathsf{T}(A) = A^t$.

(a) Show that ± 1 are the only eigenvalues of T.

(b) Describe the eigenvectors corresponding to each eigenvalue of T.

(c) Find an ordered basis β for $\mathsf{M}_{2\times 2}(R)$ such that $[\mathsf{T}]_\beta$ is a diagonal matrix.

(d) Find an ordered basis β for $\mathsf{M}_{n\times n}(R)$ such that $[\mathsf{T}]_\beta$ is a diagonal matrix for $n > 2$.

19. Let $A, B \in \mathsf{M}_{n\times n}(C)$.

(a) Prove that if B is invertible, then there exists a scalar $c \in C$ such that $A + cB$ is not invertible. *Hint:* Examine $\det(A + cB)$.

(b) Find nonzero 2×2 matrices A and B such that both A and $A + cB$ are invertible for all $c \in C$.

20. Let A be an $n \times n$ matrix with characteristic polynomial

$$f(t) = (-1)^n t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0.$$

Prove that $f(0) = a_0 = \det(A)$. Deduce that A is invertible if and only if $a_0 \neq 0$.

21. Let A and $f(t)$ be as in Exercise 20.

(a) Prove that $f(t) = (A_{11} - t)(A_{22} - t)\cdots(A_{nn} - t) + q(t)$, where $q(t)$ is a polynomial of degree at most $n-2$. *Hint:* Apply mathematical induction to n.

(b) Show that $\mathrm{tr}(A) = (-1)^{n-1}a_{n-1}$.

22.† (a) Let T be a linear operator on a vector space V over the field F, and let $g(t)$ be a polynomial with coefficients from F. Prove that if x is an eigenvector of T with corresponding eigenvalue λ, then $g(\mathsf{T})(x) = g(\lambda)x$. That is, x is an eigenvector of $g(\mathsf{T})$ with corresponding eigenvalue $g(\lambda)$.

(b) State and prove a comparable result for matrices.

(c) Verify (b) for the matrix A in Exercise 4(a) with polynomial $g(t) = 2t^2 - t + 1$, eigenvector $x = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$, and corresponding eigenvalue $\lambda = 4$.

23. Use Exercise 22 to prove that if $f(t)$ is the characteristic polynomial of a diagonalizable linear operator T, then $f(\mathsf{T}) = \mathsf{T}_0$, the zero operator. (In Section 5.4 we prove that this result does not depend on the diagonalizability of T.)

24. Use Exercise 21(a) to prove Theorem 5.3.

25. Determine the number of distinct characteristic polynomials of matrices in $\mathsf{M}_{2\times 2}(Z_2)$.

5.2 DIAGONALIZABILITY

In Section 5.1, we presented the diagonalization problem and observed that not all linear operators or matrices are diagonalizable. Although we are able to diagonalize operators and matrices and even obtain a necessary and sufficient condition for diagonalizability (Theorem 5.1 p. 247), we have not yet solved the diagonalization problem. What is still needed is a simple test to determine whether an operator or a matrix can be diagonalized, as well as a method for actually finding a basis of eigenvectors. In this section, we develop such a test and method.

In Example 6 of Section 5.1, we obtained a basis of eigenvectors by choosing one eigenvector corresponding to each eigenvalue. In general, such a procedure does not yield a basis, but the following theorem shows that any set constructed in this manner is linearly independent.

Theorem 5.5. *Let T be a linear operator on a vector space, and let $\lambda_1, \lambda_2, \ldots, \lambda_k$ be distinct eigenvalues of T. For each $i = 1, 2, \ldots, k$, let S_i be a finite set of eigenvectors of T corresponding to λ_i. If each S_i $(i = 1, 2, \ldots, k)$, is linearly independent, then $S_1 \cup S_2 \cup \cdots \cup S_k$ is linearly independent.*

Proof. The proof is by mathematical induction on k. If $k = 1$, there is nothing to prove. So assume that the theorem holds for $k - 1$ distinct eigenvalues, where $k - 1 \geq 1$, and that we have k distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_k$ of T. For each $i = 1, 2, \ldots, k$, let $S_i = \{v_{i1}, v_{i2}, \ldots, v_{in_i}\}$ be a linearly independent set of eigenvectors of T corresponding to λ_i. We wish to show that $S = S_1 \cup S_2 \cup \cdots \cup S_k$ is linearly independent.

Consider any scalars $\{a_{ij}\}$, where $i = 1, 2, \ldots, k$ and $j = 1, 2, \ldots n_i$, such that

$$\sum_{i=1}^{k} \sum_{j=1}^{n_i} a_{ij} v_{ij} = 0. \tag{1}$$

Then

$$v \in \mathsf{W}_j = \operatorname{span}(\gamma_j) \quad \text{and} \quad v \in \sum_{i \neq j} \mathsf{W}_i = \operatorname{span}\left(\bigcup_{i \neq j} \gamma_i\right).$$

Hence v is a nontrivial linear combination of both γ_j and $\bigcup_{i \neq j} \gamma_i$, so that v can be expressed as a linear combination of $\gamma_1 \cup \gamma_2 \cup \cdots \cup \gamma_k$ in more than one way. But these representations contradict Theorem 1.8 (p. 44), and so we conclude that

$$\mathsf{W}_j \cap \sum_{i \neq j} \mathsf{W}_i = \{0\},$$

proving (a). ∎

With the aid of Theorem 5.9, we are able to characterize diagonalizability in terms of direct sums.

Theorem 5.10. *A linear operator* T *on a finite-dimensional vector space* V *is diagonalizable if and only if* V *is the direct sum of the eigenspaces of* T.

Proof. Let $\lambda_1, \lambda_2, \ldots, \lambda_k$ be the distinct eigenvalues of T.

First suppose that T is diagonalizable, and for each i choose an ordered basis γ_i for the eigenspace E_{λ_i}. By Theorem 5.8, $\gamma_1 \cup \gamma_2 \cup \cdots \cup \gamma_k$ is a basis for V, and hence V is a direct sum of the E_{λ_i}'s by Theorem 5.9.

Conversely, suppose that V is a direct sum of the eigenspaces of T. For each i, choose an ordered basis γ_i of E_{λ_i}. By Theorem 5.9, the union $\gamma_1 \cup \gamma_2 \cup \cdots \cup \gamma_k$ is a basis for V. Since this basis consists of eigenvectors of T, we conclude that T is diagonalizable. ∎

Example 10

Let T be the linear operator on R^4 defined by

$$\mathsf{T}(a, b, c, d) = (a, b, 2c, 3d).$$

It is easily seen that T is diagonalizable with eigenvalues $\lambda_1 = 1$, $\lambda_2 = 2$, and $\lambda_3 = 3$. Furthermore, the corresponding eigenspaces coincide with the subspaces W_1, W_2, and W_3 of Example 9. Thus Theorem 5.10 provides us with another proof that $\mathsf{R}^4 = \mathsf{W}_1 \oplus \mathsf{W}_2 \oplus \mathsf{W}_3$. ◆

EXERCISES

1. Label the following statements as true or false.
   (a) Any linear operator on an n-dimensional vector space that has fewer than n distinct eigenvalues is not diagonalizable.

(b) Two distinct eigenvectors corresponding to the same eigenvalue are always linearly dependent.
(c) If λ is an eigenvalue of a linear operator T, then each vector in E_λ is an eigenvector of T.
(d) If λ_1 and λ_2 are distinct eigenvalues of a linear operator T, then $\mathsf{E}_{\lambda_1} \cap \mathsf{E}_{\lambda_2} = \{0\}$.
(e) Let $A \in \mathsf{M}_{n\times n}(F)$ and $\beta = \{v_1, v_2, \ldots, v_n\}$ be an ordered basis for F^n consisting of eigenvectors of A. If Q is the $n \times n$ matrix whose jth column is v_j $(1 \le j \le n)$, then $Q^{-1}AQ$ is a diagonal matrix.
(f) A linear operator T on a finite-dimensional vector space is diagonalizable if and only if the multiplicity of each eigenvalue λ equals the dimension of E_λ.
(g) Every diagonalizable linear operator on a nonzero vector space has at least one eigenvalue.

The following two items relate to the optional subsection on direct sums.

(h) If a vector space is the direct sum of subspaces $\mathsf{W}_1, \mathsf{W}_2, \ldots, \mathsf{W}_k$, then $\mathsf{W}_i \cap \mathsf{W}_j = \{0\}$ for $i \neq j$.
(i) If
$$\mathsf{V} = \sum_{i=1}^{k} \mathsf{W}_i \quad \text{and} \quad \mathsf{W}_i \cap \mathsf{W}_j = \{0\} \quad \text{for } i \neq j,$$
then $\mathsf{V} = \mathsf{W}_1 \oplus \mathsf{W}_2 \oplus \cdots \oplus \mathsf{W}_k$.

2. For each of the following matrices $A \in \mathsf{M}_{n\times n}(R)$, test A for diagonalizability, and if A is diagonalizable, find an invertible matrix Q and a diagonal matrix D such that $Q^{-1}AQ = D$.

(a) $\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$ (b) $\begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix}$ (c) $\begin{pmatrix} 1 & 4 \\ 3 & 2 \end{pmatrix}$

(d) $\begin{pmatrix} 7 & -4 & 0 \\ 8 & -5 & 0 \\ 6 & -6 & 3 \end{pmatrix}$ (e) $\begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & -1 \\ 0 & 1 & 1 \end{pmatrix}$ (f) $\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 3 \end{pmatrix}$

(g) $\begin{pmatrix} 3 & 1 & 1 \\ 2 & 4 & 2 \\ -1 & -1 & 1 \end{pmatrix}$

3. For each of the following linear operators T on a vector space V, test T for diagonalizability, and if T is diagonalizable, find a basis β for V such that $[\mathsf{T}]_\beta$ is a diagonal matrix.

(a) $\mathsf{V} = \mathsf{P}_3(R)$ and T is defined by $\mathsf{T}(f(x)) = f'(x) + f''(x)$.
(b) $\mathsf{V} = \mathsf{P}_2(R)$ and T is defined by $\mathsf{T}(ax^2 + bx + c) = cx^2 + bx + a$.

(c) $V = R^3$ and T is defined by

$$T\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} a_2 \\ -a_1 \\ 2a_3 \end{pmatrix}.$$

(d) $V = P_2(R)$ and T is defined by $T(f(x)) = f(0) + f(1)(x + x^2)$.

(e) $V = C^2$ and T is defined by $T(z, w) = (z + iw, iz + w)$.

(f) $V = M_{2\times 2}(R)$ and T is defined by $T(A) = A^t$.

4. Prove the matrix version of the corollary to Theorem 5.5: If $A \in M_{n\times n}(F)$ has n distinct eigenvalues, then A is diagonalizable.

5. State and prove the matrix version of Theorem 5.6.

6. (a) Justify the test for diagonalizability and the method for diagonalization stated in this section.
   (b) Formulate the results in (a) for matrices.

7. For

$$A = \begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix} \in M_{2\times 2}(R),$$

find an expression for A^n, where n is an arbitrary positive integer.

8. Suppose that $A \in M_{n\times n}(F)$ has two distinct eigenvalues, λ_1 and λ_2, and that $\dim(E_{\lambda_1}) = n - 1$. Prove that A is diagonalizable.

9. Let T be a linear operator on a finite-dimensional vector space V, and suppose there exists an ordered basis β for V such that $[T]_\beta$ is an upper triangular matrix.
   (a) Prove that the characteristic polynomial for T splits.
   (b) State and prove an analogous result for matrices.

   The converse of (a) is treated in Exercise 12(b).

10. Let T be a linear operator on a finite-dimensional vector space V with the distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_k$ and corresponding multiplicities $m_1, m_2, \ldots, m_k$. Suppose that β is a basis for V such that $[T]_\beta$ is an upper triangular matrix. Prove that the diagonal entries of $[T]_\beta$ are $\lambda_1, \lambda_2, \ldots, \lambda_k$ and that each λ_i occurs m_i times $(1 \le i \le k)$.

11. Let A be an $n \times n$ matrix that is similar to an upper triangular matrix and has the distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_k$ with corresponding multiplicities $m_1, m_2, \ldots, m_k$. Prove the following statements.

   (a) $\operatorname{tr}(A) = \displaystyle\sum_{i=1}^{k} m_i\lambda_i$

(b) $\det(A) = (\lambda_1)^{m_1}(\lambda_2)^{m_2}\cdots(\lambda_k)^{m_k}$.

$\boxed{12.}$ (a) Prove that if $A \in \mathsf{M}_{n\times n}(F)$ and the characteristic polynomial of A splits, then A is similar to an upper triangular matrix. (This proves the converse of Exercise 9(b).) *Hint:* Use mathematical induction on n. For the general case, let v_1 be an eigenvector of A, and extend $\{v_1\}$ to a basis $\{v_1, v_2, \ldots, v_n\}$ for F^n. Let P be the $n \times n$ matrix whose jth column is v_j, and consider $P^{-1}AP$. Exercise 13(a) in Section 5.1 and Exercise 21 in Section 4.3 can be helpful.

(b) Prove the converse of Exercise 9(a).

Visit `goo.gl/gJSjRU` for a solution.

13. Let T be an invertible linear operator on a finite-dimensional vector space V.

(a) Recall that for any eigenvalue λ of T, λ^{-1} is an eigenvalue of T^{-1} (Exercise 9 of Section 5.1). Prove that the eigenspace of T corresponding to λ is the same as the eigenspace of T^{-1} corresponding to λ^{-1}.

(b) Prove that if T is diagonalizable, then T^{-1} is diagonalizable.

14. Let $A \in \mathsf{M}_{n\times n}(F)$. Recall from Exercise 15 of Section 5.1 that A and A^t have the same characteristic polynomial and hence share the same eigenvalues with the same multiplicities. For any eigenvalue λ of A and A^t, let E_λ and E'_λ denote the corresponding eigenspaces for A and A^t, respectively.

(a) Show by way of example that for a given common eigenvalue, these two eigenspaces need not be the same.

(b) Prove that for any eigenvalue λ, $\dim(\mathsf{E}_\lambda) = \dim(\mathsf{E}'_\lambda)$.

(c) Prove that if A is diagonalizable, then A^t is also diagonalizable.

15. Find the general solution to each system of differential equations.

(a) $\begin{aligned} x' &= x + y \\ y' &= 3x - y \end{aligned}$ (b) $\begin{aligned} x_1' &= 8x_1 + 10x_2 \\ x_2' &= -5x_1 - 7x_2 \end{aligned}$

(c) $\begin{aligned} x_1' &= x_1 \phantom{{}+x_2} + x_3 \\ x_2' &= \phantom{x_1+{}} x_2 + x_3 \\ x_3' &= \phantom{x_1+x_2+{}} 2x_3 \end{aligned}$

16. Let

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

be the coefficient matrix of the system of differential equations

$$\begin{aligned} x_1' &= a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ x_2' &= a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ &\vdots \\ x_n' &= a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n. \end{aligned}$$

Suppose that A is diagonalizable and that the distinct eigenvalues of A are $\lambda_1, \lambda_2, \ldots, \lambda_k$. Prove that a differentiable function $x\colon R \to \mathsf{R}^n$ is a solution to the system if and only if x is of the form

$$x(t) = e^{\lambda_1 t}z_1 + e^{\lambda_2 t}z_2 + \cdots + e^{\lambda_k t}z_k,$$

where $z_i \in \mathsf{E}_{\lambda_i}$ for $i = 1, 2, \ldots, k$. Use this result to prove that the set of solutions to the system is an n-dimensional real vector space.

17. Let $C \in \mathsf{M}_{m\times n}(R)$, and let Y be an $n \times p$ matrix of differentiable functions. Prove $(CY)' = CY'$, where $(Y')_{ij} = Y_{ij}'$ for all i, j.

Exercises 18 through 20 are concerned with *simultaneous diagonalization.*

Definitions. *Two linear operators* T *and* U *on a finite-dimensional vector space* V *are called* **simultaneously diagonalizable** *if there exists an ordered basis* β *for* V *such that both* $[\mathsf{T}]_\beta$ *and* $[\mathsf{U}]_\beta$ *are diagonal matrices. Similarly,* $A, B \in \mathsf{M}_{n\times n}(F)$ *are called* **simultaneously diagonalizable** *if there exists an invertible matrix* $Q \in \mathsf{M}_{n\times n}(F)$ *such that both* $Q^{-1}AQ$ *and* $Q^{-1}BQ$ *are diagonal matrices.*

18. (a) Prove that if T and U are simultaneously diagonalizable linear operators on a finite-dimensional vector space V, then the matrices $[\mathsf{T}]_\beta$ and $[\mathsf{U}]_\beta$ are simultaneously diagonalizable for any ordered basis β.
    (b) Prove that if A and B are simultaneously diagonalizable matrices, then L_A and L_B are simultaneously diagonalizable linear operators.

19. (a) Prove that if T and U are simultaneously diagonalizable operators, then T and U commute (i.e., $\mathsf{TU} = \mathsf{UT}$).
    (b) Show that if A and B are simultaneously diagonalizable matrices, then A and B commute.

    The converses of (a) and (b) are established in Exercise 25 of Section 5.4.

20. Let T be a diagonalizable linear operator on a finite-dimensional vector space, and let m be any positive integer. Prove that T and T^m are simultaneously diagonalizable.

Exercises 21 through 24 are concerned with direct sums.

21. Let $\mathsf{W}_1, \mathsf{W}_2, \ldots, \mathsf{W}_k$ be subspaces of a finite-dimensional vector space V such that

$$\sum_{i=1}^{k} \mathsf{W}_i = \mathsf{V}.$$

Prove that V is the direct sum of $\mathsf{W}_1, \mathsf{W}_2, \ldots, \mathsf{W}_k$ if and only if

$$\dim(\mathsf{V}) = \sum_{i=1}^{k} \dim(\mathsf{W}_i).$$

22. Let V be a finite-dimensional vector space with a basis β, and let $\beta_1, \beta_2, \ldots, \beta_k$ be a partition of β (i.e., $\beta_1, \beta_2, \ldots, \beta_k$ are subsets of β such that $\beta = \beta_1 \cup \beta_2 \cup \cdots \cup \beta_k$ and $\beta_i \cap \beta_j = \varnothing$ if $i \neq j$). Prove that $\mathsf{V} = \mathrm{span}(\beta_1) \oplus \mathrm{span}(\beta_2) \oplus \cdots \oplus \mathrm{span}(\beta_k)$.

23. Let T be a linear operator on a finite-dimensional vector space V, and suppose that the distinct eigenvalues of T are $\lambda_1, \lambda_2, \ldots, \lambda_k$. Prove that

$$\mathrm{span}(\{x \in \mathsf{V}\colon x \text{ is an eigenvector of } \mathsf{T}\}) = \mathsf{E}_{\lambda_1} \oplus \mathsf{E}_{\lambda_2} \oplus \cdots \oplus \mathsf{E}_{\lambda_k}.$$

24. Let $\mathsf{W}_1, \mathsf{W}_2, \mathsf{K}_1, \mathsf{K}_2, \ldots, \mathsf{K}_p, \mathsf{M}_1, \mathsf{M}_2, \ldots, \mathsf{M}_q$ be subspaces of a vector space V such that $\mathsf{W}_1 = \mathsf{K}_1 \oplus \mathsf{K}_2 \oplus \cdots \oplus \mathsf{K}_p$ and $\mathsf{W}_2 = \mathsf{M}_1 \oplus \mathsf{M}_2 \oplus \cdots \oplus \mathsf{M}_q$. Prove that if $\mathsf{W}_1 \cap \mathsf{W}_2 = \{0\}$, then

$$\mathsf{W}_1 + \mathsf{W}_2 = \mathsf{W}_1 \oplus \mathsf{W}_2 = \mathsf{K}_1 \oplus \mathsf{K}_2 \oplus \cdots \oplus \mathsf{K}_p \oplus \mathsf{M}_1 \oplus \mathsf{M}_2 \oplus \cdots \oplus \mathsf{M}_q.$$

5.3* MATRIX LIMITS AND MARKOV CHAINS

In this section, we apply what we have learned thus far in Chapter 5 to study the *limit* of a sequence of powers $A, A^2, \ldots, A^n, \ldots$, where A is a square matrix with complex entries. Such sequences and their limits have practical applications in the natural and social sciences.

We assume familiarity with limits of sequences of real numbers. The limit of a sequence of complex numbers $\{z_m\colon m = 1, 2, \ldots\}$ can be defined in terms of the limits of the sequences of the real and imaginary parts: If $z_m = r_m + is_m$, where r_m and s_m are real numbers, and i is the imaginary number such that $i^2 = -1$, then

$$\lim_{m\to\infty} z_m = \lim_{m\to\infty} r_m + i \lim_{m\to\infty} s_m,$$

provided that $\lim\limits_{m\to\infty} r_m$ and $\lim\limits_{m\to\infty} s_m$ exist.

EXERCISES

1. Label the following statements as true or false.
   (a) If $A \in \mathsf{M}_{n\times n}(C)$ and $\lim_{m\to\infty} A^m = L$, then, for any invertible matrix $Q \in \mathsf{M}_{n\times n}(C)$, we have $\lim_{m\to\infty} QA^mQ^{-1} = QLQ^{-1}$.
   (b) If 2 is an eigenvalue of $A \in \mathsf{M}_{n\times n}(C)$, then $\lim_{m\to\infty} A^m$ does not exist.
   (c) Any vector
$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathsf{R}^n$$
   such that $x_1 + x_2 + \cdots + x_n = 1$ is a probability vector.
   (d) The sum of the entries of each row of a transition matrix equals 1.
   (e) The product of a transition matrix and a probability vector is a probability vector.
   (f) Let z be any complex number such that $|z| < 1$. Then the matrix
$$\begin{pmatrix} 1 & z & -1 \\ z & 1 & 1 \\ -1 & 1 & z \end{pmatrix}$$
   does not have 3 as an eigenvalue.
   (g) Every transition matrix has 1 as an eigenvalue.
   (h) No transition matrix can have -1 as an eigenvalue.
   (i) If A is a transition matrix, then $\lim_{m\to\infty} A^m$ exists.
   (j) If A is a regular transition matrix, then $\lim_{m\to\infty} A^m$ exists and has rank 1.

2. Determine whether $\lim_{m\to\infty} A^m$ exists for each of the following matrices A, and compute the limit if it exists.

(a) $\begin{pmatrix} 0.1 & 0.7 \\ 0.7 & 0.1 \end{pmatrix}$ (b) $\begin{pmatrix} -1.4 & 0.8 \\ -2.4 & 1.8 \end{pmatrix}$ (c) $\begin{pmatrix} 0.4 & 0.7 \\ 0.6 & 0.3 \end{pmatrix}$

(d) $\begin{pmatrix} -1.8 & 4.8 \\ -0.8 & 2.2 \end{pmatrix}$ (e) $\begin{pmatrix} -2 & -1 \\ 4 & 3 \end{pmatrix}$ (f) $\begin{pmatrix} 2.0 & -0.5 \\ 3.0 & -0.5 \end{pmatrix}$

(g) $\begin{pmatrix} -1.8 & 0 & -1.4 \\ -5.6 & 1 & -2.8 \\ 2.8 & 0 & 2.4 \end{pmatrix}$ (h) $\begin{pmatrix} 3.4 & -0.2 & 0.8 \\ 3.9 & 1.8 & 1.3 \\ -16.5 & -2.0 & -4.5 \end{pmatrix}$

(i) $\begin{pmatrix} -\frac{1}{2}-2i & 4i & \frac{1}{2}+5i \\ 1+2i & -3i & -1-4i \\ -1-2i & 4i & 1+5i \end{pmatrix}$

(j) $\begin{pmatrix} \dfrac{-26+i}{3} & \dfrac{-28-4i}{3} & 28 \\ \dfrac{-7+2i}{3} & \dfrac{-5+i}{3} & 7-2i \\ \dfrac{-13+6i}{6} & \dfrac{-5+6i}{6} & \dfrac{35-20i}{6} \end{pmatrix}$

3. Prove that if $A_1, A_2, \ldots$ is a sequence of $n \times p$ matrices with complex entries such that $\lim_{m\to\infty} A_m = L$, then $\lim_{m\to\infty} (A_m)^t = L^t$.

4. Prove that if $A \in \mathsf{M}_{n\times n}(C)$ is diagonalizable and $L = \lim_{m\to\infty} A^m$ exists, then either $L = I_n$ or $\operatorname{rank}(L) < n$.

5. Find 2×2 matrices A and B having real entries such that $\lim_{m\to\infty} A^m$, $\lim_{m\to\infty} B^m$, and $\lim_{m\to\infty} (AB)^m$ all exist, but

$$\lim_{m\to\infty} (AB)^m \neq (\lim_{m\to\infty} A^m)(\lim_{m\to\infty} B^m).$$

6. In the week beginning June 1, 30% of the patients who arrived by helicopter at a hospital trauma unit were ambulatory and 70% were bedridden. One week after arrival, 60% of the ambulatory patients had been released, 20% remained ambulatory, and 20% had become bedridden. After the same amount of time, 10% of the bedridden patients had been released, 20% had become ambulatory, 50% remained bedridden, and 20% had died. Determine the percentages of helicopter arrivals during the week of June 1 who were in each of the four states one week after arrival. Assuming that the given percentages continue in the future, also determine the percentages of patients who eventually end up in each of the four states.

7. A player begins a game of chance by placing a marker in box 2, marked *Start.* (See Figure 5.5.) A die is rolled, and the marker is moved one square to the left if a 1 or a 2 is rolled and one square to the right if a 3, 4, 5, or 6 is rolled. This process continues until the marker lands in square 1, in which case the player wins the game, or in square 4, in which case the player loses the game. What is the probability of winning this game? *Hint:* Instead of diagonalizing the appropriate transition matrix A, it is easier to represent e_2 as a linear combination of eigenvectors of A and then apply A^n to the result.

Win	Start		Lose
1	2	3	4

Figure 5.5

8. Which of the following transition matrices are regular?

(a) $\begin{pmatrix} 0.2 & 0.3 & 0.5 \\ 0.3 & 0.2 & 0.5 \\ 0.5 & 0.5 & 0 \end{pmatrix}$ (b) $\begin{pmatrix} 0.5 & 0 & 1 \\ 0.5 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$ (c) $\begin{pmatrix} 0.5 & 0 & 0 \\ 0.5 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$

(d) $\begin{pmatrix} 0.5 & 0 & 1 \\ 0.5 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ (e) $\begin{pmatrix} \frac{1}{3} & 0 & 0 \\ \frac{1}{3} & 1 & 0 \\ \frac{1}{3} & 0 & 1 \end{pmatrix}$ (f) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0.7 & 0.2 \\ 0 & 0.3 & 0.8 \end{pmatrix}$

(g) $\begin{pmatrix} 0 & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & 0 & 0 & 0 \\ \frac{1}{4} & \frac{1}{4} & 1 & 0 \\ \frac{1}{4} & \frac{1}{4} & 0 & 1 \end{pmatrix}$ (h) $\begin{pmatrix} \frac{1}{4} & \frac{1}{4} & 0 & 0 \\ \frac{1}{4} & \frac{1}{4} & 0 & 0 \\ \frac{1}{4} & \frac{1}{4} & 1 & 0 \\ \frac{1}{4} & \frac{1}{4} & 0 & 1 \end{pmatrix}$

9. Compute $\lim_{m\to\infty} A^m$ if it exists, for each matrix A in Exercise 8.

10. Each of the matrices that follow is a regular transition matrix for a three-state Markov chain. In all cases, the initial probability vector is

$$P = \begin{pmatrix} 0.3 \\ 0.3 \\ 0.4 \end{pmatrix}.$$

For each transition matrix, compute the proportions of objects in each state after two stages and the eventual proportions of objects in each state by determining the fixed probability vector.

(a) $\begin{pmatrix} 0.6 & 0.1 & 0.1 \\ 0.1 & 0.9 & 0.2 \\ 0.3 & 0 & 0.7 \end{pmatrix}$ (b) $\begin{pmatrix} 0.8 & 0.1 & 0.2 \\ 0.1 & 0.8 & 0.2 \\ 0.1 & 0.1 & 0.6 \end{pmatrix}$ (c) $\begin{pmatrix} 0.9 & 0.1 & 0.1 \\ 0.1 & 0.6 & 0.1 \\ 0 & 0.3 & 0.8 \end{pmatrix}$

(d) $\begin{pmatrix} 0.4 & 0.2 & 0.2 \\ 0.1 & 0.7 & 0.2 \\ 0.5 & 0.1 & 0.6 \end{pmatrix}$ (e) $\begin{pmatrix} 0.5 & 0.3 & 0.2 \\ 0.2 & 0.5 & 0.3 \\ 0.3 & 0.2 & 0.5 \end{pmatrix}$ (f) $\begin{pmatrix} 0.6 & 0 & 0.4 \\ 0.2 & 0.8 & 0.2 \\ 0.2 & 0.2 & 0.4 \end{pmatrix}$

11. In 1940, a county land-use survey showed that 10% of the county land was urban, 50% was unused, and 40% was agricultural. Five years later,

a follow-up survey revealed that 70% of the urban land had remained urban, 10% had become unused, and 20% had become agricultural. Likewise, 20% of the unused land had become urban, 60% had remained unused, and 20% had become agricultural. Finally, the 1945 survey showed that 20% of the agricultural land had become unused while 80% remained agricultural. Assuming that the trends indicated by the 1945 survey continue, compute the percentages of urban, unused, and agricultural land in the county in 1950 and the corresponding eventual percentages.

12. A diaper liner is placed in each diaper worn by a baby. If, after a diaper change, the liner is soiled, then it is replaced by a new liner. Otherwise, the liner is washed with the diapers and reused, except that each liner is replaced by a new liner after its second use, even if it has never been soiled. The probability that the baby will soil any diaper liner is one-third. If there are only new diaper liners at first, eventually what proportions of the diaper liners being used will be new, once-used, and twice-used?

13. In 1975, the automobile industry determined that 40% of American car owners drove large cars, 20% drove intermediate-sized cars, and 40% drove small cars. A second survey in 1985 showed that 70% of the large-car owners in 1975 still owned large cars in 1985, but 30% had changed to an intermediate-sized car. Of those who owned intermediate-sized cars in 1975, 10% had switched to large cars, 70% continued to drive intermediate-sized cars, and 20% had changed to small cars in 1985. Finally, of the small-car owners in 1975, 10% owned intermediate-sized cars and 90% owned small cars in 1985. Assuming that these trends continue, determine the percentages of Americans who own cars of each size in 1995 and the corresponding eventual percentages.

14. Show that if A and P are as in Example 5, then

$$A^m = \begin{pmatrix} r_m & r_{m+1} & r_{m+1} \\ r_{m+1} & r_m & r_{m+1} \\ r_{m+1} & r_{m+1} & r_m \end{pmatrix},$$

where

$$r_m = \frac{1}{3}\left[1 + \frac{(-1)^m}{2^{m-1}}\right].$$

Deduce that

$$600(A^m P) = A^m \begin{pmatrix} 300 \\ 200 \\ 100 \end{pmatrix} = \begin{pmatrix} 200 + \dfrac{(-1)^m}{2^m}(100) \\ 200 \\ 200 + \dfrac{(-1)^{m+1}}{2^m}(100) \end{pmatrix}.$$

15. Prove that if a 1-dimensional subspace W of R^n contains a nonzero vector with all nonnegative entries, then W contains a unique probability vector.

16. Prove Theorem 5.14 and its corollary.

17. Prove the two corollaries of Theorem 5.17. Visit `goo.gl/V5Hsou` for a solution.

18. Prove the corollary of Theorem 5.18.

19. Suppose that M and M' are $n \times n$ transition matrices.

(a) Prove that if M is regular, N is any $n \times n$ transition matrix, and c is a real number such that $0 < c \leq 1$, then $cM + (1-c)N$ is a regular transition matrix.
(b) Suppose that for all i, j, we have that $M'_{ij} > 0$ whenever $M_{ij} > 0$. Prove that there exists a transition matrix N and a real number c with $0 < c \leq 1$ such that $M' = cM + (1-c)N$.
(c) Deduce that if the nonzero entries of M and M' occur in the same positions, then M is regular if and only if M' is regular.

The following definition is used in Exercises 20–24.

Definition. *For $A \in \mathsf{M}_{n\times n}(C)$, define $e^A = \lim_{m\to\infty} B_m$, where*

$$B_m = I + A + \frac{A^2}{2!} + \cdots + \frac{A^m}{m!}.$$

This limit exists by Exercise 22 of Section 7.2. Thus e^A is the sum of the infinite series

$$I + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \cdots,$$

and B_m is the mth partial sum of this series. (Note the analogy with the power series

$$e^a = 1 + a + \frac{a^2}{2!} + \frac{a^3}{3!} + \cdots,$$

which is valid for all complex numbers a.)

20. Compute e^O and e^I, where O and I denote the $n \times n$ zero and identity matrices, respectively.

21. Let $P^{-1}AP = D$ be a diagonal matrix. Prove that $e^A = Pe^D P^{-1}$.

22. Let $A \in \mathsf{M}_{n\times n}(C)$ be diagonalizable. Use the result of Exercise 21 to show that e^A exists. (Exercise 22 of Section 7.2 shows that e^A exists for every $A \in \mathsf{M}_{n\times n}(C)$.)

23. Find $A, B \in \mathsf{M}_{2\times 2}(R)$ such that $e^A e^B \neq e^{A+B}$.

24. Prove that a differentiable function $x\colon R \to \mathsf{R}^n$ is a solution to the system of differential equations defined in Exercise 16 of Section 5.2 if and only if $x(t) = e^{tA}v$ for some $v \in \mathsf{R}^n$, where A is defined in that exercise.

5.4 INVARIANT SUBSPACES AND THE CAYLEY–HAMILTON THEOREM

In Section 5.1, we observed that if v is an eigenvector of a linear operator T, then T maps the span of $\{v\}$ into itself. Subspaces that are mapped into themselves are of great importance in the study of linear operators (see, e.g., Exercises 29–33 of Section 2.1).

Definition. *Let* T *be a linear operator on a vector space* V. *A subspace* W *of* V *is called a* T-***invariant subspace*** *of* V *if* $\mathsf{T}(\mathsf{W}) \subseteq \mathsf{W}$, *that is, if* $\mathsf{T}(v) \in \mathsf{W}$ *for all* $v \in \mathsf{W}$.

Example 1

Suppose that T is a linear operator on a vector space V. Then the following subspaces of V are T-invariant:

1. $\{\mathit{0}\}$
2. V
3. $\mathsf{R}(\mathsf{T})$
4. $\mathsf{N}(\mathsf{T})$
5. E_λ, for any eigenvalue λ of T.

The proofs that these subspaces are T-invariant are left as exercises. (See Exercise 3.) ♦

Example 2

Let T be the linear operator on R^3 defined by

$$\mathsf{T}(a, b, c) = (a + b, b + c, 0).$$

Then the xy-plane $= \{(x, y, 0) \colon x, y \in R\}$ and the x-axis $= \{(x, 0, 0) \colon x \in R\}$ are T-invariant subspaces of R^3. ♦

Example 9

Let

$$B_1 = \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}, \quad B_2 = (3), \quad \text{and} \quad B_3 = \begin{pmatrix} 1 & 2 & 1 \\ 1 & 2 & 3 \\ 1 & 1 & 1 \end{pmatrix}.$$

Then

$$B_1 \oplus B_2 \oplus B_3 = \left(\begin{array}{cc|c|ccc} 1 & 2 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 3 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{array}\right). \quad \blacklozenge$$

The final result of this section relates direct sums of matrices to direct sums of invariant subspaces. It is an extension of Exercise 33 to the case $k \geq 2$.

Theorem 5.24. *Let* T *be a linear operator on a finite-dimensional vector space* V, *and let* $\mathsf{W}_1, \mathsf{W}_2, \ldots, \mathsf{W}_k$ *be* T*-invariant subspaces of* V *such that* $\mathsf{V} = \mathsf{W}_1 \oplus \mathsf{W}_2 \oplus \cdots \oplus \mathsf{W}_k$. *For each* i, *let* β_i *be an ordered basis for* W_i, *and let* $\beta = \beta_1 \cup \beta_2 \cup \cdots \cup \beta_k$. *Let* $A = [\mathsf{T}]_\beta$ *and* $B_i = [\mathsf{T}_{\mathsf{W}_i}]_{\beta_i}$ *for* $i = 1, 2, \ldots, k$. *Then* $A = B_1 \oplus B_2 \oplus \cdots \oplus B_k$.

Proof. See Exercise 34. ∎

EXERCISES

1. Label the following statements as true or false.
   (a) There exists a linear operator T with no T-invariant subspace.
   (b) If T is a linear operator on a finite-dimensional vector space V and W is a T-invariant subspace of V, then the characteristic polynomial of T_W divides the characteristic polynomial of T.
   (c) Let T be a linear operator on a finite-dimensional vector space V, and let v and w be in V. If W is the T-cyclic subspace generated by v, W' is the T-cyclic subspace generated by w, and $\mathsf{W} = \mathsf{W}'$, then $v = w$.
   (d) If T is a linear operator on a finite-dimensional vector space V, then for any $v \in \mathsf{V}$ the T-cyclic subspace generated by v is the same as the T-cyclic subspace generated by $\mathsf{T}(v)$.
   (e) Let T be a linear operator on an n-dimensional vector space. Then there exists a polynomial $g(t)$ of degree n such that $g(\mathsf{T}) = \mathsf{T}_0$.
   (f) Any polynomial of degree n with leading coefficient $(-1)^n$ is the characteristic polynomial of some linear operator.

(g) If T is a linear operator on a finite-dimensional vector space V, and if V is the direct sum of k T-invariant subspaces, then there is an ordered basis β for V such that $[\mathsf{T}]_\beta$ is a direct sum of k matrices.

2. For each of the following linear operators T on the vector space V, determine whether the given subspace W is a T-invariant subspace of V.

(a) $\mathsf{V} = \mathsf{P}_3(R)$, $\mathsf{T}(f(x)) = f'(x)$, and $\mathsf{W} = \mathsf{P}_2(R)$

(b) $\mathsf{V} = \mathsf{P}(R)$, $\mathsf{T}(f(x)) = xf(x)$, and $\mathsf{W} = \mathsf{P}_2(R)$

(c) $\mathsf{V} = \mathsf{R}^3$, $\mathsf{T}(a, b, c) = (a + b + c, a + b + c, a + b + c)$, and $\mathsf{W} = \{(t, t, t) \colon t \in R\}$

(d) $\mathsf{V} = \mathsf{C}([0, 1])$, $\mathsf{T}(f(t)) = \left[\int_0^1 f(x)\, dx\right] t$, and $\mathsf{W} = \{f \in \mathsf{V} \colon f(t) = at + b \text{ for some } a \text{ and } b\}$

(e) $\mathsf{V} = \mathsf{M}_{2\times 2}(R)$, $\mathsf{T}(A) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} A$, and $\mathsf{W} = \{A \in \mathsf{V} \colon A^t = A\}$

3. Let T be a linear operator on a finite-dimensional vector space V. Prove that the following subspaces are T-invariant.

(a) $\{0\}$ and V

(b) N(T) and R(T)

(c) E_λ, for any eigenvalue λ of T

4. Let T be a linear operator on a vector space V, and let W be a T-invariant subspace of V. Prove that W is $g(\mathsf{T})$-invariant for any polynomial $g(t)$.

5. Let T be a linear operator on a vector space V. Prove that the intersection of any collection of T-invariant subspaces of V is a T-invariant subspace of V.

6. For each linear operator T on the vector space V, find an ordered basis for the T-cyclic subspace generated by the vector z.

(a) $\mathsf{V} = \mathsf{R}^4$, $\mathsf{T}(a, b, c, d) = (a + b, b - c, a + c, a + d)$, and $z = e_1$.

(b) $\mathsf{V} = \mathsf{P}_3(R)$, $\mathsf{T}(f(x)) = f''(x)$, and $z = x^3$.

(c) $\mathsf{V} = \mathsf{M}_{2\times 2}(R)$, $\mathsf{T}(A) = A^t$, and $z = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

(d) $\mathsf{V} = \mathsf{M}_{2\times 2}(R)$, $\mathsf{T}(A) = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix} A$, and $z = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

7. Prove that the restriction of a linear operator T to a T-invariant subspace is a linear operator on that subspace.

8. Let T be a linear operator on a vector space with a T-invariant subspace W. Prove that if v is an eigenvector of T_W with corresponding eigenvalue λ, then v is also an eigenvector of T with corresponding eigenvalue λ.

9. For each linear operator T and cyclic subspace W in Exercise 6, compute the characteristic polynomial of $\mathsf{T_W}$ in two ways, as in Example 6.

10. For each linear operator in Exercise 6, find the characteristic polynomial $f(t)$ of T, and verify that the characteristic polynomial of $\mathsf{T_W}$ (computed in Exercise 9) divides $f(t)$.

11. Let T be a linear operator on a vector space V, let v be a nonzero vector in V, and let W be the T-cyclic subspace of V generated by v. Prove that

(a) W is T-invariant.
(b) Any T-invariant subspace of V containing v also contains W.

12. Prove that $A = \begin{pmatrix} B_1 & B_2 \\ O & B_3 \end{pmatrix}$ in the proof of Theorem 5.20.

13. Let T be a linear operator on a vector space V, let v be a nonzero vector in V, and let W be the T-cyclic subspace of V generated by v. For any $w \in \mathsf{V}$, prove that $w \in \mathsf{W}$ if and only if there exists a polynomial $g(t)$ such that $w = g(\mathsf{T})(v)$.

14. Prove that the polynomial $g(t)$ of Exercise 13 can always be chosen so that its degree is less than $\dim(\mathsf{W})$.

15. Use the Cayley–Hamilton theorem (Theorem 5.22) to prove its corollary for matrices. *Warning:* If $f(t) = \det(A - tI)$ is the characteristic polynomial of A, it is tempting to "prove" that $f(A) = O$ by saying "$f(A) = \det(A - AI) = \det(O) = 0$." Why is this argument incorrect? Visit `goo.gl/ZMVn9i` for a solution.

16. Let T be a linear operator on a finite-dimensional vector space V.

(a) Prove that if the characteristic polynomial of T splits, then so does the characteristic polynomial of the restriction of T to any T-invariant subspace of V.
(b) Deduce that if the characteristic polynomial of T splits, then any nontrivial T-invariant subspace of V contains an eigenvector of T.

17. Let A be an $n \times n$ matrix. Prove that

$$\dim(\operatorname{span}(\{I_n, A, A^2, \ldots\})) \leq n.$$

18. Let A be an $n \times n$ matrix with characteristic polynomial

$$f(t) = (-1)^n t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0.$$

(a) Prove that A is invertible if and only if $a_0 \neq 0$.

(b) Prove that if A is invertible, then

$$A^{-1} = (-1/a_0)[(-1)^n A^{n-1} + a_{n-1}A^{n-2} + \cdots + a_1 I_n].$$

(c) Use (b) to compute A^{-1} for

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 2 & 3 \\ 0 & 0 & -1 \end{pmatrix}.$$

19. Let A denote the $k \times k$ matrix

$$\begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & \cdots & 0 & -a_2 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_{k-1} \end{pmatrix},$$

where $a_0, a_1, \ldots, a_{k-1}$ are arbitrary scalars. Prove that the characteristic polynomial of A is

$$(-1)^k(a_0 + a_1 t + \cdots + a_{k-1}t^{k-1} + t^k).$$

Hint: Use mathematical induction on k, computing the determinant by cofactor expansion along the first row.

20. Let T be a linear operator on a vector space V, and suppose that V is a T-cyclic subspace of itself. Prove that if U is a linear operator on V, then UT = TU if and only if U = g(T) for some polynomial $g(t)$. *Hint:* Suppose that V is generated by v. Choose $g(t)$ according to Exercise 13 so that $g(\mathsf{T})(v) = \mathsf{U}(v)$.

21. Let T be a linear operator on a two-dimensional vector space V. Prove that either V is a T-cyclic subspace of itself or T = cI for some scalar c.

22. Let T be a linear operator on a two-dimensional vector space V and suppose that T $\neq c$I for any scalar c. Show that if U is any linear operator on V such that UT = TU, then U = g(T) for some polynomial $g(t)$.

23. Let T be a linear operator on a finite-dimensional vector space V, and let W be a T-invariant subspace of V. Suppose that $v_1, v_2, \ldots, v_k$ are eigenvectors of T corresponding to distinct eigenvalues. Prove that if $v_1 + v_2 + \cdots + v_k$ is in W, then $v_i \in$ W for all i. *Hint:* Use mathematical induction on k.

24. Prove that the restriction of a diagonalizable linear operator T to any nontrivial T-invariant subspace is also diagonalizable. *Hint:* Use the result of Exercise 23.

25. (a) Prove the converse to Exercise 19(a) of Section 5.2: If T and U are diagonalizable linear operators on a finite-dimensional vector space V such that UT = TU, then T and U are simultaneously diagonalizable. (See the definitions in the exercises of Section 5.2.) *Hint:* For any eigenvalue λ of T, show that E_λ is U-invariant, and apply Exercise 24 to obtain a basis for E_λ of eigenvectors of U.

(b) State and prove a matrix version of (a).

26. Let T be a linear operator on an n-dimensional vector space V such that T has n distinct eigenvalues. Prove that V is a T-cyclic subspace of itself. *Hint:* Use Exercise 23 to find a vector v such that $\{v, \mathsf{T}(v), \ldots, \mathsf{T}^{n-1}(v)\}$ is linearly independent.

Exercises 27 through 31 require familiarity with quotient spaces as defined in Exercise 31 of Section 1.3. Before attempting these exercises, the reader should first review the other exercises treating quotient spaces: Exercise 35 of Section 1.6, Exercise 42 of Section 2.1, and Exercise 24 of Section 2.4.

For the purposes of Exercises 27 through 31, T is a fixed linear operator on a finite-dimensional vector space V, and W is a nonzero T-invariant subspace of V. We require the following definition.

Definition. *Let* T *be a linear operator on a vector space* V, *and let* W *be a* T*-invariant subspace of* V. *Define* $\overline{\mathsf{T}}\colon \mathsf{V}/\mathsf{W} \to \mathsf{V}/\mathsf{W}$ *by*

$$\overline{\mathsf{T}}(v+\mathsf{W}) = \mathsf{T}(v)+\mathsf{W} \quad \text{for any } v+\mathsf{W} \in \mathsf{V}/\mathsf{W}.$$

27. (a) Prove that $\overline{\mathsf{T}}$ is well defined. That is, show that $\overline{\mathsf{T}}(v+\mathsf{W}) = \overline{\mathsf{T}}(v'+\mathsf{W})$ whenever $v+\mathsf{W} = v'+\mathsf{W}$.

(b) Prove that $\overline{\mathsf{T}}$ is a linear operator on V/W.

(c) Let $\eta\colon \mathsf{V} \to \mathsf{V}/\mathsf{W}$ be the linear transformation defined in Exercise 42 of Section 2.1 by $\eta(v) = v+\mathsf{W}$. Show that the diagram of Figure 5.6 commutes; that is, prove that $\eta\mathsf{T} = \overline{\mathsf{T}}\eta$. (This exercise does not require the assumption that V is finite-dimensional.)

$$\begin{tikzcd} \mathsf{V} \arrow[r, "\mathsf{T}"] \arrow[d, "\eta"'] & \mathsf{V} \arrow[d, "\eta"] \\ \mathsf{V}/\mathsf{W} \arrow[r, "\overline{\mathsf{T}}"] & \mathsf{V}/\mathsf{W} \end{tikzcd}$$

Figure 5.6

28. Let $f(t)$, $g(t)$, and $h(t)$ be the characteristic polynomials of T, $\mathsf{T_W}$, and $\overline{\mathsf{T}}$, respectively. Prove that $f(t) = g(t)h(t)$. *Hint:* Extend an ordered basis $\gamma = \{v_1, v_2, \ldots, v_k\}$ for W to an ordered basis $\beta = \{v_1, v_2, \ldots, v_k, v_{k+1}, \ldots, v_n\}$ for V. Then show that the collection of cosets $\alpha = \{v_{k+1} + \mathsf{W}, v_{k+2} + \mathsf{W}, \ldots, v_n + \mathsf{W}\}$ is an ordered basis for V/W, and prove that

$$[\mathsf{T}]_\beta = \begin{pmatrix} B_1 & B_2 \\ O & B_3 \end{pmatrix},$$

where $B_1 = [\mathsf{T}]_\gamma$ and $B_3 = [\overline{\mathsf{T}}]_\alpha$.

29. Use the hint in Exercise 28 to prove that if T is diagonalizable, then so is $\overline{\mathsf{T}}$.

30. Prove that if both $\mathsf{T_W}$ and $\overline{\mathsf{T}}$ are diagonalizable and have no common eigenvalues, then T is diagonalizable.

The results of Theorem 5.21 and Exercise 28 are useful in devising methods for computing characteristic polynomials without the use of determinants. This is illustrated in the next exercise.

31. Let $A = \begin{pmatrix} 1 & 1 & -3 \\ 2 & 3 & 4 \\ 1 & 2 & 1 \end{pmatrix}$, let $\mathsf{T} = \mathsf{L}_A$, and let W be the cyclic subspace of R^3 generated by e_1.

(a) Use Theorem 5.21 to compute the characteristic polynomial of $\mathsf{T_W}$.
(b) Show that $\{e_2 + \mathsf{W}\}$ is a basis for R^3/W, and use this fact to compute the characteristic polynomial of $\overline{\mathsf{T}}$.
(c) Use the results of (a) and (b) to find the characteristic polynomial of A.

Exercises 32 through 39 are concerned with direct sums.

32. Let T be a linear operator on a vector space V, and let $\mathsf{W}_1, \mathsf{W}_2, \ldots, \mathsf{W}_k$ be T-invariant subspaces of V. Prove that $\mathsf{W}_1 + \mathsf{W}_2 + \cdots + \mathsf{W}_k$ is also a T-invariant subspace of V.

33. Give a direct proof of Theorem 5.24 for the case $k = 2$. (This result is used in the proof of Theorem 5.23.)

34. Prove Theorem 5.24. *Hint:* Begin with Exercise 33 and extend it using mathematical induction on k, the number of subspaces.

35. Let T be a linear operator on a finite-dimensional vector space V. Prove that T is diagonalizable if and only if V is the direct sum of one-dimensional T-invariant subspaces.

36. Let T be a linear operator on a finite-dimensional vector space V, and let $W_1, W_2, \ldots, W_k$ be T-invariant subspaces of V such that $V = W_1 \oplus W_2 \oplus \cdots \oplus W_k$. Prove that

$$\det(T) = \det(T_{W_1}) \cdot \det(T_{W_2}) \cdot \cdots \cdot \det(T_{W_k}).$$

37. Let T be a linear operator on a finite-dimensional vector space V, and let $W_1, W_2, \ldots, W_k$ be T-invariant subspaces of V such that $V = W_1 \oplus W_2 \oplus \cdots \oplus W_k$. Prove that T is diagonalizable if and only if T_{W_i} is diagonalizable for all i.

38. Let $\mathcal{C}$ be a collection of diagonalizable linear operators on a finite-dimensional vector space V. Prove that there is an ordered basis β such that $[T]_\beta$ is a diagonal matrix for all $T \in \mathcal{C}$ if and only if the operators of $\mathcal{C}$ commute under composition. (This is an extension of Exercise 25.) *Hints for the case that the operators commute:* The result is trivial if each operator has only one eigenvalue. Otherwise, establish the general result by mathematical induction on $\dim(V)$, using the fact that V is the direct sum of the eigenspaces of some operator in $\mathcal{C}$ that has more than one eigenvalue.

39. Let $B_1, B_2, \ldots, B_k$ be square matrices with entries in the same field, and let $A = B_1 \oplus B_2 \oplus \cdots \oplus B_k$. Prove that the characteristic polynomial of A is the product of the characteristic polynomials of the B_i's.

40. Let

$$A = \begin{pmatrix} 1 & 2 & \cdots & n \\ n+1 & n+2 & \cdots & 2n \\ \vdots & \vdots & & \vdots \\ n^2-n+1 & n^2-n+2 & \cdots & n^2 \end{pmatrix}.$$

Find the characteristic polynomial of A. *Hint:* First prove that A has rank 2 and that $\text{span}(\{(1,1,\ldots,1),(1,2,\ldots,n)\})$ is L_A-invariant.

41. Let $A \in M_{n\times n}(R)$ be the matrix defined by $A_{ij} = 1$ for all i and j. Find the characteristic polynomial of A.

INDEX OF DEFINITIONS FOR CHAPTER 5

Also,

$$\langle f_n, f_n \rangle = \frac{1}{2\pi}\int_0^{2\pi} e^{i(n-n)t}\,dt = \frac{1}{2\pi}\int_0^{2\pi} 1\,dt = 1.$$

In other words, $\langle f_m, f_n \rangle = \delta_{mn}$. ◆

EXERCISES

1. Label the following statements as true or false.
   (a) An inner product is a scalar-valued function on the set of ordered pairs of vectors.
   (b) An inner product space must be over the field of real or complex numbers.
   (c) An inner product is linear in both components.
   (d) There is exactly one inner product on the vector space R^n.
   (e) The triangle inequality only holds in finite-dimensional inner product spaces.
   (f) Only square matrices have a conjugate-transpose.
   (g) If x, y, and z are vectors in an inner product space such that $\langle x, y \rangle = \langle x, z \rangle$, then $y = z$.
   (h) If $\langle x, y \rangle = 0$ for all x in an inner product space, then $y = \mathit{0}$.

2. Let $x = (2, 1+i, i)$ and $y = (2-i, 2, 1+2i)$ be vectors in C^3. Compute $\langle x, y \rangle$, $\|x\|$, $\|y\|$, and $\|x+y\|$. Then verify both the Cauchy–Schwarz inequality and the triangle inequality.

3. In $\mathsf{C}([0,1])$, let $f(t) = t$ and $g(t) = e^t$. Compute $\langle f, g \rangle$ (as defined in Example 3), $\|f\|$, $\|g\|$, and $\|f+g\|$. Then verify both the Cauchy–Schwarz inequality and the triangle inequality.

4. (a) Complete the proof in Example 5 that $\langle \cdot, \cdot \rangle$ is an inner product (the Frobenius inner product) on $\mathsf{M}_{n\times n}(F)$.
   (b) Use the Frobenius inner product to compute $\|A\|$, $\|B\|$, and $\langle A, B \rangle$ for

$$A = \begin{pmatrix} 1 & 2+i \\ 3 & i \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 1+i & 0 \\ i & -i \end{pmatrix}.$$

5. In C^2, show that $\langle x, y \rangle = xAy^*$ is an inner product, where

$$A = \begin{pmatrix} 1 & i \\ -i & 2 \end{pmatrix}.$$

   Compute $\langle x, y \rangle$ for $x = (1-i, 2+3i)$ and $y = (2+i, 3-2i)$.

6. Complete the proof of Theorem 6.1.

7. Complete the proof of Theorem 6.2.

8. Provide reasons why each of the following is not an inner product on the given vector spaces.

 (a) $\langle (a,b),(c,d)\rangle = ac - bd$ on R^2.
 (b) $\langle A,B\rangle = \operatorname{tr}(A+B)$ on $\mathsf{M}_{2\times 2}(R)$.
 (c) $\langle f(x),g(x)\rangle = \int_0^1 f'(t)g(t)\,dt$ on $\mathsf{P}(R)$, where $'$ denotes differentiation.

9. Let β be a basis for a finite-dimensional inner product space.

 (a) Prove that if $\langle x,z\rangle = 0$ for all $z\in\beta$, then $x = 0$.
 (b) Prove that if $\langle x,z\rangle = \langle y,z\rangle$ for all $z\in\beta$, then $x = y$.

$\boxed{10.^\dagger}$ Let V be an inner product space, and suppose that x and y are orthogonal vectors in V. Prove that $\|x+y\|^2 = \|x\|^2 + \|y\|^2$. Deduce the Pythagorean theorem in R^2. Visit `goo.gl/1iTZzC` for a solution.

11. Prove the *parallelogram law* on an inner product space V; that is, show that

$$\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2 \quad \text{for all } x,y\in\mathsf{V}.$$

What does this equation state about parallelograms in R^2?

12.† Let $\{v_1,v_2,\ldots,v_k\}$ be an orthogonal set in V, and let $a_1,a_2,\ldots,a_k$ be scalars. Prove that

$$\left\|\sum_{i=1}^{k} a_i v_i\right\|^2 = \sum_{i=1}^{k} |a_i|^2\|v_i\|^2.$$

13. Suppose that $\langle\cdot,\cdot\rangle_1$ and $\langle\cdot,\cdot\rangle_2$ are two inner products on a vector space V. Prove that $\langle\cdot,\cdot\rangle = \langle\cdot,\cdot\rangle_1 + \langle\cdot,\cdot\rangle_2$ is another inner product on V.

14. Let A and B be $n\times n$ matrices, and let c be a scalar. Prove that $(A+cB)^* = A^* + \bar{c}B^*$.

15. (a) Prove that if V is an inner product space, then $|\langle x,y\rangle| = \|x\|\cdot\|y\|$ if and only if one of the vectors x or y is a multiple of the other. *Hint:* If the identity holds and $y\neq 0$, let

$$a = \frac{\langle x,y\rangle}{\|y\|^2},$$

and let $z = x - ay$. Prove that y and z are orthogonal and

$$|a| = \frac{\|x\|}{\|y\|}.$$

Then apply Exercise 10 to $\|x\|^2 = \|ay+z\|^2$ to obtain $\|z\| = 0$.

(b) Derive a similar result for the equality $\|x + y\| = \|x\| + \|y\|$, and generalize it to the case of n vectors.

16. (a) Show that the vector space H with $\langle \cdot, \cdot \rangle$ defined on page 330 is an inner product space.

(b) Let $\mathsf{V} = \mathsf{C}([0,1])$, and define

$$\langle f, g \rangle = \int_0^{1/2} f(t)g(t)\,dt.$$

Is this an inner product on V?

17. Let T be a linear operator on an inner product space V, and suppose that $\|\mathsf{T}(x)\| = \|x\|$ for all x. Prove that T is one-to-one.

18. Let V be a vector space over F, where $F = R$ or $F = C$, and let W be an inner product space over F with inner product $\langle \cdot, \cdot \rangle$. If $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$ is linear, prove that $\langle x, y \rangle' = \langle \mathsf{T}(x), \mathsf{T}(y) \rangle$ defines an inner product on V if and only if T is one-to-one.

19. Let V be an inner product space. Prove that

(a) $\|x \pm y\|^2 = \|x\|^2 \pm 2\Re\langle x, y \rangle + \|y\|^2$ for all $x, y \in \mathsf{V}$, where $\Re\langle x, y \rangle$ denotes the real part of the complex number $\langle x, y \rangle$.

(b) $|\,\|x\| - \|y\|\,| \leq \|x - y\|$ for all $x, y \in \mathsf{V}$.

20. Let V be an inner product space over F. Prove the *polar identities*: For all $x, y \in \mathsf{V}$,

(a) $\langle x, y \rangle = \frac{1}{4}\|x + y\|^2 - \frac{1}{4}\|x - y\|^2 \quad$ if $F = R$;

(b) $\langle x, y \rangle = \frac{1}{4}\sum_{k=1}^{4} i^k \|x + i^k y\|^2 \quad$ if $F = C$, where $i^2 = -1$.

21. Let A be an $n \times n$ matrix. Define

$$A_1 = \frac{1}{2}(A + A^*) \quad \text{and} \quad A_2 = \frac{1}{2i}(A - A^*).$$

(a) Prove that $A_1^* = A_1$, $A_2^* = A_2$, and $A = A_1 + iA_2$. Would it be reasonable to define A_1 and A_2 to be the real and imaginary parts, respectively, of the matrix A?

(b) Let A be an $n \times n$ matrix. Prove that the representation in (a) is unique. That is, prove that if $A = B_1 + iB_2$, where $B_1^* = B_1$ and $B_2^* = B_2$, then $B_1 = A_1$ and $B_2 = A_2$.

22. Let V be a real or complex vector space (possibly infinite-dimensional), and let β be a basis for V. For $x, y \in \mathsf{V}$ there exist $v_1, v_2, \ldots, v_n \in \beta$ such that

$$x = \sum_{i=1}^{n} a_i v_i \quad \text{and} \quad y = \sum_{i=1}^{n} b_i v_i.$$

Define

$$\langle x, y\rangle = \sum_{i=1}^{n} a_i \bar{b}_i.$$

(a) Prove that $\langle\cdot,\cdot\rangle$ is an inner product on V and that β is an orthonormal basis for V. Thus every real or complex vector space may be regarded as an inner product space.

(b) Prove that if $\mathsf{V} = \mathsf{R}^n$ or $\mathsf{V} = \mathsf{C}^n$ and β is the standard ordered basis, then the inner product defined above is the standard inner product.

23. Let $\mathsf{V} = \mathsf{F}^n$, and let $A \in \mathsf{M}_{n\times n}(F)$.

(a) Prove that $\langle x, Ay\rangle = \langle A^*x, y\rangle$ for all $x, y \in \mathsf{V}$.

(b) Suppose that for some $B \in \mathsf{M}_{n\times n}(F)$, we have $\langle x, Ay\rangle = \langle Bx, y\rangle$ for all $x, y \in \mathsf{V}$. Prove that $B = A^*$.

(c) Let α be the standard ordered basis for V. For any orthonormal basis β for V, let Q be the $n \times n$ matrix whose columns are the vectors in β. Prove that $Q^* = Q^{-1}$.

(d) Define linear operators T and U on V by $\mathsf{T}(x) = Ax$ and $\mathsf{U}(x) = A^*x$. Show that $[\mathsf{U}]_\beta = [\mathsf{T}]_\beta^*$ for any orthonormal basis β for V.

24. Let V be a complex inner product space with an inner product $\langle\cdot,\cdot\rangle$. Let $[\cdot,\cdot]$ be the real-valued function such that $[x, y]$ is the real part of the complex number $\langle x, y\rangle$ for all $x, y \in \mathsf{V}$. Prove that $[\cdot,\cdot]$ is an inner product for V, where V is regarded as a vector space over R. Prove, furthermore, that $[x, ix] = 0$ for all $x \in \mathsf{V}$.

25. Let V be a vector space over C, and suppose that $[\cdot,\cdot]$ is a real inner product on V, where V is regarded as a vector space over R, such that $[x, ix] = 0$ for all $x \in \mathsf{V}$. Let $\langle\cdot,\cdot\rangle$ be the complex-valued function defined by

$$\langle x, y\rangle = [x, y] + i[x, iy] \qquad \text{for } x, y \in \mathsf{V}.$$

Prove that $\langle\cdot,\cdot\rangle$ is a complex inner product on V.

The following definition is used in Exercises 26–30.

Definition. *Let* V *be a vector space over* F*, where* F *is either* R *or* C*. Regardless of whether* V *is or is not an inner product space, we may still define a norm* $\|\cdot\|_{\mathsf{V}}$ *as a real-valued function on* V *satisfying the following three conditions for all* $x, y \in \mathsf{V}$ *and* $a \in F$:

(1) $\|x\|_{\mathsf{V}} \geq 0$, *and* $\|x\|_{\mathsf{V}} = 0$ *if and only if* $x = 0$.

(2) $\|ax\|_{\mathsf{V}} = |a| \cdot \|x\|_{\mathsf{V}}$.

(3) $\|x + y\|_{\mathsf{V}} \leq \|x\|_{\mathsf{V}} + \|y\|_{\mathsf{V}}$.

26. Prove that the following are norms on the given vector spaces V.

(a) $V = R^2$; $\quad \|(a,b)\|_V = |a| + |b| \quad$ for all $(a,b) \in V$

(b) $V = C([0,1])$; $\quad \|f\|_V = \max\limits_{t\in[0,1]} |f(t)| \quad$ for all $f \in V$

(c) $V = C([0,1])$; $\quad \|f\|_V = \int_0^1 |f(t)|\,dt \quad$ for all $f \in V$

(d) $V = M_{m\times n}(F)$; $\quad \|A\|_V = \max\limits_{i,j} |A_{ij}| \quad$ for all $A \in V$

27. Use Exercise 11 to show that there is no inner product $\langle\cdot,\cdot\rangle$ on R^2 such that $\|x\|_V^2 = \langle x, x\rangle$ for all $x \in R^2$ if the norm is defined as in Exercise 26(a).

28. Let $\|\cdot\|_V$ be a norm on a vector space V, and define, for each ordered pair of vectors, the scalar $d(x,y) = \|x-y\|_V$, called the **distance** between x and y. Prove the following results for all $x,y,z \in V$.

(a) $d(x,y) \geq 0$.
(b) $d(x,y) = d(y,x)$.
(c) $d(x,y) \leq d(x,z) + d(z,y)$.
(d) $d(x,x) = 0$ if and only if $x = 0$.
(e) $d(x,y) \neq 0$ if $x \neq y$.

29. Let $\|\cdot\|_V$ be a norm on a real vector space V satisfying the parallelogram law given in Exercise 11. Define

$$\langle x,y\rangle = \frac{1}{4}\left[\|x+y\|_V^2 - \|x-y\|_V^2\right].$$

Prove that $\langle\cdot,\cdot\rangle$ defines an inner product on V such that $\|x\|_V^2 = \langle x,x\rangle$ for all $x \in V$. *Hints:*

(a) Prove $\langle x, 2y\rangle = 2\langle x,y\rangle$ for all $x,y \in V$.
(b) Prove $\langle x+u, y\rangle = \langle x,y\rangle + \langle u,y\rangle$ for all $x,u,y \in V$.
(c) Prove $\langle nx, y\rangle = n\langle x,y\rangle$ for every positive integer n and every $x,y \in V$.
(d) Prove $m\left\langle \frac{1}{m}x, y\right\rangle = \langle x,y\rangle$ for every positive integer m and every $x,y \in V$.
(e) Prove $\langle rx, y\rangle = r\langle x,y\rangle$ for every rational number r and every $x,y \in V$.
(f) Prove $|\langle x,y\rangle| \leq \|x\|_V\|y\|_V$ for every $x,y \in V$. *Hint:* Condition (3) in the definition of norm can be helpful.
(g) Prove that for every $c \in R$, every rational number r, and every $x,y \in V$,

$$\begin{aligned}|c\langle x,y\rangle - \langle cx,y\rangle| &= |(c-r)\langle x,y\rangle - \langle (c-r)x,y\rangle| \\ &\leq 2|c-r|\|x\|_V\|y\|_V.\end{aligned}$$

(h) Use the fact that for any $c \in R$, $|c - r|$ can be made arbitrarily small, where r varies over the set of rational numbers, to establish item (b) of the definition of *inner product*.

30. Let $\|\cdot\|_{\mathsf{V}}$ be a norm (as defined on page 337) on a complex vector space V satisfying the parallelogram law given in Exercise 11. Prove that there is an inner product $\langle\cdot,\cdot\rangle$ on V such that $\|x\|_{\mathsf{V}}^2 = \langle x, x\rangle$ for all $x \in \mathsf{V}$. *Hint:* Apply Exercise 29 to V regarded as a vector space over R. Then apply Exercise 25.

6.2 THE GRAM–SCHMIDT ORTHOGONALIZATION PROCESS AND ORTHOGONAL COMPLEMENTS

In previous chapters, we have seen the special role of the standard ordered bases for C^n and R^n. The special properties of these bases stem from the fact that the basis vectors form an orthonormal set. Just as bases are the building blocks of vector spaces, bases that are also orthonormal sets are the building blocks of inner product spaces. We now name such bases.

Definition. *Let* V *be an inner product space. A subset of* V *is an* ***orthonormal basis*** *for* V *if it is an ordered basis that is orthonormal.*

Example 1

The standard ordered basis for F^n is an orthonormal basis for F^n. ◆

Example 2

The set

$$\left\{\left(\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}\right), \left(\frac{2}{\sqrt{5}}, \frac{-1}{\sqrt{5}}\right)\right\}$$

is an orthonormal basis for R^2. ◆

The next theorem and its corollaries illustrate why orthonormal sets and, in particular, orthonormal bases are so important.

Theorem 6.3. *Let* V *be an inner product space and* $S = \{v_1, v_2, \ldots, v_k\}$ *be an orthogonal subset of* V *consisting of nonzero vectors. If* $y \in \operatorname{span}(S)$, *then*

$$y = \sum_{i=1}^{k} \frac{\langle y, v_i\rangle}{\|v_i\|^2} v_i.$$

Proof. Write $y = \displaystyle\sum_{i=1}^{k} a_i v_i$, where $a_1, a_2, \ldots, a_k \in F$. Then, for $1 \le j \le k$, we have

$$\langle y, v_j\rangle = \left\langle \sum_{i=1}^{k} a_i v_i, v_j \right\rangle = \sum_{i=1}^{k} a_i \langle v_i, v_j\rangle = a_j \langle v_j, v_j\rangle = a_j \|v_j\|^2.$$

If $x \in \mathsf{W}^\perp$, then $\langle x, v_i \rangle = 0$ for $1 \le i \le k$. Therefore

$$x = \sum_{i=k+1}^{n} \langle x, v_i \rangle\, v_i \in \operatorname{span}(S_1).$$

(c) Let W be a subspace of V. It is a finite-dimensional inner product space because V is, and so it has an orthonormal basis $\{v_1, v_2, \ldots, v_k\}$. By (a) and (b), we have

$$\dim(\mathsf{V}) = n = k + (n - k) = \dim(\mathsf{W}) + \dim(\mathsf{W}^\perp). \qquad \blacksquare$$

Example 11

Let $\mathsf{W} = \operatorname{span}(\{e_1, e_2\})$ in F^3. Then $x = (a, b, c) \in \mathsf{W}^\perp$ if and only if $0 = \langle x, e_1 \rangle = a$ and $0 = \langle x, e_2 \rangle = b$. So $x = (0, 0, c)$, and therefore $\mathsf{W}^\perp = \operatorname{span}(\{e_3\})$. One can deduce the same result by noting that $e_3 \in \mathsf{W}^\perp$ and, from (c), that $\dim(\mathsf{W}^\perp) = 3 - 2 = 1$. ♦

EXERCISES

1. Label the following statements as true or false.
   (a) The Gram–Schmidt orthogonalization process produces an orthonormal set from an arbitrary linearly independent set.
   (b) Every nonzero finite-dimensional inner product space has an orthonormal basis.
   (c) The orthogonal complement of any set is a subspace.
   (d) If $\{v_1, v_2, \ldots, v_n\}$ is a basis for an inner product space V, then for any $x \in \mathsf{V}$ the scalars $\langle x, v_i \rangle$ are the Fourier coefficients of x.
   (e) An orthonormal basis must be an ordered basis.
   (f) Every orthogonal set is linearly independent.
   (g) Every orthonormal set is linearly independent.

2. In each part, apply the Gram–Schmidt process to the given subset S of the inner product space V to obtain an orthogonal basis for $\operatorname{span}(S)$. Then normalize the vectors in this basis to obtain an orthonormal basis β for $\operatorname{span}(S)$, and compute the Fourier coefficients of the given vector relative to β. Finally, use Theorem 6.5 to verify your result.
   (a) $\mathsf{V} = \mathsf{R}^3$, $S = \{(1,0,1), (0,1,1), (1,3,3)\}$, and $x = (1,1,2)$
   (b) $\mathsf{V} = \mathsf{R}^3$, $S = \{(1,1,1), (0,1,1), (0,0,1)\}$, and $x = (1,0,1)$
   (c) $\mathsf{V} = \mathsf{P}_2(R)$ with the inner product $\langle f(x), g(x) \rangle = \int_0^1 f(t)g(t)\,dt$, $S = \{1, x, x^2\}$, and $h(x) = 1 + x$
   (d) $\mathsf{V} = \operatorname{span}(S)$, where $S = \{(1, i, 0), (1 - i, 2, 4i)\}$, and $x = (3 + i, 4i, -4)$

(e) $V = R^4$, $S = \{(2,-1,-2,4),(-2,1,-5,5),(-1,3,7,11)\}$, and $x = (-11,8,-4,18)$

(f) $V = R^4$, $S = \{(1,-2,-1,3),(3,6,3,-1),(1,4,2,8)\}$, and $x = (-1,2,1,1)$

(g) $V = M_{2\times 2}(R)$, $S = \left\{\begin{pmatrix} 3 & 5 \\ -1 & 1 \end{pmatrix}, \begin{pmatrix} -1 & 9 \\ 5 & -1 \end{pmatrix}, \begin{pmatrix} 7 & -17 \\ 2 & -6 \end{pmatrix}\right\}$, and $A = \begin{pmatrix} -1 & 27 \\ -4 & 8 \end{pmatrix}$.

(h) $V = M_{2\times 2}(R)$, $S = \left\{\begin{pmatrix} 2 & 2 \\ 2 & 1 \end{pmatrix}, \begin{pmatrix} 11 & 4 \\ 2 & 5 \end{pmatrix}, \begin{pmatrix} 4 & -12 \\ 3 & -16 \end{pmatrix}\right\}$, and $A = \begin{pmatrix} 8 & 6 \\ 25 & -13 \end{pmatrix}$

(i) $V = \operatorname{span}(S)$ with the inner product $\langle f,g\rangle = \int_0^\pi f(t)g(t)\,dt$, $S = \{\sin t, \cos t, 1, t\}$, and $h(t) = 2t+1$

(j) $V = C^4$, $S = \{(1,i,2-i,-1),(2+3i,3i,1-i,2i), (-1+7i,6+10i,11-4i,3+4i)\}$, and $x = (-2+7i,6+9i,9-3i,4+4i)$

(k) $V = C^4$, $S = \{(-4,3-2i,i,1-4i), (-1-5i,5-4i,-3+5i,7-2i),(-27-i,-7-6i,-15+25i,-7-6i)\}$, and $x = (-13-7i,-12+3i,-39-11i,-26+5i)$

(l) $V = M_{2\times 2}(C)$, $S = \left\{\begin{pmatrix} 1-i & -2-3i \\ 2+2i & 4+i \end{pmatrix}, \begin{pmatrix} 8i & 4 \\ -3-3i & -4+4i \end{pmatrix}, \begin{pmatrix} -25-38i & -2-13i \\ 12-78i & -7+24i \end{pmatrix}\right\}$, and $A = \begin{pmatrix} -2+8i & -13+i \\ 10-10i & 9-9i \end{pmatrix}$

(m) $V = M_{2\times 2}(C)$, $S = \left\{\begin{pmatrix} -1+i & -i \\ 2-i & 1+3i \end{pmatrix}, \begin{pmatrix} -1-7i & -9-8i \\ 1+10i & -6-2i \end{pmatrix}, \begin{pmatrix} -11-132i & -34-31i \\ 7-126i & -71-5i \end{pmatrix}\right\}$, and $A = \begin{pmatrix} -7+5i & 3+18i \\ 9-6i & -3+7i \end{pmatrix}$

3. In R^2, let

$$\beta = \left\{\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right), \left(\frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}\right)\right\}.$$

Find the Fourier coefficients of $(3,4)$ relative to β.

4. Let $S = \{(1,0,i),(1,2,1)\}$ in C^3. Compute $S^\perp$.

5. Let $S_0 = \{x_0\}$, where x_0 is a nonzero vector in R^3. Describe $S_0^\perp$ geometrically. Now suppose that $S = \{x_1, x_2\}$ is a linearly independent subset of R^3. Describe $S^\perp$ geometrically.

6. Let V be an inner product space, and let W be a finite-dimensional subspace of V. If $x \notin \mathsf{W}$, prove that there exists $y \in \mathsf{V}$ such that $y \in \mathsf{W}^{\perp}$, but $\langle x, y\rangle \neq 0$. *Hint:* Use Theorem 6.6.

7. Let β be a basis for a subspace W of an inner product space V, and let $z \in \mathsf{V}$. Prove that $z \in \mathsf{W}^{\perp}$ if and only if $\langle z, v\rangle = 0$ for every $v \in \beta$.

8. Prove that if $\{w_1, w_2, \ldots, w_n\}$ is an orthogonal set of nonzero vectors, then the vectors $v_1, v_2, \ldots, v_n$ derived from the Gram–Schmidt process satisfy $v_i = w_i$ for $i = 1, 2, \ldots, n$. *Hint:* Use mathematical induction.

9. Let $\mathsf{W} = \operatorname{span}(\{(i, 0, 1)\})$ in C^3. Find orthonormal bases for W and $\mathsf{W}^{\perp}$.

10. Let W be a finite-dimensional subspace of an inner product space V. Prove that $\mathsf{V} = \mathsf{W} \oplus \mathsf{W}^{\perp}$. Using the definition on page 76, prove that there exists a projection T on W along $\mathsf{W}^{\perp}$ that satisfies $\mathsf{N}(\mathsf{T}) = \mathsf{W}^{\perp}$. In addition, prove that $\|\mathsf{T}(x)\| \leq \|x\|$ for all $x \in \mathsf{V}$. *Hint:* Use Theorem 6.6 and Exercise 10 of Section 6.1.

$\boxed{11.}$ Let A be an $n \times n$ matrix with complex entries. Prove that $AA^* = I$ if and only if the rows of A form an orthonormal basis for C^n. Visit `goo.gl/iKcC4S` for a solution.

12. Prove that for any matrix $A \in \mathsf{M}_{m\times n}(F)$, $(\mathsf{R}(\mathsf{L}_{A^*}))^{\perp} = \mathsf{N}(\mathsf{L}_A)$.

13. Let V be an inner product space, S and S_0 be subsets of V, and W be a finite-dimensional subspace of V. Prove the following results.
   (a) $S_0 \subseteq S$ implies that $S^{\perp} \subseteq S_0^{\perp}$.
   (b) $S \subseteq (S^{\perp})^{\perp}$; so $\operatorname{span}(S) \subseteq (S^{\perp})^{\perp}$.
   (c) $\mathsf{W} = (\mathsf{W}^{\perp})^{\perp}$. *Hint:* Use Exercise 6.
   (d) $\mathsf{V} = \mathsf{W} \oplus \mathsf{W}^{\perp}$. (See the exercises of Section 1.3.)

14. Let W_1 and W_2 be subspaces of a finite-dimensional inner product space. Prove that $(\mathsf{W}_1+\mathsf{W}_2)^{\perp} = \mathsf{W}_1^{\perp}\cap\mathsf{W}_2^{\perp}$ and $(\mathsf{W}_1\cap\mathsf{W}_2)^{\perp} = \mathsf{W}_1^{\perp}+\mathsf{W}_2^{\perp}$. (See the definition of the sum of subsets of a vector space on page 22.) *Hint for the second equation:* Apply Exercise 13(c) to the first equation.

15. Let V be a finite-dimensional inner product space over F.
   (a) *Parseval's Identity.* Let $\{v_1, v_2, \ldots, v_n\}$ be an orthonormal basis for V. For any $x, y \in \mathsf{V}$ prove that

$$\langle x, y\rangle = \sum_{i=1}^{n} \langle x, v_i\rangle \overline{\langle y, v_i\rangle}.$$

   (b) Use (a) to prove that if β is an orthonormal basis for V with inner product $\langle \cdot, \cdot\rangle$, then for any $x, y \in \mathsf{V}$

$$\langle \phi_\beta(x), \phi_\beta(y)\rangle' = \langle [x]_\beta, [y]_\beta\rangle' = \langle x, y\rangle,$$

where $\langle \cdot, \cdot \rangle'$ is the standard inner product on F^n.

16. (a) *Bessel's Inequality.* Let V be an inner product space, and let $S = \{v_1, v_2, \ldots, v_n\}$ be an orthonormal subset of V. Prove that for any $x \in \mathsf{V}$ we have

$$\|x\|^2 \geq \sum_{i=1}^{n} |\langle x, v_i \rangle|^2.$$

Hint: Apply Theorem 6.6 to $x \in \mathsf{V}$ and $\mathsf{W} = \text{span}(S)$. Then use Exercise 10 of Section 6.1.

(b) In the context of (a), prove that Bessel's inequality is an equality if and only if $x \in \text{span}(S)$.

17. Let T be a linear operator on an inner product space V. If $\langle \mathsf{T}(x), y \rangle = 0$ for all $x, y \in \mathsf{V}$, prove that $\mathsf{T} = \mathsf{T}_0$. In fact, prove this result if the equality holds for all x and y in some basis for V.

18. Let $\mathsf{V} = \mathsf{C}([-1,1])$. Suppose that W_e and W_o denote the subspaces of V consisting of the even and odd functions, respectively. (See Exercise 22 of Section 1.3.) Prove that $\mathsf{W}_e^{\perp} = \mathsf{W}_o$, where the inner product on V is defined by

$$\langle f, g \rangle = \int_{-1}^{1} f(t)g(t)\, dt.$$

19. In each of the following parts, find the orthogonal projection of the given vector on the given subspace W of the inner product space V.

(a) $\mathsf{V} = \mathsf{R}^2$, $u = (2,6)$, and $\mathsf{W} = \{(x,y)\colon y = 4x\}$.
(b) $\mathsf{V} = \mathsf{R}^3$, $u = (2,1,3)$, and $\mathsf{W} = \{(x,y,z)\colon x + 3y - 2z = 0\}$.
(c) $\mathsf{V} = \mathsf{P}(R)$ with the inner product $\langle f(x), g(x) \rangle = \int_0^1 f(t)g(t)\, dt$, $h(x) = 4 + 3x - 2x^2$, and $\mathsf{W} = \mathsf{P}_1(R)$.

20. In each part of Exercise 19, find the distance from the given vector to the subspace W.

21. Let $\mathsf{V} = \mathsf{C}([-1,1])$ with the inner product $\langle f, g \rangle = \int_{-1}^{1} f(t)g(t)\, dt$, and let W be the subspace $\mathsf{P}_2(R)$, viewed as a space of functions. Use the orthonormal basis obtained in Example 5 to compute the "best" (closest) second-degree polynomial approximation of the function $h(t) = e^t$ on the interval $[-1,1]$.

22. Let $\mathsf{V} = \mathsf{C}([0,1])$ with the inner product $\langle f, g \rangle = \int_0^1 f(t)g(t)\, dt$. Let W be the subspace spanned by the linearly independent set $\{t, \sqrt{t}\}$.

(a) Find an orthonormal basis for W.
(b) Let $h(t) = t^2$. Use the orthonormal basis obtained in (a) to obtain the "best" (closest) approximation of h in W.

23. Let V be the vector space defined in Example 5 of Section 1.2, the space of all sequences σ in F (where $F = R$ or $F = C$) such that $\sigma(n) \neq 0$ for only finitely many positive integers n. For $\sigma, \mu \in \mathsf{V}$, we define $\langle \sigma, \mu \rangle = \sum_{n=1}^{\infty} \sigma(n)\overline{\mu(n)}$. Since all but a finite number of terms of the series are zero, the series converges.

(a) Prove that $\langle \cdot, \cdot \rangle$ is an inner product on V, and hence V is an inner product space.

(b) For each positive integer n, let e_n be the sequence defined by $e_n(k) = \delta_{nk}$, where δ_{nk} is the Kronecker delta. Prove that $\{e_1, e_2, \ldots\}$ is an orthonormal basis for V.

(c) Let $\sigma_n = e_1 + e_n$ and $\mathsf{W} = \operatorname{span}(\{\sigma_n : n \geq 2\}$.

(i) Prove that $e_1 \notin \mathsf{W}$, so $\mathsf{W} \neq \mathsf{V}$.

(ii) Prove that $\mathsf{W}^{\perp} = \{0\}$, and conclude that $\mathsf{W} \neq (\mathsf{W}^{\perp})^{\perp}$.

Thus the assumption in Exercise 13(c) that W is finite-dimensional is essential.

6.3 THE ADJOINT OF A LINEAR OPERATOR

In Section 6.1, we defined the conjugate transpose A^* of a matrix A. For a linear operator T on an inner product space V, we now define a related linear operator on V called the *adjoint* of T, whose matrix representation with respect to any orthonormal basis β for V is $[\mathsf{T}]_\beta^*$. The analogy between conjugation of complex numbers and adjoints of linear operators will become apparent. We first need a preliminary result.

Let V be an inner product space, and let $y \in \mathsf{V}$. The function $\mathsf{g}\colon \mathsf{V} \to F$ defined by $\mathsf{g}(x) = \langle x, y \rangle$ is clearly linear. More interesting is the fact that if V is finite-dimensional, every linear transformation from V into F is of this form.

Theorem 6.8. *Let* V *be a finite-dimensional inner product space over* F, *and let* $\mathsf{g}\colon \mathsf{V} \to F$ *be a linear transformation. Then there exists a unique vector* $y \in \mathsf{V}$ *such that* $\mathsf{g}(x) = \langle x, y \rangle$ *for all* $x \in \mathsf{V}$.

Proof. Let $\beta = \{v_1, v_2, \ldots, v_n\}$ be an orthonormal basis for V, and let

$$y = \sum_{i=1}^{n} \overline{\mathsf{g}(v_i)} v_i.$$

Define $\mathsf{h}\colon \mathsf{V} \to F$ by $\mathsf{h}(x) = \langle x, y \rangle$, which is clearly linear. Furthermore, for $1 \leq j \leq n$ we have

$$\mathsf{h}(v_j) = \langle v_j, y \rangle = \left\langle v_j, \sum_{i=1}^{n} \overline{\mathsf{g}(v_i)} v_i \right\rangle$$

Example 3

Consider the system

$$\begin{aligned} x + 2y + \ \ z &= \ \ 4 \\ x - \ \ y + 2z &= -11 \\ x + 5y \qquad &= \ \ 19. \end{aligned}$$

Let

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 1 & -1 & 2 \\ 1 & 5 & 0 \end{pmatrix} \quad \text{and} \quad b = \begin{pmatrix} 4 \\ -11 \\ 19 \end{pmatrix}.$$

To find the minimal solution to this system, we must first find some solution u to $AA^*x = b$. Now

$$AA^* = \begin{pmatrix} 6 & 1 & 11 \\ 1 & 6 & -4 \\ 11 & -4 & 26 \end{pmatrix};$$

so we consider the system

$$\begin{aligned} 6x + \ \ y + 11z &= \ \ 4 \\ x + 6y - \ \ 4z &= -11 \\ 11x - 4y + 26z &= \ \ 19, \end{aligned}$$

for which one solution is

$$u = \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix}.$$

(Any solution will suffice.) Hence

$$s = A^*u = \begin{pmatrix} -1 \\ 4 \\ -3 \end{pmatrix}$$

is the minimal solution to the given system. ◆

EXERCISES

1. Label the following statements as true or false. Assume that the underlying inner product spaces are finite-dimensional.

   (a) Every linear operator has an adjoint.
   (b) Every linear operator on V has the form $x \to \langle x, y \rangle$ for some $y \in \mathsf{V}$.
   (c) For every linear operator T on V and every ordered basis β for V, we have $[\mathsf{T}^*]_\beta = ([\mathsf{T}]_\beta)^*$.
   (d) The adjoint of a linear operator is unique.

(e) For any linear operators T and U and scalars a and b,

$$(a\mathsf{T} + b\mathsf{U})^* = a\mathsf{T}^* + b\mathsf{U}^*.$$

(f) For any $n \times n$ matrix A, we have $(\mathsf{L}_A)^* = \mathsf{L}_{A^*}$.
(g) For any linear operator T, we have $(\mathsf{T}^*)^* = \mathsf{T}$.

2. For each of the following inner product spaces V (over F) and linear transformations $\mathsf{g}\colon \mathsf{V} \to F$, find a vector y such that $\mathsf{g}(x) = \langle x, y\rangle$ for all $x \in \mathsf{V}$.
   (a) $\mathsf{V} = \mathsf{R}^3$, $\mathsf{g}(a_1, a_2, a_3) = a_1 - 2a_2 + 4a_3$
   (b) $\mathsf{V} = \mathsf{C}^2$, $\mathsf{g}(z_1, z_2) = z_1 - 2z_2$
   (c) $\mathsf{V} = \mathsf{P}_2(R)$ with $\langle f(x), h(x)\rangle = \int_0^1 f(t)h(t)\,dt$, $\mathsf{g}(f) = f(0) + f'(1)$

3. For each of the following inner product spaces V and linear operators T on V, evaluate T^* at the given vector in V.
   (a) $\mathsf{V} = \mathsf{R}^2$, $\mathsf{T}(a, b) = (2a + b, a - 3b)$, $x = (3, 5)$.
   (b) $\mathsf{V} = \mathsf{C}^2$, $\mathsf{T}(z_1, z_2) = (2z_1 + iz_2, (1 - i)z_1)$, $x = (3 - i, 1 + 2i)$.
   (c) $\mathsf{V} = \mathsf{P}_1(R)$ with $\langle f(x), g(x)\rangle = \int_{-1}^1 f(t)g(t)\,dt$, $\mathsf{T}(f) = f' + 3f$, $f(t) = 4 - 2t$

4. Complete the proof of Theorem 6.11.

5. (a) Complete the proof of the corollary to Theorem 6.11 by using Theorem 6.11, as in the proof of (c).
   (b) State a result for nonsquare matrices that is analogous to the corollary to Theorem 6.11, and prove it using a matrix argument.

6. Let T be a linear operator on an inner product space V. Let $\mathsf{U}_1 = \mathsf{T} + \mathsf{T}^*$ and $\mathsf{U}_2 = \mathsf{T}\mathsf{T}^*$. Prove that $\mathsf{U}_1 = \mathsf{U}_1^*$ and $\mathsf{U}_2 = \mathsf{U}_2^*$.

7. Give an example of a linear operator T on an inner product space V such that $\mathsf{N}(\mathsf{T}) \neq \mathsf{N}(\mathsf{T}^*)$.

8. Let V be a finite-dimensional inner product space, and let T be a linear operator on V. Prove that if T is invertible, then T^* is invertible and $(\mathsf{T}^*)^{-1} = (\mathsf{T}^{-1})^*$.

9. Prove that if $\mathsf{V} = \mathsf{W} \oplus \mathsf{W}^\perp$ and T is the projection on W along $\mathsf{W}^\perp$, then $\mathsf{T} = \mathsf{T}^*$. *Hint:* Recall that $\mathsf{N}(\mathsf{T}) = \mathsf{W}^\perp$. (For definitions, see the exercises of Sections 1.3 and 2.1.)

10. Let T be a linear operator on an inner product space V. Prove that $\|\mathsf{T}(x)\| = \|x\|$ for all $x \in \mathsf{V}$ if and only if $\langle \mathsf{T}(x), \mathsf{T}(y)\rangle = \langle x, y\rangle$ for all $x, y \in \mathsf{V}$. *Hint:* Use Exercise 20 of Section 6.1.

11. For a linear operator T on an inner product space V, prove that $\mathsf{T}^*\mathsf{T} = \mathsf{T}_0$ implies $\mathsf{T} = \mathsf{T}_0$. Is the same result true if we assume that $\mathsf{T}\mathsf{T}^* = \mathsf{T}_0$?

12. Let V be an inner product space, and let T be a linear operator on V. Prove the following results.

(a) $\mathsf{R}(\mathsf{T}^*)^\perp = \mathsf{N}(\mathsf{T})$.
(b) If V is finite-dimensional, then $\mathsf{R}(\mathsf{T}^*) = \mathsf{N}(\mathsf{T})^\perp$. *Hint:* Use Exercise 13(c) of Section 6.2.

13. Let T be a linear operator on a finite-dimensional inner product space V. Prove the following results.

(a) $\mathsf{N}(\mathsf{T}^*\mathsf{T}) = \mathsf{N}(\mathsf{T})$. Deduce that $\operatorname{rank}(\mathsf{T}^*\mathsf{T}) = \operatorname{rank}(\mathsf{T})$.
(b) $\operatorname{rank}(\mathsf{T}) = \operatorname{rank}(\mathsf{T}^*)$. Deduce from (a) that $\operatorname{rank}(\mathsf{T}\mathsf{T}^*) = \operatorname{rank}(\mathsf{T})$.
(c) For any $n \times n$ matrix A, $\operatorname{rank}(A^*A) = \operatorname{rank}(AA^*) = \operatorname{rank}(A)$.

14. Let V be an inner product space, and let $y, z \in \mathsf{V}$. Define $\mathsf{T}\colon \mathsf{V} \to \mathsf{V}$ by $\mathsf{T}(x) = \langle x, y\rangle z$ for all $x \in \mathsf{V}$. First prove that T is linear. Then show that T^* exists, and find an explicit expression for it.

The following definition is used in Exercises 15–17 and is an extension of the definition of the *adjoint* of a linear operator.

Definition. *Let* $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$ *be a linear transformation, where* V *and* W *are finite-dimensional inner product spaces with inner products* $\langle\cdot,\cdot\rangle_1$ *and* $\langle\cdot,\cdot\rangle_2$*, respectively. A function* $\mathsf{T}^*\colon \mathsf{W} \to \mathsf{V}$ *is called an* **adjoint** *of* T *if* $\langle \mathsf{T}(x), y\rangle_2 = \langle x, \mathsf{T}^*(y)\rangle_1$ *for all* $x \in \mathsf{V}$ *and* $y \in \mathsf{W}$.

15. Let $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$ be a linear transformation, where V and W are finite-dimensional inner product spaces with inner products $\langle\cdot,\cdot\rangle_1$ and $\langle\cdot,\cdot\rangle_2$, respectively. Prove the following results.

(a) There is a unique adjoint T^* of T, and T^* is linear.
(b) If β and γ are orthonormal bases for V and W, respectively, then $[\mathsf{T}^*]_\gamma^\beta = ([\mathsf{T}]_\beta^\gamma)^*$.
(c) $\operatorname{rank}(\mathsf{T}^*) = \operatorname{rank}(\mathsf{T})$.
(d) $\langle \mathsf{T}^*(x), y\rangle_1 = \langle x, \mathsf{T}(y)\rangle_2$ for all $x \in \mathsf{W}$ and $y \in \mathsf{V}$.
(e) For all $x \in \mathsf{V}$, $\mathsf{T}^*\mathsf{T}(x) = \mathit{0}$ if and only if $\mathsf{T}(x) = \mathit{0}$.

16. State and prove a result that extends the first four parts of Theorem 6.11 using the preceding definition.

17. Let $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$ be a linear transformation, where V and W are finite-dimensional inner product spaces. Prove that $(\mathsf{R}(\mathsf{T}^*))^\perp = \mathsf{N}(\mathsf{T})$, using the preceding definition.

18.† Let A be an $n \times n$ matrix. Prove that $\det(A^*) = \overline{\det(A)}$. Visit `goo.gl/csqoFY` for a solution.

19. Suppose that A is an $m \times n$ matrix in which no two columns are identical. Prove that A^*A is a diagonal matrix if and only if every pair of columns of A is orthogonal.

20. For each of the sets of data that follows, use the least squares approximation to find the best fits with both (i) a linear function and (ii) a quadratic function. Compute the error E in both cases.

(a) $\{(-3,9),(-2,6),(0,2),(1,1)\}$
(b) $\{(1,2),(3,4),(5,7),(7,9),(9,12)\}$
(c) $\{(-2,4),(-1,3),(0,1),(1,-1),(2,-3)\}$

21. In physics, *Hooke's law* states that (within certain limits) there is a linear relationship between the length x of a spring and the force y applied to (or exerted by) the spring. That is, $y = cx + d$, where c is called the **spring constant**. Use the following data to estimate the spring constant (the length is given in inches and the force is given in pounds).

Length x	Force y
3.5	1.0
4.0	2.2
4.5	2.8
5.0	4.3

22. Find the minimal solution to each of the following systems of linear equations.

(a) $x + 2y - z = 12$

(b) $$\begin{aligned} x + 2y - z &= 1 \\ 2x + 3y + z &= 2 \\ 4x + 7y - z &= 4 \end{aligned}$$

(c) $$\begin{aligned} x + y - z &= 0 \\ 2x - y + z &= 3 \\ x - y + z &= 2 \end{aligned}$$

(d) $$\begin{aligned} x + y + z - w &= 1 \\ 2x - y \qquad + w &= 1 \end{aligned}$$

23. Consider the problem of finding the least squares line $y = ct + d$ corresponding to the m observations $(t_1, y_1), (t_2, y_2), \ldots, (t_m, y_m)$.

(a) Show that the equation $(A^*A)x_0 = A^*y$ of Theorem 6.12 takes the form of the *normal equations:*

$$\left(\sum_{i=1}^{m} t_i^2\right) c + \left(\sum_{i=1}^{m} t_i\right) d = \sum_{i=1}^{m} t_i y_i$$

and

$$\left(\sum_{i=1}^{m} t_i\right) c + md = \sum_{i=1}^{m} y_i.$$

These equations may also be obtained from the error E by setting the partial derivatives of E with respect to both c and d equal to zero.

(b) Use the second normal equation of (a) to show that the least squares line must pass through the *center of mass*, $(\bar{t}, \bar{y})$, where

$$\bar{t} = \frac{1}{m}\sum_{i=1}^{m} t_i \quad \text{and} \quad \bar{y} = \frac{1}{m}\sum_{i=1}^{m} y_i.$$

24. Let V and $\{e_1, e_2, \ldots\}$ be defined as in Exercise 23 of Section 6.2. Define $\mathsf{T}\colon \mathsf{V} \to \mathsf{V}$ by

$$\mathsf{T}(\sigma)(k) = \sum_{i=k}^{\infty} \sigma(i) \quad \text{for every positive integer } k.$$

Notice that the infinite series in the definition of T converges because $\sigma(i) \neq 0$ for only finitely many i.

(a) Prove that T is a linear operator on V.
(b) Prove that for any positive integer n, $\mathsf{T}(e_n) = \sum_{i=1}^{n} e_i$.
(c) Prove that T has no adjoint. *Hint:* By way of contradiction, suppose that T^* exists. Prove that for any positive integer n, $\mathsf{T}^*(e_n)(k) \neq 0$ for infinitely many k.

6.4 NORMAL AND SELF-ADJOINT OPERATORS

We have seen the importance of diagonalizable operators in Chapter 5. For an operator on a vector space V to be diagonalizble, it is necessary and sufficient for V to contain a basis of eigenvectors for this operator. As V is an inner product space in this chapter, it is reasonable to seek conditions that guarantee that V has an orthonormal basis of eigenvectors. A very important result that helps achieve our goal is Schur's theorem (Theorem 6.14). The formulation that follows is in terms of linear operators. The next section contains the more familiar matrix form. We begin with a lemma.

Lemma. *Let* T *be a linear operator on a finite-dimensional inner product space* V. *If* T *has an eigenvector, then so does* T^*.

Proof. Suppose that v is an eigenvector of T with corresponding eigenvalue λ. Then for any $x \in \mathsf{V}$,

$$0 = \langle \mathit{0}, x\rangle = \langle (\mathsf{T} - \lambda\mathsf{I})(v), x\rangle = \langle v, (\mathsf{T} - \lambda\mathsf{I})^*(x)\rangle = \left\langle v, (\mathsf{T}^* - \overline{\lambda}\mathsf{I})(x)\right\rangle,$$

λ is real, the characteristic polynomial splits over R. But T_A has the same characteristic polynomial as A, which has the same characteristic polynomial as T. Therefore the characteristic polynomial of T splits. ∎

We are now able to establish one of the major results of this chapter.

Theorem 6.17. *Let T be a linear operator on a finite-dimensional real inner product space V. Then T is self-adjoint if and only if there exists an orthonormal basis β for V consisting of eigenvectors of T.*

Proof. Suppose that T is self-adjoint. By the lemma, we may apply Schur's theorem to obtain an orthonormal basis β for V such that the matrix $A = [\mathsf{T}]_\beta$ is upper triangular. But

$$A^* = [\mathsf{T}]_\beta^* = [\mathsf{T}^*]_\beta = [\mathsf{T}]_\beta = A.$$

So A and A^* are both upper triangular, and therefore A is a diagonal matrix. Thus β must consist of eigenvectors of T.

The converse is left as an exercise. ∎

We restate this theorem in matrix form in the next section (as Theorem 6.20 on p. 381).

Example 4

As we noted earlier, real symmetric matrices are self-adjoint, and self-adjoint matrices are normal. The following matrix A is complex and symmetric:

$$A = \begin{pmatrix} i & i \\ i & 1 \end{pmatrix} \quad \text{and} \quad A^* = \begin{pmatrix} -i & -i \\ -i & 1 \end{pmatrix}.$$

But A is not normal, because $(AA^*)_{12} = 1+i$ and $(A^*A)_{12} = 1-i$. Therefore complex symmetric matrices need not be normal. ◆

EXERCISES

1. Label the following statements as true or false. Assume that the underlying inner product spaces are finite-dimensional.
   (a) Every self-adjoint operator is normal.
   (b) Operators and their adjoints have the same eigenvectors.
   (c) If T is an operator on an inner product space V, then T is normal if and only if $[\mathsf{T}]_\beta$ is normal, where β is any ordered basis for V.
   (d) A real or complex matrix A is normal if and only if L_A is normal.
   (e) The eigenvalues of a self-adjoint operator must all be real.
   (f) The identity and zero operators are self-adjoint.

(g) Every normal operator is diagonalizable.
(h) Every self-adjoint operator is diagonalizable.

2. For each linear operator T on an inner product space V, determine whether T is normal, self-adjoint, or neither. If possible, produce an orthonormal basis of eigenvectors of T for V and list the corresponding eigenvalues.

(a) $\mathsf{V} = \mathsf{R}^2$ and T is defined by $\mathsf{T}(a, b) = (2a - 2b, -2a + 5b)$.
(b) $\mathsf{V} = \mathsf{R}^3$ and T is defined by $\mathsf{T}(a, b, c) = (-a + b, 5b, 4a - 2b + 5c)$.
(c) $\mathsf{V} = \mathsf{C}^2$ and T is defined by $\mathsf{T}(a, b) = (2a + ib, a + 2b)$.
(d) $\mathsf{V} = \mathsf{P}_2(R)$ and T is defined by $\mathsf{T}(f) = f'$, where

$$\langle f(x), g(x) \rangle = \int_0^1 f(t)g(t)\, dt.$$

(e) $\mathsf{V} = \mathsf{M}_{2\times 2}(R)$ and T is defined by $\mathsf{T}(A) = A^t$.
(f) $\mathsf{V} = \mathsf{M}_{2\times 2}(R)$ and T is defined by $\mathsf{T}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} c & d \\ a & b \end{pmatrix}$.

3. Give an example of a linear operator T on R^2 and an ordered basis for R^2 that provides a counterexample to the statement in Exercise 1(c).

4. Let T and U be self-adjoint operators on an inner product space V. Prove that TU is self-adjoint if and only if TU = UT.

5. Prove (b) of Theorem 6.15.

6. Let V be a complex inner product space, and let T be a linear operator on V. Define

$$\mathsf{T}_1 = \frac{1}{2}(\mathsf{T} + \mathsf{T}^*) \quad \text{and} \quad \mathsf{T}_2 = \frac{1}{2i}(\mathsf{T} - \mathsf{T}^*).$$

(a) Prove that T_1 and T_2 are self-adjoint and that $\mathsf{T} = \mathsf{T}_1 + i\,\mathsf{T}_2$.
(b) Suppose also that $\mathsf{T} = \mathsf{U}_1 + i\mathsf{U}_2$, where U_1 and U_2 are self-adjoint. Prove that $\mathsf{U}_1 = \mathsf{T}_1$ and $\mathsf{U}_2 = \mathsf{T}_2$.
(c) Prove that T is normal if and only if $\mathsf{T}_1\mathsf{T}_2 = \mathsf{T}_2\mathsf{T}_1$.

7. Let T be a linear operator on an inner product space V, and let W be a T-invariant subspace of V. Prove the following results.

(a) If T is self-adjoint, then T_W is self-adjoint.
(b) $\mathsf{W}^\perp$ is T^*-invariant.
(c) If W is both T- and T^*-invariant, then $(\mathsf{T}_\mathsf{W})^* = (\mathsf{T}^*)_\mathsf{W}$.
(d) If W is both T- and T^*-invariant and T is normal, then T_W is normal.

8. Let T be a normal operator on a finite-dimensional complex inner product space V, and let W be a subspace of V. Prove that if W is T-invariant, then W is also T*-invariant. *Hint:* Use Exercise 24 of Section 5.4.

9. Let T be a normal operator on a finite-dimensional inner product space V. Prove that $N(T) = N(T^*)$ and $R(T) = R(T^*)$. *Hint:* Use Theorem 6.15 and Exercise 12 of Section 6.3.

10. Let T be a self-adjoint operator on a finite-dimensional inner product space V. Prove that for all $x \in V$

$$\|T(x) \pm ix\|^2 = \|T(x)\|^2 + \|x\|^2.$$

Deduce that $T - iI$ is invertible and that the adjoint of $(T - iI)^{-1}$ is $(T + iI)^{-1}$.

11. Assume that T is a linear operator on a complex (not necessarily finite-dimensional) inner product space V with an adjoint T^*. Prove the following results.

(a) If T is self-adjoint, then $\langle T(x), x\rangle$ is real for all $x \in V$.
(b) If T satisfies $\langle T(x), x\rangle = 0$ for all $x \in V$, then $T = T_0$. *Hint:* Replace x by $x + y$ and then by $x + iy$, and expand the resulting inner products.
(c) If $\langle T(x), x\rangle$ is real for all $x \in V$, then T is self-adjoint.

12. Let T be a normal operator on a finite-dimensional real inner product space V whose characteristic polynomial splits. Prove that V has an orthonormal basis of eigenvectors of T. Hence prove that T is self-adjoint.

13. An $n \times n$ real matrix A is said to be a **Gramian** matrix if there exists a real (square) matrix B such that $A = B^tB$. Prove that A is a Gramian matrix if and only if A is symmetric and all of its eigenvalues are nonnegative. *Hint:* Apply Theorem 6.17 to $T = L_A$ to obtain an orthonormal basis $\{v_1, v_2, \ldots, v_n\}$ of eigenvectors with the associated eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$. Define the linear operator U by $U(v_i) = \sqrt{\lambda_i}v_i$.

14. *Simultaneous Diagonalization.* Let V be a finite-dimensional real inner product space, and let U and T be self-adjoint linear operators on V such that $UT = TU$. Prove that there exists an orthonormal basis for V consisting of vectors that are eigenvectors of both U and T. (The complex version of this result appears as Exercise 10 of Section 6.6.) *Hint:* For any eigenspace $W = E_\lambda$ of T, we have that W is both T- and U-invariant. By Exercise 7, we have that $W^\perp$ is both T- and U-invariant. Apply Theorem 6.17 and Theorem 6.6 (p. 347).

15. Let A and B be symmetric $n \times n$ matrices such that $AB = BA$. Use Exercise 14 to prove that there exists an orthogonal matrix P such that P^tAP and P^tBP are both diagonal matrices.

16. Prove the *Cayley–Hamilton theorem* for a complex $n \times n$ matrix A. That is, if $f(t)$ is the characteristic polynomial of A, prove that $f(A) = O$. *Hint:* Use Schur's theorem to show that A may be assumed to be upper triangular, in which case

$$f(t) = \prod_{i=1}^{n} (A_{ii} - t).$$

Now if $\mathsf{T} = \mathsf{L}_A$, we have $(A_{jj}\mathsf{I} - \mathsf{T})(e_j) \in \text{span}(\{e_1, e_2, \ldots, e_{j-1}\})$ for $j \geq 2$, where $\{e_1, e_2, \ldots, e_n\}$ is the standard ordered basis for C^n. (The general case is proved in Section 5.4.)

The following definitions are used in Exercises 17 through 23.

Definitions. *A linear operator* T *on a finite-dimensional inner product space is called* **positive definite** [**positive semidefinite**] *if* T *is self-adjoint and* $\langle \mathsf{T}(x), x\rangle > 0$ [$\langle \mathsf{T}(x), x\rangle \geq 0$] *for all* $x \neq 0$.

An $n \times n$ *matrix* A *with entries from* R *or* C *is called* **positive definite** [**positive semidefinite**] *if* L_A *is positive definite* [*positive semidefinite*].

17. Let T and U be self-adjoint linear operators on an n-dimensional inner product space V, and let $A = [\mathsf{T}]_\beta$, where β is an orthonormal basis for V. Prove the following results.

(a) T is positive definite [semidefinite] if and only if all of its eigenvalues are positive [nonnegative].

(b) T is positive definite if and only if

$$\sum_{i,j} A_{ij}a_j\bar{a}_i > 0 \text{ for all nonzero } n\text{-tuples } (a_1, a_2, \ldots, a_n).$$

(c) T is positive semidefinite if and only if $A = B^*B$ for some square matrix B.

(d) If T and U are positive semidefinite operators such that $\mathsf{T}^2 = \mathsf{U}^2$, then $\mathsf{T} = \mathsf{U}$.

(e) If T and U are positive definite operators such that $\mathsf{TU} = \mathsf{UT}$, then TU is positive definite.

(f) T is positive definite [semidefinite] if and only if A is positive definite [semidefinite].

Because of (f), results analogous to items (a) through (d) hold for matrices as well as operators.

18. Let $\mathsf{T}: \mathsf{V} \to \mathsf{W}$ be a linear transformation, where V and W are finite-dimensional inner product spaces. Prove the following results.

(a) $\mathsf{T}^*\mathsf{T}$ and $\mathsf{T}\mathsf{T}^*$ are positive semidefinite. (See Exercise 15 of Section 6.3.)

(b) $\text{rank}(\mathsf{T}^*\mathsf{T}) = \text{rank}(\mathsf{T}\mathsf{T}^*) = \text{rank}(\mathsf{T})$.

19. Let T and U be positive definite operators on an inner product space V. Prove the following results.

(a) $\mathsf{T} + \mathsf{U}$ is positive definite.

(b) If $c > 0$, then $c\mathsf{T}$ is positive definite.

(c) T^{-1} is positive definite.

Visit `goo.gl/cQch7i` for a solution.

20. Let V be an inner product space with inner product $\langle \cdot, \cdot \rangle$, and let T be a positive definite linear operator on V. Prove that $\langle x, y \rangle' = \langle \mathsf{T}(x), y \rangle$ defines another inner product on V.

21. Let V be a finite-dimensional inner product space, and let T and U be self-adjoint operators on V such that T is positive definite. Prove that both TU and UT are diagonalizable linear operators that have only real eigenvalues. *Hint:* Show that UT is self-adjoint with respect to the inner product $\langle x, y \rangle' = \langle \mathsf{T}(x), y \rangle$. To show that TU is self-adjoint, repeat the argument with T^{-1} in place of T.

22. This exercise provides a converse to Exercise 20. Let V be a finite-dimensional inner product space with inner product $\langle \cdot, \cdot \rangle$, and let $\langle \cdot, \cdot \rangle'$ be any other inner product on V.

(a) Prove that there exists a unique linear operator T on V such that $\langle x, y \rangle' = \langle \mathsf{T}(x), y \rangle$ for all x and y in V. *Hint:* Let $\beta = \{v_1, v_2, \ldots, v_n\}$ be an orthonormal basis for V with respect to $\langle \cdot, \cdot \rangle$, and define a matrix A by $A_{ij} = \langle v_j, v_i \rangle'$ for all i and j. Let T be the unique linear operator on V such that $[\mathsf{T}]_\beta = A$.

(b) Prove that the operator T of (a) is positive definite with respect to both inner products.

23. Let U be a diagonalizable linear operator on a finite-dimensional inner product space V such that all of the eigenvalues of U are real. Prove that there exist positive definite linear operators T_1 and T_1' and self-adjoint linear operators T_2 and T_2' such that $\mathsf{U} = \mathsf{T}_2\mathsf{T}_1 = \mathsf{T}_1'\mathsf{T}_2'$. *Hint:* Let $\langle \cdot, \cdot \rangle$ be the inner product associated with V, β a basis of eigenvectors for U, $\langle \cdot, \cdot \rangle'$ the inner product on V with respect to which β is orthonormal (see Exercise 22(a) of Section 6.1), and T_1 the positive definite operator according to Exercise 22. Show that U is self-adjoint with respect to $\langle \cdot, \cdot \rangle'$ and $\mathsf{U} = \mathsf{T}_1^{-1}\mathsf{U}^*\mathsf{T}_1$ (the adjoint is with respect to $\langle \cdot, \cdot \rangle$). Let $\mathsf{T}_2 = \mathsf{T}_1{}^{-1}\mathsf{U}^*$.

which is the matrix representation of a rotation through the angle $\theta = \sin^{-1}\left(-\frac{2}{\sqrt{5}}\right) \approx -63.4°$. This possibility produces the same ellipse as the one in Figure 6.4, but interchanges the names of the x'- and y'-axes.

EXERCISES

1. Label the following statements as true or false. Assume that the underlying inner product spaces are finite-dimensional.
   (a) Every unitary operator is normal.
   (b) Every orthogonal operator is diagonalizable.
   (c) A matrix is unitary if and only if it is invertible.
   (d) If two matrices are unitarily equivalent, then they are also similar.
   (e) The sum of unitary matrices is unitary.
   (f) The adjoint of a unitary operator is unitary.
   (g) If T is an orthogonal operator on V, then $[\mathsf{T}]_\beta$ is an orthogonal matrix for any ordered basis β for V.
   (h) If all the eigenvalues of a linear operator are 1, then the operator must be unitary or orthogonal.
   (i) A linear operator may preserve norms without preserving inner products.

2. For each of the following matrices A, find an orthogonal or unitary matrix P and a diagonal matrix D such that $P^*AP = D$.

   (a) $\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ (b) $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ (c) $\begin{pmatrix} 2 & 3-3i \\ 3+3i & 5 \end{pmatrix}$

   (d) $\begin{pmatrix} 0 & 2 & 2 \\ 2 & 0 & 2 \\ 2 & 2 & 0 \end{pmatrix}$ (e) $\begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$

3. Prove that the composite of unitary [orthogonal] operators is unitary [orthogonal].

4. For $z \in C$, define $\mathsf{T}_z \colon C \to C$ by $\mathsf{T}_z(u) = zu$. Characterize those z for which T_z is normal, self-adjoint, or unitary.

5. Which of the following pairs of matrices are unitarily equivalent?

   (a) $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ and $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ (b) $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ and $\begin{pmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{pmatrix}$

(c) $\begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ and $\begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$

(d) $\begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ and $\begin{pmatrix} 1 & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & -i \end{pmatrix}$

(e) $\begin{pmatrix} 1 & 1 & 0 \\ 0 & 2 & 2 \\ 0 & 0 & 3 \end{pmatrix}$ and $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$

6. Let V be the inner product space of complex-valued continuous functions on $[0, 1]$ with the inner product

$$\langle f, g \rangle = \int_0^1 f(t)\overline{g(t)}\, dt.$$

Let $h \in \mathsf{V}$, and define $\mathsf{T}\colon \mathsf{V} \to \mathsf{V}$ by $\mathsf{T}(f) = hf$. Prove that T is a unitary operator if and only if $|h(t)| = 1$ for $0 \leq t \leq 1$.

red *Hint for the "only if" part:* Suppose that T is unitary. Set $f(t) = 1 - |h(t)|^2$ and $g(t) = 1$. Show that

$$\int_0^1 (1 - |h(t)|^2)^2\, dt = 0,$$

and use the fact that if the integral of a nonnegative continuous function is zero, then the function is identically zero.

7. Prove that if T is a unitary operator on a finite-dimensional inner product space V, then T has a unitary *square root*; that is, there exists a unitary operator U such that $\mathsf{T} = \mathsf{U}^2$. Visit `goo.gl/jADTaS` for a solution.

8. Let T be a self-adjoint linear operator on a finite-dimensional inner product space. Prove that $(\mathsf{T}+i\mathsf{I})(\mathsf{T}-i\mathsf{I})^{-1}$ is unitary using Exercise 10 of Section 6.4.

9. Let U be a linear operator on a finite-dimensional inner product space V. If $\|\mathsf{U}(x)\| = \|x\|$ for all x in some orthonormal basis for V, must U be unitary? Justify your answer with a proof or a counterexample.

10. Let A be an $n \times n$ real symmetric or complex normal matrix. Prove that

$$\operatorname{tr}(A) = \sum_{i=1}^{n} \lambda_i \quad \text{and} \quad \operatorname{tr}(A^*A) = \sum_{i=1}^{n} |\lambda_i|^2,$$

where the λ_i's are the (not necessarily distinct) eigenvalues of A.

11. Find an orthogonal matrix whose first row is $(\frac{1}{3}, \frac{2}{3}, \frac{2}{3})$.

12. Let A be an $n \times n$ real symmetric or complex normal matrix. Prove that

$$\det(A) = \prod_{i=1}^{n} \lambda_i,$$

where the λ_i's are the (not necessarily distinct) eigenvalues of A.

13. Suppose that A and B are diagonalizable matrices. Prove or disprove that A is similar to B if and only if A and B are unitarily equivalent.

14. Prove that if A and B are unitarily equivalent matrices, then A is positive definite [semidefinite] if and only if B is positive definite [semidefinite]. (See the definitions in the exercises in Section 6.4.)

15. Let U be a unitary operator on an inner product space V, and let W be a finite-dimensional U-invariant subspace of V. Prove that

 (a) $\mathsf{U}(\mathsf{W}) = \mathsf{W}$;
 (b) $\mathsf{W}^{\perp}$ is U-invariant.

 Contrast (b) with Exercise 16.

16. Find an example of a unitary operator U on an inner product space and a U-invariant subspace W such that $\mathsf{W}^{\perp}$ is not U-invariant.

17. Prove that a matrix that is both unitary and upper triangular must be a diagonal matrix.

18. Show that "is unitarily equivalent to" is an equivalence relation on $\mathsf{M}_{n\times n}(C)$.

19. Let W be a finite-dimensional subspace of an inner product space V. By Theorem 6.7 (p. 349) and the exercises of Section 1.3, $\mathsf{V} = \mathsf{W} \oplus \mathsf{W}^{\perp}$. Define $\mathsf{U}\colon \mathsf{V} \to \mathsf{V}$ by $\mathsf{U}(v_1 + v_2) = v_1 - v_2$, where $v_1 \in \mathsf{W}$ and $v_2 \in \mathsf{W}^{\perp}$. Prove that U is a self-adjoint unitary operator.

20. Let V be a finite-dimensional inner product space. A linear operator U on V is called a **partial isometry** if there exists a subspace W of V such that $\|\mathsf{U}(x)\| = \|x\|$ for all $x \in \mathsf{W}$ and $\mathsf{U}(x) = 0$ for all $x \in \mathsf{W}^{\perp}$. Observe that W need *not* be U-invariant. Suppose that U is such an operator and $\{v_1, v_2, \ldots, v_k\}$ is an orthonormal basis for W. Prove the following results.

 (a) $\langle \mathsf{U}(x), \mathsf{U}(y)\rangle = \langle x, y\rangle$ for all $x, y \in \mathsf{W}$. *Hint:* Use Exercise 20 of Section 6.1.
 (b) $\{\mathsf{U}(v_1), \mathsf{U}(v_2), \ldots, \mathsf{U}(v_k)\}$ is an orthonormal basis for $\mathsf{R}(\mathsf{U})$.

(c) There exists an orthonormal basis γ for V such that the first k columns of $[\mathsf{U}]_\gamma$ form an orthonormal set and the remaining columns are zero.

(d) Let $\{w_1, w_2, \ldots, w_j\}$ be an orthonormal basis for $\mathsf{R(U)}^\perp$ and $\beta = \{\mathsf{U}(v_1), \mathsf{U}(v_2), \ldots, \mathsf{U}(v_k), w_1, \ldots, w_j\}$. Then β is an orthonormal basis for V.

(e) Let T be the linear operator on V that satisfies $\mathsf{T}(\mathsf{U}(v_i)) = v_i$ $(1 \le i \le k)$ and $\mathsf{T}(w_i) = 0$ $(1 \le i \le j)$. Then T is well defined, and $\mathsf{T} = \mathsf{U}^*$. *Hint:* Show that $\langle \mathsf{U}(x), y\rangle = \langle x, \mathsf{T}(y)\rangle$ for all $x, y \in \beta$. There are four cases.

(f) U^* is a partial isometry.

This exercise is continued in Exercise 9 of Section 6.6.

21. Let A and B be $n \times n$ matrices that are unitarily equivalent.

(a) Prove that $\mathrm{tr}(A^*A) = \mathrm{tr}(B^*B)$.

(b) Use (a) to prove that

$$\sum_{i,j=1}^{n} |A_{ij}|^2 = \sum_{i,j=1}^{n} |B_{ij}|^2.$$

(c) Use (b) to show that the matrices

$$\begin{pmatrix} 1 & 2 \\ 2 & i \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} i & 4 \\ 1 & 1 \end{pmatrix}$$

are *not* unitarily equivalent.

22. Let V be a real inner product space.

(a) Prove that any translation on V is a rigid motion.

(b) Prove that the composite of any two rigid motions on V is a rigid motion on V.

23. Prove the following variant of Theorem 6.22: If $f\colon \mathsf{V} \to \mathsf{V}$ is a rigid motion on a finite-dimensional real inner product space V, then there exists a unique orthogonal operator T on V and a unique translation g on V such that $f = \mathsf{T} \circ g$. (Note that the conclusion of Theorem 6.22 has $f = g \circ \mathsf{T}$).

24. Let T and U be orthogonal operators on R^2. Use Theorem 6.23 to prove the following results.

(a) If T and U are both reflections about lines through the origin, then UT is a rotation.

(b) If T is a rotation and U is a reflection about a line through the origin, then both UT and TU are reflections about lines through the origin.

25. Suppose that T and U are reflections of R^2 about the respective lines L and L′ through the origin and that ϕ and ψ are the angles from the positive x-axis to L and L′, respectively. By Exercise 24, UT is a rotation. Find its angle of rotation.

26. Suppose that T and U are orthogonal operators on R^2 such that T is the rotation by the angle ϕ and U is the reflection about the line L through the origin. Let ψ be the angle from the positive x-axis to L. By Exercise 24, both UT and TU are reflections about lines L_1 and L_2, respectively, through the origin.

(a) Find the angle θ from the positive x-axis to L_1..
(b) Find the angle θ from the positive x-axis to L_2.

27. Find new coordinates x', y' so that the following quadratic forms can be written as $\lambda_1(x')^2 + \lambda_2(y')^2$.

(a) $x^2 + 4xy + y^2$
(b) $2x^2 + 2xy + 2y^2$
(c) $x^2 - 12xy - 4y^2$
(d) $3x^2 + 2xy + 3y^2$
(e) $x^2 - 2xy + y^2$

28. Consider the expression X^tAX, where $X^t = (x, y, z)$ and A is as defined in Exercise 2(e). Find a change of coordinates x', y', z' so that the preceding expression is of the form $\lambda_1(x')^2 + \lambda_2(y')^2 + \lambda_3(z')^2$.

29. *QR-Factorization.* Let $w_1, w_2, \ldots, w_n$ be linearly independent vectors in F^n, and let $v_1, v_2, \ldots, v_n$ be the orthogonal vectors obtained from $w_1, w_2, \ldots, w_n$ by the Gram–Schmidt process. Let $u_1, u_2, \ldots, u_n$ be the orthonormal basis obtained by normalizing the v_i's.

(a) Solving (1) in Section 6.2 for w_k in terms of u_k, show that

$$w_k = \|v_k\| u_k + \sum_{j=1}^{k-1} \langle w_k, u_j \rangle\, u_j \quad (1 \le k \le n).$$

(b) Let A and Q denote the $n \times n$ matrices in which the kth columns are w_k and u_k, respectively. Define $R \in \mathsf{M}_{n\times n}(F)$ by

$$R_{jk} = \begin{cases} \|v_j\| & \text{if } j = k \\ \langle w_k, u_j \rangle & \text{if } j < k \\ 0 & \text{if } j > k. \end{cases}$$

Prove $A = QR$.

(c) Compute Q and R as in (b) for the 3×3 matrix whose columns are the vectors $(1, 1, 0)$, $(2, 0, 1)$, and $(2, 2, 1)$.

(d) Since Q is unitary [orthogonal] and R is upper triangular in (b), we have shown that every invertible matrix is the product of a unitary [orthogonal] matrix and an upper triangular matrix. Suppose that $A \in \mathsf{M}_{n\times n}(F)$ is invertible and $A = Q_1R_1 = Q_2R_2$, where $Q_1, Q_2 \in \mathsf{M}_{n\times n}(F)$ are unitary and $R_1, R_2 \in \mathsf{M}_{n\times n}(F)$ are upper triangular. Prove that $D = R_2R_1^{-1}$ is a unitary diagonal matrix. *Hint:* Use Exercise 17.

(e) The QR factorization described in (b) provides an orthogonalization method for solving a linear system $Ax = b$ when A is invertible. Decompose A to QR, by the Gram–Schmidt process or other means, where Q is unitary and R is upper triangular. Then $QRx = b$, and hence $Rx = Q^*b$. This last system can be easily solved since R is upper triangular. [1]

Use the orthogonalization method and (c) to solve the system

$$\begin{aligned} x_1 + 2x_2 + 2x_3 &= 1 \\ x_1 + 2x_3 &= 11 \\ x_2 + x_3 &= -1. \end{aligned}$$

30. Suppose that β and γ are ordered bases for an n-dimensional real [complex] inner product space V. Prove that if Q is an orthogonal [unitary] $n \times n$ matrix that changes γ-coordinates into β-coordinates, then β is orthonormal if and only if γ is orthonormal.

The following definition is used in Exercises 31 and 32.

Definition. *Let* V *be a finite-dimensional complex [real] inner product space, and let u be a unit vector in* V. *Define the* **Householder** *operator* $\mathsf{H}_u \colon \mathsf{V} \to \mathsf{V}$ *by* $\mathsf{H}_u(x) = x - 2\langle x, u\rangle u$ *for all* $x \in \mathsf{V}$.

31. Let H_u be a Householder operator on a finite-dimensional inner product space V. Prove the following results.

(a) H_u is linear.
(b) $\mathsf{H}_u(x) = x$ if and only if x is orthogonal to u.
(c) $\mathsf{H}_u(u) = -u$.
(d) $\mathsf{H}_u^* = \mathsf{H}_u$ and $\mathsf{H}_u^2 = \mathsf{I}$, and hence H_u is a unitary [orthogonal] operator on V.

(*Note:* If V is a real inner product space, then in the language of Section 6.11, H_u is a reflection.)

---

[1] At one time, because of its great stability, this method for solving large systems of linear equations with a computer was being advocated as a better method than Gaussian elimination even though it requires about three times as much work. (Later, however, J. H. Wilkinson showed that if Gaussian elimination is done "properly," then it is nearly as stable as the orthogonalization method.)

32. Let V be a finite-dimensional inner product space over F. Let x and y be linearly independent vectors in V such that $\|x\| = \|y\|$.

(a) If $F = C$, prove that there exists a unit vector u in V and a complex number θ with $|\theta| = 1$ such that $\mathsf{H}_u(x) = \theta y$. *Hint:* Choose θ so that $\langle x, \theta y\rangle$ is real, and set $u = \dfrac{1}{\|x - \theta y\|}(x - \theta y)$.

(b) If $F = R$, prove that there exists a unit vector u in V such that $\mathsf{H}_u(x) = y$.

6.6 ORTHOGONAL PROJECTIONS AND THE SPECTRAL THEOREM

In this section, we rely heavily on Theorems 6.16 (p. 369) and 6.17 (p. 371) to develop an elegant representation of a normal (if $F = C$) or a self-adjoint (if $F = R$) operator T on a finite-dimensional inner product space. We prove that T can be written in the form $\lambda_1\mathsf{T}_1 + \lambda_2\mathsf{T}_2 + \cdots + \lambda_k\mathsf{T}_k$, where $\lambda_1, \lambda_2, \ldots, \lambda_k$ are the distinct eigenvalues of T and $\mathsf{T}_1, \mathsf{T}_2, \ldots, \mathsf{T}_k$ are *orthogonal projections.* We must first develop some results about these special projections.

We assume that the reader is familiar with the results about direct sums developed at the end of Section 5.2. The special case where V is a direct sum of two subspaces is considered in the exercises of Section 1.3.

Recall from the exercises of Section 2.1 that if $\mathsf{V} = \mathsf{W}_1 \oplus \mathsf{W}_2$, then a linear operator T on V is the **projection on W_1 along W_2** if, whenever $x = x_1 + x_2$, with $x_1 \in \mathsf{W}_1$ and $x_2 \in \mathsf{W}_2$, we have $\mathsf{T}(x) = x_1$. By Exercise 27 of Section 2.1, we have

$$\mathsf{R}(\mathsf{T}) = \mathsf{W}_1 = \{x \in \mathsf{V} \colon \mathsf{T}(x) = x\} \quad \text{and} \quad \mathsf{N}(\mathsf{T}) = \mathsf{W}_2.$$

So $\mathsf{V} = \mathsf{R}(\mathsf{T}) \oplus \mathsf{N}(\mathsf{T})$. Thus there is no ambiguity if we refer to T as a "projection on W_1" or simply as a "projection." In fact, it can be shown (see Exercise 17 of Section 2.3) that T is a projection if and only if $\mathsf{T} = \mathsf{T}^2$. Because $\mathsf{V} = \mathsf{W}_1 \oplus \mathsf{W}_2 = \mathsf{W}_1 \oplus \mathsf{W}_3$ does *not* imply that $\mathsf{W}_2 = \mathsf{W}_3$, we see that W_1 does not uniquely determine T. For an *orthogonal* projection T, however, T is uniquely determined by its range.

Definition. *Let* V *be an inner product space, and let* $\mathsf{T}\colon \mathsf{V} \to \mathsf{V}$ *be a projection. We say that* T *is an* **orthogonal projection** *if* $\mathsf{R}(\mathsf{T})^\perp = \mathsf{N}(\mathsf{T})$ *and* $\mathsf{N}(\mathsf{T})^\perp = \mathsf{R}(\mathsf{T})$.

Note that by Exercise 13(c) of Section 6.2, if V is finite-dimensional, we need only assume that one of the equalities in this definition holds. For example, if $\mathsf{R}(\mathsf{T})^\perp = \mathsf{N}(\mathsf{T})$, then $\mathsf{R}(\mathsf{T}) = \mathsf{R}(\mathsf{T})^{\perp\perp} = \mathsf{N}(\mathsf{T})^\perp$.

An orthogonal projection is *not* the same as an orthogonal operator. In Figure 6.5, T is an orthogonal projection, but T is clearly not an orthogonal operator because $\|\mathsf{T}(v)\| \neq \|v\|$.

Proof. If T is unitary, then T is normal and every eigenvalue of T has absolute value 1 by Corollary 2 to Theorem 6.18 (p. 379).

Let $\mathsf{T} = \lambda_1\mathsf{T}_1 + \lambda_2\mathsf{T}_2 + \cdots + \lambda_k\mathsf{T}_k$ be the spectral decomposition of T. If $|\lambda| = 1$ for every eigenvalue λ of T, then by (c) of the spectral theorem,

$$\begin{aligned}\mathsf{TT}^* &= (\lambda_1\mathsf{T}_1 + \lambda_2\mathsf{T}_2 + \cdots + \lambda_k\mathsf{T}_k)(\overline{\lambda_1}\mathsf{T}_1 + \overline{\lambda_2}\mathsf{T}_2 + \cdots + \overline{\lambda_k}\mathsf{T}_k)\\ &= |\lambda_1|^2\mathsf{T}_1 + |\lambda_2|^2\mathsf{T}_2 + \cdots + |\lambda_k|^2\mathsf{T}_k\\ &= \mathsf{T}_1 + \mathsf{T}_2 + \cdots + \mathsf{T}_k\\ &= \mathsf{I}.\end{aligned}$$

Hence T is unitary. ∎

Corollary 3. *If $F = C$, then* T *is self-adjoint if and only if* T *is normal and every eigenvalue of* T *is real.*

Proof. Suppose that T is normal and that its eigenvalues are real. Let $\mathsf{T} = \lambda_1\mathsf{T}_1 + \lambda_2\mathsf{T}_2 + \cdots + \lambda_k\mathsf{T}_k$ be the spectral decomposition of T. Then

$$\mathsf{T}^* = \overline{\lambda_1}\mathsf{T}_1 + \overline{\lambda_2}\mathsf{T}_2 + \cdots + \overline{\lambda_k}\mathsf{T}_k = \lambda_1\mathsf{T}_1 + \lambda_2\mathsf{T}_2 + \cdots + \lambda_k\mathsf{T}_k = \mathsf{T}.$$

Hence T is self-adjoint.

Now suppose that T is self-adjoint and hence normal. That its eigenvalues are real has been proved in the lemma to Theorem 6.17 (p. 371). ∎

Corollary 4. *Let* T *be as in the spectral theorem with spectral decomposition* $\mathsf{T} = \lambda_1\mathsf{T}_1 + \lambda_2\mathsf{T}_2 + \cdots + \lambda_k\mathsf{T}_k$. *Then each* T_j *is a polynomial in* T.

Proof. Choose a polynomial g_j $(1 \le j \le k)$ such that $g_j(\lambda_i) = \delta_{ij}$. Then

$$\begin{aligned}g_j(\mathsf{T}) &= g_j(\lambda_1)\mathsf{T}_1 + g_j(\lambda_2)\mathsf{T}_2 + \cdots + g_j(\lambda_k)\mathsf{T}_k\\ &= \delta_{1j}\mathsf{T}_1 + \delta_{2j}\mathsf{T}_2 + \cdots + \delta_{kj}\mathsf{T}_k = \mathsf{T}_j.\end{aligned}$$ ∎

EXERCISES

1. Label the following statements as true or false. Assume that the underlying inner product spaces are finite-dimensional.
   (a) All projections are self-adjoint.
   (b) An orthogonal projection is uniquely determined by its range.
   (c) Every self-adjoint operator is a linear combination of orthogonal projections.
   (d) If T is a projection on W, then $\mathsf{T}(x)$ is the vector in W that is closest to x.

(e) Every orthogonal projection is a unitary operator.

2. Let $\mathsf{V} = \mathsf{R}^2$, $\mathsf{W} = \text{span}(\{(1,2)\})$, and β be the standard ordered basis for V. Compute $[\mathsf{T}]_\beta$, where T is the orthogonal projection of V on W. Do the same for $\mathsf{V} = \mathsf{R}^3$ and $\mathsf{W} = \text{span}(\{(1,0,1)\})$.

3. For each of the matrices A in Exercise 2 of Section 6.5:

   (1) Verify that L_A possesses a spectral decomposition.
   (2) For each eigenvalue of L_A, explicitly define the orthogonal projection on the corresponding eigenspace.
   (3) Verify your results using the spectral theorem.

4. Let W be a finite-dimensional subspace of an inner product space V. Show that if T is the orthogonal projection of V on W, then $\mathsf{I} - \mathsf{T}$ is the orthogonal projection of V on $\mathsf{W}^\perp$.

5. Let T be a linear operator on a finite-dimensional inner product space V.

   (a) If T is an orthogonal projection, prove that $\|\mathsf{T}(x)\| \le \|x\|$ for all $x \in \mathsf{V}$. Give an example of a projection for which this inequality does not hold. What can be concluded about a projection for which the inequality is actually an equality for all $x \in \mathsf{V}$?
   (b) Suppose that T is a projection such that $\|\mathsf{T}(x)\| \le \|x\|$ for all $x \in \mathsf{V}$. Prove that T is an orthogonal projection.

6. Let T be a normal operator on a finite-dimensional inner product space. Prove that if T is a projection, then T is also an orthogonal projection.

7. Let T be a normal operator on a finite-dimensional complex inner product space V. Use the spectral decomposition $\lambda_1\mathsf{T}_1 + \lambda_2\mathsf{T}_2 + \cdots + \lambda_k\mathsf{T}_k$ of T to prove the following results.

   (a) If g is a polynomial, then

   $$g(\mathsf{T}) = \sum_{i=1}^{k} g(\lambda_i)\mathsf{T}_i.$$

   (b) If $\mathsf{T}^n = \mathsf{T}_0$ for some n, then $\mathsf{T} = \mathsf{T}_0$.
   (c) Let U be a linear operator on V. Then U commutes with T if and only if U commutes with each T_i.
   (d) There exists a normal operator U on V such that $\mathsf{U}^2 = \mathsf{T}$.
   (e) T is invertible if and only if $\lambda_i \neq 0$ for $1 \le i \le k$.
   (f) T is a projection if and only if every eigenvalue of T is 1 or 0.
   (g) $\mathsf{T} = -\mathsf{T}^*$ if and only if every λ_i is an imaginary number.

8. Use Corollary 1 of the spectral theorem to show that if T is a normal operator on a complex finite-dimensional inner product space and U is a linear operator that commutes with T, then U commutes with T^*.

9. Referring to Exercise 20 of Section 6.5, prove the following facts about a partial isometry U.

   (a) U^*U is an orthogonal projection on W.
   (b) $UU^*U = U$.

10. *Simultaneous diagonalization.* Let U and T be normal operators on a finite-dimensional complex inner product space V such that $TU = UT$. Prove that there exists an orthonormal basis for V consisting of vectors that are eigenvectors of both T and U. *Hint:* Use the hint of Exercise 14 of Section 6.4 along with Exercise 8.

$\boxed{11.}$ Prove (c) of the spectral theorem. Visit `goo.gl/utQ9Pb` for a solution.

6.7* THE SINGULAR VALUE DECOMPOSITION AND THE PSEUDOINVERSE

In Section 6.4, we characterized normal operators on complex spaces and self-adjoint operators on real spaces in terms of orthonormal bases of eigenvectors and their corresponding eigenvalues (Theorems 6.16, p. 369, and 6.17, p. 371). In this section, we establish a comparable theorem whose scope is the entire class of linear transformations on both complex and real finite-dimensional inner product spaces—*the singular value theorem for linear transformations* (Theorem 6.26). There are similarities and differences among these theorems. All rely on the use of orthonormal bases and numerical invariants. However, because of its general scope, the singular value theorem is concerned with two (usually distinct) inner product spaces and with two (usually distinct) orthonormal bases. If the two spaces and the two bases are identical, then the transformation would, in fact, be a normal or self-adjoint operator. Another difference is that the numerical invariants in the singular value theorem, the *singular values*, are nonnegative, in contrast to their counterparts, the eigenvalues, for which there is no such restriction. This property is necessary to guarantee the uniqueness of singular values.

The singular value theorem encompasses both real and complex spaces. For brevity, in this section we use the terms *unitary operator* and *unitary matrix* to include orthogonal operators and orthogonal matrices in the context of real spaces. Thus any operator T for which $\langle \mathsf{T}(x), \mathsf{T}(y)\rangle = \langle x, y\rangle$, or any matrix A for which $\langle Ax, Ay\rangle = \langle x, y\rangle$, for all x and y is called *unitary* for the purposes of this section.

In Exercise 15 of Section 6.3, the definition of the adjoint of an operator is extended to any linear transformation $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$, where V and W are

Finally, suppose that y is any vector in F^n such that $Az = Ay = c$. Then

$$A^\dagger c = A^\dagger Az = A^\dagger AA^\dagger b = A^\dagger b = z$$

by Exercise 23; hence we may apply part (a) of this theorem to the system $Ax = c$ to conclude that $\|z\| \leq \|y\|$ with equality if and only if $z = y$. ∎

Note that the vector $z = A^\dagger b$ in Theorem 6.30 is the vector x_0 described in Theorem 6.12 that arises in the least squares application on pages 358–361.

Example 7

Consider the linear systems

$$\begin{aligned} x_1 + x_2 - x_3 &= 1 \\ x_1 + x_2 - x_3 &= 1 \end{aligned} \qquad \text{and} \qquad \begin{aligned} x_1 + x_2 - x_3 &= 1 \\ x_1 + x_2 - x_3 &= 2. \end{aligned}$$

The first system has infinitely many solutions. Let $A = \begin{pmatrix} 1 & 1 & -1 \\ 1 & 1 & -1 \end{pmatrix}$, the coefficient matrix of the system, and let $b = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. By Example 6,

$$A^\dagger = \frac{1}{6}\begin{pmatrix} 1 & 1 \\ 1 & 1 \\ -1 & -1 \end{pmatrix},$$

and therefore

$$z = A^\dagger b = \frac{1}{6}\begin{pmatrix} 1 & 1 \\ 1 & 1 \\ -1 & -1 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$$

is the solution of minimal norm by Theorem 6.30(a).

The second system is obviously inconsistent. Let $b = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$. Thus, although

$$z = A^\dagger b = \frac{1}{6}\begin{pmatrix} 1 & 1 \\ 1 & 1 \\ -1 & -1 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$$

is not a solution to the second system, it is the "best approximation" to a solution having minimum norm, as described in Theorem 6.30(b). ◆

EXERCISES

1. Label the following statements as true or false.

(a) The singular values of any linear operator on a finite-dimensional vector space are also eigenvalues of the operator.
(b) The singular values of any matrix A are the eigenvalues of A^*A.
(c) For any matrix A and any scalar c, if σ is a singular value of A, then $|c|\sigma$ is a singular value of cA.
(d) The singular values of any linear operator are nonnegative.
(e) If λ is an eigenvalue of a self-adjoint matrix A, then λ is a singular value of A.
(f) For any $m\times n$ matrix A and any $b \in \mathsf{F}^n$, the vector $A^\dagger b$ is a solution to $Ax = b$.
(g) The pseudoinverse of any linear operator exists even if the operator is not invertible.

2. Let $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$ be a linear transformation of rank r, where V and W are finite-dimensional inner product spaces. In each of the following, find orthonormal bases $\{v_1, v_2, \ldots, v_n\}$ for V and $\{u_1, u_2, \ldots, u_m\}$ for W, and the nonzero singular values $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r$ of T such that $\mathsf{T}(v_i) = \sigma_i u_i$ for $1 \leq i \leq r$.

(a) $\mathsf{T}\colon \mathsf{R}^2 \to \mathsf{R}^3$ defined by $\mathsf{T}(x_1, x_2) = (x_1,\ x_1 + x_2,\ x_1 - x_2)$
(b) $\mathsf{T}\colon \mathsf{P}_2(R) \to \mathsf{P}_1(R)$, where $\mathsf{T}(f(x)) = f''(x)$, and the inner products are defined as in Example 1
(c) Let $\mathsf{V} = \mathsf{W} = \mathrm{span}(\{1, \sin x, \cos x\})$ with the inner product defined by $\langle f, g\rangle = \int_0^{2\pi} f(t)g(t)\,dt$, and T is defined by $\mathsf{T}(f) = f' + 2f$
(d) $\mathsf{T}\colon \mathsf{C}^2 \to \mathsf{C}^2$ defined by $\mathsf{T}(z_1, z_2) = ((1-i)z_2,\ (1+i)z_1 + z_2)$

3. Find a singular value decomposition for each of the following matrices.

(a) $\begin{pmatrix} 1 & 1 \\ 1 & 1 \\ -1 & -1 \end{pmatrix}$ (b) $\begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & -1 \end{pmatrix}$ (c) $\begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \\ 1 & 1 \end{pmatrix}$

(d) $\begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{pmatrix}$ (e) $\begin{pmatrix} 1+i & 1 \\ 1-i & -i \end{pmatrix}$ (f) $\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 0 & -2 & 1 \\ 1 & -1 & 1 & 1 \end{pmatrix}$

4. Find a polar decomposition for each of the following matrices.

(a) $\begin{pmatrix} 1 & 1 \\ 2 & -2 \end{pmatrix}$ (b) $\begin{pmatrix} 20 & 4 & 0 \\ 0 & 0 & 1 \\ 4 & 20 & 0 \end{pmatrix}$

5. Find an explicit formula for each of the following expressions.

(a) $\mathsf{T}^\dagger(x_1, x_2, x_3)$, where T is the linear transformation of Exercise 2(a)
(b) $\mathsf{T}^\dagger(a + bx + cx^2)$, where T is the linear transformation of Exercise 2(b)

(c) $\mathsf{T}^\dagger(a + b\sin x + c\cos x)$, where T is the linear transformation of Exercise 2(c)
(d) $\mathsf{T}^\dagger(z_1, z_2)$, where T is the linear transformation of Exercise 2(d)

6. Use the results of Exercise 3 to find the pseudoinverse of each of the following matrices.

(a) $\begin{pmatrix} 1 & 1 \\ 1 & 1 \\ -1 & -1 \end{pmatrix}$ (b) $\begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & -1 \end{pmatrix}$ (c) $\begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \\ 1 & 1 \end{pmatrix}$

(d) $\begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{pmatrix}$ (e) $\begin{pmatrix} 1+i & 1 \\ 1-i & -i \end{pmatrix}$ (f) $\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 0 & -2 & 1 \\ 1 & -1 & 1 & 1 \end{pmatrix}$

7. For each of the given linear transformations $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$,

(i) Describe the subspace Z_1 of V such that $\mathsf{T}^\dagger\mathsf{T}$ is the orthogonal projection of V on Z_1.

(ii) Describe the subspace Z_2 of W such that $\mathsf{T}\mathsf{T}^\dagger$ is the orthogonal projection of W on Z_2.

(a) T is the linear transformation of Exercise 2(a)
(b) T is the linear transformation of Exercise 2(b)
(c) T is the linear transformation of Exercise 2(c)
(d) T is the linear transformation of Exercise 2(d)

8. For each of the given systems of linear equations,

(i) If the system is consistent, find the unique solution having minimum norm.

(ii) If the system is inconsistent, find the "best approximation to a solution" having minimum norm, as described in Theorem 6.30(b).

(Use your answers to parts (a) and (f) of Exercise 6.)

(a) $\begin{aligned} x_1 + x_2 &= 1 \\ x_1 + x_2 &= 2 \\ -x_1 + -x_2 &= 0 \end{aligned}$ (b) $\begin{aligned} x_1 + x_2 + x_3 + x_4 &= 2 \\ x_1 - 2x_3 + x_4 &= -1 \\ x_1 - x_2 + x_3 + x_4 &= 2 \end{aligned}$

9. Let V and W be finite-dimensional inner product spaces over F, and suppose that $\{v_1, v_2, \ldots, v_n\}$ and $\{u_1, u_2, \ldots, u_m\}$ are orthonormal bases for V and W, respectively. Let $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$ be a linear transformation of rank r, and suppose that $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$ are such that

$$\mathsf{T}(v_i) = \begin{cases} \sigma_i u_i & \text{if } 1 \leq i \leq r \\ 0 & \text{if } r < i. \end{cases}$$

(a) Prove that $\{u_1, u_2, \ldots, u_m\}$ is a set of eigenvectors of TT^* with corresponding eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_m$, where

$$\lambda_i = \begin{cases} \sigma_i^2 & \text{if } 1 \le i \le r \\ 0 & \text{if } r < i. \end{cases}$$

(b) Let A be an $m \times n$ matrix with real or complex entries. Prove that the nonzero singular values of A are the positive square roots of the nonzero eigenvalues of AA^*, including repetitions.

(c) Prove that TT^* and $\mathsf{T}^*\mathsf{T}$ have the same nonzero eigenvalues, including repetitions.

(d) State and prove a result for matrices analogous to (c).

10. Use Exercise 8 of Section 2.5 to obtain another proof of Theorem 6.27, the singular value decomposition theorem for matrices.

11. This exercise relates the singular values of a well-behaved linear operator or matrix to its eigenvalues.

(a) Let T be a normal linear operator on an n-dimensional inner product space with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$. Prove that the singular values of T are $|\lambda_1|, |\lambda_2|, \ldots, |\lambda_n|$.

(b) State and prove a result for matrices analogous to (a).

12. Let A be a normal matrix with an orthonormal basis of eigenvectors $\beta = \{v_1, v_2, \ldots, v_n\}$ and corresponding eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$. Let V be the $n \times n$ matrix whose columns are the vectors in β. Prove that for each i there is a scalar θ_i of absolute value 1 such that if U is the $n \times n$ matrix with $\theta_i v_i$ as column i and Σ is the diagonal matrix such that $\Sigma_{ii} = |\lambda_i|$ for each i, then $U\Sigma V^*$ is a singular value decomposition of A.

13. Prove that if A is a positive semidefinite matrix, then the singular values of A are the same as the eigenvalues of A.

14. Prove that if A is a positive definite matrix and $A = U\Sigma V^*$ is a singular value decomposition of A, then $U = V$.

15. Let A be a square matrix with a polar decomposition $A = WP$.

(a) Prove that A is normal if and only if $WP^2 = P^2W$.

(b) Use (a) to prove that A is normal if and only if $WP = PW$.

16. Let A be a square matrix. Prove an alternate form of the polar decomposition for A: There exists a unitary matrix W and a positive semidefinite matrix P such that $A = PW$.

17. Let T and U be linear operators on R^2 defined for all $(x_1, x_2) \in \mathsf{R}^2$ by

$$\mathsf{T}(x_1, x_2) = (x_1, 0) \text{ and } \mathsf{U}(x_1, x_2) = (x_1 + x_2, 0).$$

(a) Prove that $(\mathsf{UT})^\dagger \neq \mathsf{T}^\dagger \mathsf{U}^\dagger$.
(b) Exhibit matrices A and B such that AB is defined, but $(AB)^\dagger \neq B^\dagger A^\dagger$.

18. Let A be an $m \times n$ matrix. Prove the following results.

(a) For any $m \times m$ unitary matrix G, $(GA)^\dagger = A^\dagger G^*$.
(b) For any $n \times n$ unitary matrix H, $(AH)^\dagger = H^* A^\dagger$.

19. Let A be a matrix with real or complex entries. Prove the following results.

(a) The nonzero singular values of A are the same as the nonzero singular values of A^*, which are the same as the nonzero singular values of A^t.
(b) $(A^\dagger)^* = (A^*)^\dagger$.
(c) $(A^\dagger)^t = (A^t)^\dagger$.

20. Let A be a square matrix such that $A^2 = O$. Prove that $(A^\dagger)^2 = O$.

21. Let V and W be finite-dimensional inner product spaces, and let $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$ be linear. Prove the following results.

(a) $\mathsf{T}\mathsf{T}^\dagger\mathsf{T} = \mathsf{T}$.
(b) $\mathsf{T}^\dagger\mathsf{T}\mathsf{T}^\dagger = \mathsf{T}^\dagger$.
(c) Both $\mathsf{T}^\dagger\mathsf{T}$ and $\mathsf{T}\mathsf{T}^\dagger$ are self-adjoint.

Visit `goo.gl/Dz3WQE` for a solution. The preceding three statements are called the **Penrose conditions**, and they characterize the pseudoinverse of a linear transformation as shown in Exercise 22.

22. Let V and W be finite-dimensional inner product spaces. Let $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$ and $\mathsf{U}\colon \mathsf{W} \to \mathsf{V}$ be linear transformations such that $\mathsf{TUT} = \mathsf{T}$, $\mathsf{UTU} = \mathsf{U}$, and both UT and TU are self-adjoint. Prove that $\mathsf{U} = \mathsf{T}^\dagger$.

23. State and prove a result for matrices that is analogous to the result of Exercise 21.

24. State and prove a result for matrices that is analogous to the result of Exercise 22.

25. Let V and W be finite-dimensional inner product spaces, and let $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$ be linear. Prove the following results.

(a) If T is one-to-one, then $\mathsf{T}^*\mathsf{T}$ is invertible and $\mathsf{T}^\dagger = (\mathsf{T}^*\mathsf{T})^{-1}\mathsf{T}^*$.
(b) If T is onto, then $\mathsf{T}\mathsf{T}^*$ is invertible and $\mathsf{T}^\dagger = \mathsf{T}^*(\mathsf{T}\mathsf{T}^*)^{-1}$.

26. Let V and W be finite-dimensional inner product spaces with orthonormal bases β and γ, respectively, and let T: V $\to$ W be linear. Prove that $([\mathsf{T}]_\beta^\gamma)^\dagger = [\mathsf{T}^\dagger]_\gamma^\beta$.

27. Let V and W be finite-dimensional inner product spaces, and let T: V $\to$ W be a linear transformation. Prove part (b) of the lemma to Theorem 6.30: $\mathsf{TT}^\dagger$ is the orthogonal projection of W on R(T).

6.8* BILINEAR AND QUADRATIC FORMS

There is a certain class of scalar-valued functions of two variables defined on a vector space that arises in the study of such diverse subjects as geometry and multivariable calculus. This is the class of *bilinear forms.* We study the basic properties of this class with a special emphasis on symmetric bilinear forms, and we consider some of its applications to quadratic surfaces and multivariable calculus. In this section, F denotes any field that does not have characteristic two, as defined on page 549.

Bilinear Forms

Definition. *Let* V *be a vector space over a field F. A function H from the set* V $\times$ V *of ordered pairs of vectors to F is called a* **bilinear form** *on* V *if H is linear in each variable when the other variable is held fixed; that is, H is a bilinear form on* V *if*

(a) $H(ax_1 + x_2, y) = aH(x_1, y) + H(x_2, y)$ for all $x_1, x_2, y \in$ V and $a \in F$
(b) $H(x, ay_1 + y_2) = aH(x, y_1) + H(x, y_2)$ for all $x, y_1, y_2 \in$ V and $a \in F$.

We denote the set of all bilinear forms on V by $\mathcal{B}$(V). Observe that an inner product on a vector space is a bilinear form if the underlying field is real, but not if the underlying field is complex.

Example 1

Define a function H: R^2 $\times$ R^2 $\to R$ by

$$H\left(\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}\right) = 2a_1b_1 + 3a_1b_2 + 4a_2b_1 - a_2b_2 \quad \text{for } \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \in \mathsf{R}^2.$$

We could verify directly that H is a bilinear form on R^2. However, it is more enlightening and less tedious to observe that if

$$A = \begin{pmatrix} 2 & 3 \\ 4 & -1 \end{pmatrix}, \quad x = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, \quad \text{and} \quad y = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix},$$

then

$$H(x, y) = x^t A y.$$

The bilinearity of H now follows directly from the distributive property of matrix multiplication over matrix addition. ◆

The matrix J_{pr} acts as a *canonical form* for the theory of real symmetric matrices. The next corollary, whose proof is contained in the proof of Corollary 2, describes the role of J_{pr}.

Corollary 3. *A real symmetric $n \times n$ matrix A has index p and rank r if and only if A is congruent to J_{pr} (as just defined).*

Example 12

Let

$$A = \begin{pmatrix} 1 & 1 & -3 \\ -1 & 2 & 1 \\ 3 & 1 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 3 & 2 \\ 1 & 2 & 1 \end{pmatrix}, \quad \text{and} \quad C = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 1 & 2 & 1 \end{pmatrix}.$$

We apply Corollary 2 to determine which pairs of the matrices A, B, and C are congruent.

The matrix A is the 3×3 matrix of Example 6, where it is shown that A is congruent to a diagonal matrix with diagonal entries 1, 1, and -24. Therefore A has rank 3 and index 2. Using the methods of Example 6 (it is not necessary to compute Q), it can be shown that B and C are congruent, respectively, to the diagonal matrices

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -4 \end{pmatrix}.$$

It follows that both A and C have rank 3 and index 2, while B has rank 3 and index 1. We conclude that A and C are congruent but that B is congruent to neither A nor C. ◆

EXERCISES

1. Label the following statements as true or false.
   (a) Every quadratic form is a bilinear form.
   (b) If two matrices are congruent, they have the same eigenvalues.
   (c) Symmetric bilinear forms have symmetric matrix representations.
   (d) Any symmetric matrix is congruent to a diagonal matrix.
   (e) The sum of two symmetric bilinear forms is a symmetric bilinear form.
   (f) If two symmetric matrices over a field not of characteristic two have the the same characteristic polynomial, then they are matrix representations of the same bilinear form.
   (g) There exists a bilinear form H such that $H(x, y) \neq 0$ for all x and y.

(h) If V is a vector space of dimension n, then $\dim(\mathcal{B}(V)) = 2n$.
(i) Let H be a bilinear form on a finite-dimensional vector space V with $\dim(\mathsf{V}) > 1$. For any $x \in \mathsf{V}$, there exists $y \in \mathsf{V}$ such that $y \neq 0$, but $H(x, y) = 0$.
(j) If H is any bilinear form on a finite-dimensional real inner product space V, then there exists an ordered basis β for V such that $\psi_\beta(H)$ is a diagonal matrix.

2. Prove properties 1, 2, 3, and 4 on page 420.

3. (a) Prove that the sum of two bilinear forms is a bilinear form.
(b) Prove that the product of a scalar and a bilinear form is a bilinear form.
(c) Prove Theorem 6.31.

4. Determine which of the mappings that follow are bilinear forms. Justify your answers.

(a) Let $\mathsf{V} = \mathsf{C}[0, 1]$ be the space of continuous real-valued functions on the closed interval $[0, 1]$. For $f, g \in \mathsf{V}$, define

$$H(f, g) = \int_0^1 f(t)g(t)dt.$$

(b) Let V be a vector space over F, and let $J \in \mathcal{B}(\mathsf{V})$ be nonzero. Define $H\colon \mathsf{V} \times \mathsf{V} \to F$ by

$$H(x, y) = [J(x, y)]^2 \quad \text{for all } x, y \in \mathsf{V}.$$

(c) Define $H\colon R \times R \to R$ by $H(t_1, t_2) = t_1 + 2t_2$.
(d) Consider the vectors of R^2 as column vectors, and let $H\colon \mathsf{R}^2 \to R$ be the function defined by $H(x, y) = \det(x, y)$, the determinant of the 2×2 matrix with columns x and y.
(e) Let V be a real inner product space, and let $H\colon \mathsf{V} \times \mathsf{V} \to R$ be the function defined by $H(x, y) = \langle x, y\rangle$ for $x, y \in \mathsf{V}$.
(f) Let V be a complex inner product space, and let $H\colon \mathsf{V} \times \mathsf{V} \to C$ be the function defined by $H(x, y) = \langle x, y\rangle$ for $x, y \in \mathsf{V}$.

5. Verify that each of the given mappings is a bilinear form. Then compute its matrix representation with respect to the given ordered basis β.

(a) $H\colon \mathsf{R}^3 \times \mathsf{R}^3 \to R$, where

$$H\left(\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}, \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}\right) = a_1b_1 - 2a_1b_2 + a_2b_1 - a_3b_3$$

and

$$\beta = \left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\}.$$

(b) Let $\mathsf{V} = \mathsf{M}_{2\times 2}(R)$ and

$$\beta = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}.$$

Define $H\colon \mathsf{V} \times \mathsf{V} \to R$ by $H(A, B) = \operatorname{tr}(A) \cdot \operatorname{tr}(B)$.

(c) Let $\beta = \{\cos t, \sin t, \cos 2t, \sin 2t\}$. Then β is an ordered basis for $\mathsf{V} = \operatorname{span}(\beta)$, a four-dimensional subspace of the space of all continuous functions on R. Let $H\colon \mathsf{V} \times \mathsf{V} \to R$ be the function defined by $H(f, g) = f'(0) \cdot g''(0)$.

6. Let $H\colon \mathsf{R}^2 \times \mathsf{R}^2 \to R$ be the function defined by

$$H\left(\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \right) = a_1 b_2 + a_2 b_1 \quad \text{for } \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \in \mathsf{R}^2.$$

(a) Prove that H is a bilinear form.
(b) Find the 2×2 matrix A such that $H(x, y) = x^t A y$ for all $x, y \in \mathsf{R}^2$.

7. Let V and W be vector spaces over the same field, and let $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$ be a linear transformation. For any $H \in \mathcal{B}(\mathsf{W})$, define $\widehat{\mathsf{T}}(H)\colon \mathsf{V} \times \mathsf{V} \to F$ by $\widehat{\mathsf{T}}(H)(x, y) = H(\mathsf{T}(x), \mathsf{T}(y))$ for all $x, y \in \mathsf{V}$. Prove the following results.

(a) If $H \in \mathcal{B}(\mathsf{W})$, then $\widehat{\mathsf{T}}(H) \in \mathcal{B}(\mathsf{V})$.
(b) $\widehat{\mathsf{T}}\colon \mathcal{B}(\mathsf{W}) \to \mathcal{B}(\mathsf{V})$ is a linear transformation.
(c) If T is an isomorphism, then so is $\widehat{\mathsf{T}}$.

8. Assume the notation of Theorem 6.32.

(a) Prove that for any ordered basis β, ψ_β is linear.
(b) Let β be an ordered basis for an n-dimensional space V over F, and let $\phi_\beta\colon \mathsf{V} \to \mathsf{F}^n$ be the standard representation of V with respect to β. For $A \in \mathsf{M}_{n\times n}(F)$, define $H\colon \mathsf{V} \times \mathsf{V} \to F$ by $H(x, y) = [\phi_\beta(x)]^t A [\phi_\beta(y)]$. Prove that $H \in \mathcal{B}(\mathsf{V})$. Can you establish this as a corollary to Exercise 7?
(c) Prove the converse of (b): Let H be a bilinear form on V. If $A = \psi_\beta(H)$, then $H(x, y) = [\phi_\beta(x)]^t A [\phi_\beta(y)]$.

9. **(a)** Prove Corollary 1 to Theorem 6.32.
(b) For a finite-dimensional vector space V, describe a method for finding an ordered basis for $\mathcal{B}(\mathsf{V})$.

10. Prove Corollary 2 to Theorem 6.32.

11. Prove Corollary 3 to Theorem 6.32.

12. Prove that the relation of congruence is an equivalence relation.

13. Use Corollary 2 to Theorem 6.32 and Theorem 2.22(b) to obtain an alternate proof to Theorem 6.33.

14. Let V be a finite-dimensional vector space and $H \in \mathcal{B}(\mathsf{V})$. Prove that, for any ordered bases β and γ of V, $\operatorname{rank}(\psi_\beta(H)) = \operatorname{rank}(\psi_\gamma(H))$.

15. Prove the following results.
    (a) Any square diagonal matrix is symmetric.
    (b) Any matrix congruent to a diagonal matrix is symmetric.
    (c) the corollary to Theorem 6.35

16. Let V be a vector space over a field F not of characteristic two, and let H be a symmetric bilinear form on V. Prove that if $K(x) = H(x,x)$ is the quadratic form associated with H, then, for all $x, y \in \mathsf{V}$,

$$H(x,y) = \frac{1}{2}[K(x+y) - K(x) - K(y)].$$

17. For each of the given quadratic forms K on a real inner product space V, find a symmetric bilinear form H such that $K(x) = H(x,x)$ for all $x \in \mathsf{V}$. Then find an orthonormal basis β for V such that $\psi_\beta(H)$ is a diagonal matrix.
    (a) $K\colon \mathsf{R}^2 \to R$ defined by $K\begin{pmatrix} t_1 \\ t_2 \end{pmatrix} = -2t_1^2 + 4t_1t_2 + t_2^2$
    (b) $K\colon \mathsf{R}^2 \to R$ defined by $K\begin{pmatrix} t_1 \\ t_2 \end{pmatrix} = 7t_1^2 - 8t_1t_2 + t_2^2$
    (c) $K\colon \mathsf{R}^3 \to R$ defined by $K\begin{pmatrix} t_1 \\ t_2 \\ t_3 \end{pmatrix} = 3t_1^2 + 3t_2^2 + 3t_3^2 - 2t_1t_3$

18. Let S be the set of all $(t_1, t_2, t_3) \in \mathsf{R}^3$ for which

$$3t_1^2 + 3t_2^2 + 3t_3^2 - 2t_1t_3 + 2\sqrt{2}(t_1 + t_3) + 1 = 0.$$

Find an orthonormal basis β for R^3 for which the equation relating the coordinates of points of S relative to β is simpler. Describe S geometrically.

19. Prove the following refinement of Theorem 6.37(d).

(a) If $0 < \text{rank}(A) < n$ and A has no negative eigenvalues, then f has no local maximum at p.
(b) If $0 < \text{rank}(A) < n$ and A has no positive eigenvalues, then f has no local minimum at p.

20. Prove the following variation of the second-derivative test for the case $n = 2$: Define

$$D = \left[\frac{\partial^2 f(p)}{\partial t_1^2}\right]\left[\frac{\partial^2 f(p)}{\partial t_2^2}\right] - \left[\frac{\partial^2 f(p)}{\partial t_1 \partial t_2}\right]^2.$$

(a) If $D > 0$ and $\partial^2 f(p)/\partial t_1^2 > 0$, then f has a local minimum at p.
(b) If $D > 0$ and $\partial^2 f(p)/\partial t_1^2 < 0$, then f has a local maximum at p.
(c) If $D < 0$, then f has no local extremum at p.
(d) If $D = 0$, then the test is inconclusive.

Hint: Observe that, as in Theorem 6.37, $D = \det(A) = \lambda_1\lambda_2$, where λ_1 and λ_2 are the eigenvalues of A.

21. Let A and E be in $\mathsf{M}_{n\times n}(F)$, with E an elementary matrix. In Section 3.1, it was shown that AE can be obtained from A by means of an elementary column operation. Prove that E^tA can be obtained by means of the same elementary operation performed on the rows rather than on the columns of A. *Hint:* Note that $E^tA = (A^tE)^t$.

22. For each of the following matrices A with entries from R, find a diagonal matrix D and an invertible matrix Q such that $Q^tAQ = D$.

(a) $\begin{pmatrix} 1 & 3 \\ 3 & 2 \end{pmatrix}$ (b) $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ (c) $\begin{pmatrix} 3 & 1 & 2 \\ 1 & 4 & 0 \\ 2 & 0 & -1 \end{pmatrix}$

Hint for (b): Use an elementary operation other than interchanging columns.

23. Prove that if the diagonal entries of a diagonal matrix are permuted, then the resulting diagonal matrix is congruent to the original one.

24. Let T be a linear operator on a real inner product space V, and define $H: \mathsf{V} \times \mathsf{V} \to R$ by $H(x, y) = \langle x, \mathsf{T}(y)\rangle$ for all $x, y \in \mathsf{V}$.

(a) Prove that H is a bilinear form.
(b) Prove that H is symmetric if and only if T is self-adjoint.
(c) What properties must T have for H to be an inner product on V?
(d) Explain why H may fail to be a bilinear form if V is a complex inner product space.

25. Prove the converse to Exercise 24(a): Let V be a finite-dimensional real inner product space, and let H be a bilinear form on V. Then there exists a unique linear operator T on V such that $H(x, y) = \langle x, \mathsf{T}(y) \rangle$ for all $x, y \in \mathsf{V}$. *Hint:* Choose an orthonormal basis β for V, let $A = \psi_\beta(H)$, and let T be the linear operator on V such that $[\mathsf{T}]_\beta = A$. Visit `goo.gl/bGAfSy` for a solution.

26. Prove that the number of distinct equivalence classes of congruent $n \times n$ real symmetric matrices is

$$\frac{(n+1)(n+2)}{2}.$$

6.9* EINSTEIN'S SPECIAL THEORY OF RELATIVITY

As a consequence of physical experiments performed in the latter half of the nineteenth century (most notably the Michelson–Morley experiment of 1887), physicists concluded that *the results obtained in measuring the speed of light c are independent of the velocity of the instrument used to measure it.* For example, suppose that while on Earth, an experimenter measures the speed of light emitted from the sun and finds it to be 186,000 miles per second. Now suppose that the experimenter places the measuring equipment in a spaceship that leaves Earth traveling at 100,000 miles per second in a direction away from the sun. A repetition of the same experiment from the spaceship yields the same result: Light is traveling at 186,000 miles per second relative to the spaceship, rather than 86,000 miles per second as one might expect!

This revelation led to a new way of relating coordinate systems used to locate events in space–time. The result was Albert Einstein's *special theory of relativity.* In this section, we develop via a linear algebra viewpoint the essence of Einstein's theory.

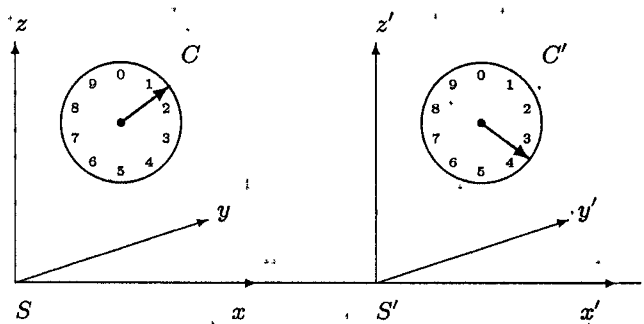


Figure 6.8

The basic problem is to compare two different inertial (nonaccelerating) coordinate systems S and S' that are in motion relative to each other under

Thus

$$\begin{pmatrix} \dfrac{1}{\sqrt{1-v^2}} & 0 & \dfrac{-v}{\sqrt{1-v^2}} \\ 0 & 1 & 0 \\ \dfrac{-v}{\sqrt{1-v^2}} & 0 & \dfrac{1}{\sqrt{1-v^2}} \end{pmatrix} \begin{pmatrix} vt \\ 0 \\ t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ t' \end{pmatrix}.$$

From the preceding equation, we obtain $\dfrac{-v^2t}{\sqrt{1-v^2}} + \dfrac{t}{\sqrt{1-v^2}} = t'$, or

$$t' = t\sqrt{1-v^2}. \tag{26}$$

This is the desired result.

A dramatic consequence of time contraction is that distances are contracted along the line of motion (see Exercise 7).

Let us make one additional point. Suppose that we choose units of distance and time commonly used in the study of motion, such as the mile, the kilometer, and the second. Recall that the velocity v we have been using is actually the ratio of the velocity using these units with the speed of light c, using the same units. For this reason, we can replace v in any of the equations given in this section with the ratio v/c, where v and c are given using the same units of measurement. Thus, for example, given a set of units of distance and time, (26) becomes

$$t' = t\sqrt{1 - \frac{v^2}{c^2}}.$$

EXERCISES

1. Complete the proof of Theorem 6.40 for the case $t < 0$.

2. For

$$w_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \quad \text{and} \quad w_2 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix},$$

show that

(a) $\{w_1, w_2\}$ is an orthogonal basis for $\text{span}(\{e_1, e_3\})$;
(b) $\text{span}(\{e_1, e_3\})$ is $\mathsf{T}_v^*\mathsf{L}_A\mathsf{T}_v$-invariant.

3. Derive (24), and prove that

$$\mathsf{T}_v \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \dfrac{-v}{\sqrt{1-v^2}} \\ 0 \\ \dfrac{1}{\sqrt{1-v^2}} \end{pmatrix}. \tag{25}$$

Hint: Use a technique similar to the derivation of (22).

4. Consider three coordinate systems S, S', and S'' with the corresponding axes (x,x',x''; y,y',y''; and z,z',z'') parallel and such that the x-, x'-, and x''-axes coincide. Suppose that S' is moving past S at a velocity $v_1 > 0$ (as measured on S), S'' is moving past S' at a velocity $v_2 > 0$ (as measured on S'), and S'' is moving past S at a velocity $v_3 > 0$ (as measured on S), and that there are three clocks C, C', and C'' such that C is stationary relative to S, C' is stationary relative to S', and C'' is stationary relative to S''. Suppose that when measured on any of the three clocks, all the origins of S, S', and S'' coincide at time 0. Assuming that $\mathsf{T}_{v_3} = \mathsf{T}_{v_2}\mathsf{T}_{v_1}$ (i.e., $B_{v_3} = B_{v_2}B_{v_1}$), prove that

$$v_3 = \frac{v_1 + v_2}{1 + v_1v_2}.$$

Note that substituting $v_2 = 1$ in this equation yields $v_3 = 1$. This tells us that the speed of light as measured in S or S' is the same. Why would we be surprised if this were not the case?

$\boxed{5.}$ Compute $(B_v)^{-1}$. Show $(B_v)^{-1} = B_{(-v)}$. Conclude that if S' moves at a negative velocity v relative to S, then $[\mathsf{T}_v]_\beta = B_v$, where B_v is of the form given in Theorem 6.43. Visit goo.gl/9gWNYu for a solution.

6. Suppose that an astronaut left Earth in the year 2000 and traveled to a star 99 light years away from Earth at 99% of the speed of light and that upon reaching the star immediately turned around and returned to Earth at the same speed. Assuming Einstein's special theory of relativity, show that if the astronaut was 20 years old at the time of departure, then he or she would return to Earth at age 48.2 in the year 2200. Explain the use of Exercise 4 in solving this problem.

7. Recall the moving space vehicle considered in the study of time contraction. Suppose that the vehicle is moving toward a fixed star located on the x-axis of S at a distance b units from the origin of S. If the space vehicle moves toward the star at velocity v, Earthlings (who remain "almost" stationary relative to S) compute the time it takes for the vehicle to reach the star as $t = b/v$. Due to the phenomenon of time contraction, the astronaut perceives a time span of $t' = t\sqrt{1-v^2} = (b/v)\sqrt{1-v^2}$. A paradox appears in that the astronaut perceives a time span inconsistent with a distance of b and a velocity of v. The paradox is resolved by observing that the distance from the solar system to the star as measured by the astronaut is less than b.

Assuming that the coordinate systems S and S' and clocks C and C' are as in the discussion of time contraction, prove the following results.

(a) At time t (as measured on C), the space–time coordinates of the star relative to S and C are

$$\begin{pmatrix} b \\ 0 \\ t \end{pmatrix}.$$

(b) At time t (as measured on C), the space–time coordinates of the star relative to S' and C' are

$$\begin{pmatrix} \dfrac{b - vt}{\sqrt{1-v^2}} \\ 0 \\ \dfrac{t - bv}{\sqrt{1-v^2}} \end{pmatrix}.$$

(c) For

$$x' = \frac{b - tv}{\sqrt{1-v^2}} \quad \text{and} \quad t' = \frac{t - bv}{\sqrt{1-v^2}},$$

we have $x' = b\sqrt{1-v^2} - t'v$.

This result may be interpreted to mean that at time t' as measured by the astronaut, the distance from the astronaut to the star, as measured by the astronaut, (see Figure 6.9) is

$$b\sqrt{1-v^2} - t'v.$$

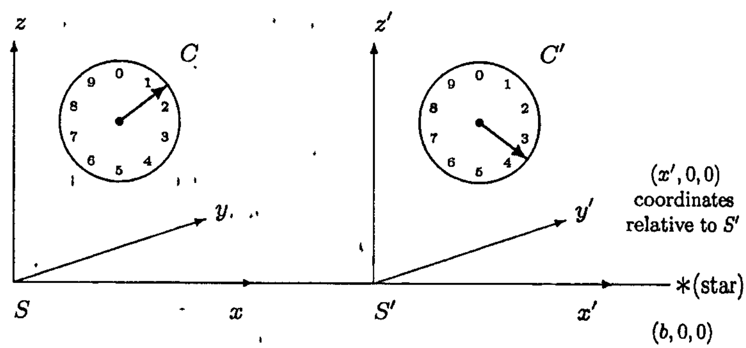


Figure 6.9

(d) Conclude from the preceding equation that
   (1) the speed of the space vehicle relative to the star, as measured by the astronaut, is v;

(2) the distance from Earth to the star, as measured by the astronaut, is $b\sqrt{1-v^2}$.

Thus distances along the line of motion of the space vehicle appear to be contracted by a factor of $\sqrt{1-v^2}$.

6.10* CONDITIONING AND THE RAYLEIGH QUOTIENT

In Section 3.4, we studied specific techniques that allow us to solve systems of linear equations in the form $Ax = b$, where A is an $m \times n$ matrix and b is an $m \times 1$ vector. Such systems often arise in applications to the real world. The coefficients in the system are frequently obtained from experimental data, and, in many cases, both m and n are so large that a computer must be used in the calculation of the solution. Thus two types of errors must be considered. First, experimental errors arise in the collection of data since no instruments can provide completely accurate measurements. Second, computers introduce roundoff errors. One might intuitively feel that small relative changes in the coefficients of the system cause small relative errors in the solution. A system that has this property is called **well-conditioned**; otherwise, the system is called **ill-conditioned**.

We now consider several examples of these types of errors, concentrating primarily on changes in b rather than on changes in the entries of A. In addition, we assume that A is a square, complex (or real), invertible matrix since this is the case most frequently encountered in applications.

Example 1

Consider the system

$$\begin{aligned} x_1 + x_2 &= 5 \\ x_1 - x_2 &= 1. \end{aligned}$$

The solution to this system is

$$\begin{pmatrix} 3 \\ 2 \end{pmatrix}.$$

Now suppose that we change the system somewhat and consider the new system

$$\begin{aligned} x_1 + x_2 &= 5 \\ x_1 - x_2 &= 1.0001. \end{aligned}$$

This modified system has the solution

$$\begin{pmatrix} 3.00005 \\ 1.99995 \end{pmatrix}.$$

We see that a change of 10^{-4} in one coefficient has caused a change of less than 10^{-4} in each coordinate of the new solution. More generally, the system

$$\begin{aligned} x_1 + x_2 &= 5 \\ x_1 - x_2 &= 1 + \delta \end{aligned}$$

(b) *If* λ_1 *and* λ_n *are the largest and smallest eigenvalues, respectively, of* A^*A, *then* $\text{cond}(A) = \sqrt{\lambda_1/\lambda_n}$.

Proof. Statement (a) follows from the previous inequalities, and (b) follows from Corollaries 1 and 2 to Theorem 6.44. ∎

It should be noted that the definition of cond(A) depends on how the norm of A is defined. There are many reasonable ways of defining the norm of a matrix. In fact, the only property needed to establish Theorem 6.45(a) and the two displayed inequalities preceding it is that $\|Ax\| \leq \|A\|_E \cdot \|x\|$ for all x.

It is clear from Theorem 6.45(a) that $\text{cond}(A) \geq 1$. It is left as an exercise to prove that $\text{cond}(A) = 1$ if and only if A is a scalar multiple of a unitary or orthogonal matrix. Moreover, it can be shown with some work that equality can be obtained in (a) by an appropriate choice of b and δb.

We can see immediately from (a) that if cond(A) is close to 1, then a small relative error in b forces a small relative error in x. If cond(A) is large, however, then the relative error in x may be small even though the relative error in b is large, or the relative error in x may be large even though the relative error in b is small! In short, cond(A) merely indicates the *potential* for large relative errors.

We have so far considered only errors in the vector b. If there is an error δA in the coefficient matrix of the system $Ax = b$, the situation is more complicated. For example, $A + \delta A$ may fail to be invertible. But under the appropriate assumptions, it can be shown that a bound for the relative error in x can be given in terms of cond(A). For example, Charles Cullen (Charles G. Cullen, *An Introduction to Numerical Linear Algebra*, PWS Publishing Co., Boston 1994, p. 60) shows that if $A + \delta A$ is invertible, then

$$\frac{\|\delta x\|}{\|x + \delta x\|} \leq \text{cond}(A)\frac{\|\delta A\|_E}{\|A\|_E}.$$

It should be mentioned that, in practice, one never computes cond(A) from its definition, for it would be an unnecessary waste of time to compute A^{-1} merely to determine its norm. In fact, if a computer is used to find A^{-1}, the computed inverse of A in all likelihood only approximates A^{-1}, and the error in the computed inverse is affected by the size of cond(A). So we are caught in a vicious circle! There are, however, some situations in which a usable approximation of cond(A) can be found. Thus, in most cases, the estimate of the relative error in x is based on an estimate of cond(A).

EXERCISES

1. Label the following statements as true or false.

(a) If $Ax = b$ is well-conditioned, then $\text{cond}(A)$ is small.
(b) If $\text{cond}(A)$ is large, then $Ax = b$ is ill-conditioned.
(c) If $\text{cond}(A)$ is small, then $Ax = b$ is well-conditioned.
(d) The norm of A equals the Rayleigh quotient.
(e) The norm of A always equals the largest eigenvalue of A.

2. Compute the norms of the following matrices.

(a) $\begin{pmatrix} 4 & 0 \\ 1 & 3 \end{pmatrix}$ (b) $\begin{pmatrix} 5 & 3 \\ -3 & 3 \end{pmatrix}$ (c) $\begin{pmatrix} 1 & \frac{-2}{\sqrt{3}} & 0 \\ 0 & \frac{-2}{\sqrt{3}} & 1 \\ 0 & \frac{2}{\sqrt{3}} & 1 \end{pmatrix}$

3. Prove that if B is symmetric, then $\|B\|_E$ is the largest eigenvalue of B.

4. Let A and A^{-1} be as follows:

$$A = \begin{pmatrix} 6 & 13 & -17 \\ 13 & 29 & -38 \\ -17 & -38 & 50 \end{pmatrix} \quad \text{and} \quad A^{-1} = \begin{pmatrix} 6 & -4 & 1 \\ -4 & 11 & 7 \\ -1 & 7 & 5 \end{pmatrix}.$$

The eigenvalues of A are approximately 84.74, 0.2007, and 0.0588.

(a) Approximate $\|A\|_E$, $\|A^{-1}\|_E$, and $\text{cond}(A)$. (Note Exercise 3.)
(b) Suppose that we have vectors x and $\tilde{x}$ such that $Ax = b$ and $\|b - A\tilde{x}\| \leq 0.001$. Use (a) to determine upper bounds for $\|\tilde{x} - A^{-1}b\|$ (the absolute error) and $\|\tilde{x} - A^{-1}b\|/\|A^{-1}b\|$ (the relative error).

5. Suppose that x is the actual solution of $Ax = b$ and that a computer arrives at an approximate solution $\tilde{x}$. If $\text{cond}(A) = 100$, $\|b\| = 1$, and $\|b - A\tilde{x}\| = 0.1$, obtain upper and lower bounds for $\|x - \tilde{x}\|/\|x\|$.

6. Let

$$B = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}.$$

Compute

$$R\begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix}, \quad \|B\|_E, \quad \text{and} \quad \text{cond}(B).$$

7. Let B be a symmetric matrix. Prove that $\min_{x \neq 0} R(x)$ equals the smallest eigenvalue of B.

8. Prove that if λ is an eigenvalue of AA^*, then λ is an eigenvalue of A^*A. This completes the proof of the lemma to Corollary 2 to Theorem 6.44.

9. Prove that if A is an invertible matrix and $Ax = b$, then

$$\frac{1}{\|A\|_E \cdot \|A^{-1}\|_E}\left(\frac{\|\delta b\|}{\|b\|}\right) \leq \frac{\|\delta x\|}{\|x\|}.$$

10. Prove the left inequality of (a) in Theorem 6.45.

11. Prove that $\text{cond}(A) = 1$ if and only if A is a scalar multiple of a unitary or orthogonal matrix.

$\boxed{12.}$ (a) Let A and B be square matrices that are unitarily equivalent. Prove that $\|A\|_E = \|B\|_E$.
(b) Let T be a linear operator on a finite-dimensional inner product space V. Define

$$\|\mathsf{T}\|_E = \max_{x \neq 0} \frac{\|\mathsf{T}(x)\|}{\|x\|}.$$

Prove that $\|\mathsf{T}\|_E = \|[\mathsf{T}]_\beta\|_E$, where β is any orthonormal basis for V.
(c) Let V be an infinite-dimensional inner product space with an orthonormal basis $\{v_1, v_2, \ldots\}$. Let T be the linear operator on V such that $\mathsf{T}(v_k) = kv_k$. Prove that $\|\mathsf{T}\|_E$ (defined in (b)) does not exist.

Visit goo.gl/B8Uw33 for a solution.

The next exercise assumes the definitions of *singular value* and *pseudoinverse* and the results of Section 6.7.

13. Let A be an $n \times n$ matrix of rank r with the nonzero singular values $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r$. Prove each of the following results.
(a) $\|A\|_E = \sigma_1$.
(b) $\|A^\dagger\|_E = \dfrac{1}{\sigma_r}$.
(c) If A is invertible (and hence $r = n$), then $\text{cond}(A) = \dfrac{\sigma_1}{\sigma_n}$.

6.11* THE GEOMETRY OF ORTHOGONAL OPERATORS

By Theorem 6.22 (p. 383), any rigid motion on a finite-dimensional real inner product space is the composite of an orthogonal operator and a translation. Thus, to understand the geometry of rigid motions thoroughly, we must analyze the structure of orthogonal operators. In this section, we show that any orthogonal operator on a finite-dimensional real inner product space can be described in terms of rotations and reflections.

two vectors if p is even and one vector if p is odd. For each $i = 1, 2, \ldots, r$, let $\mathsf{W}_{k+i} = \text{span}(\beta_i)$. Then, clearly, $\{\mathsf{W}_1, \mathsf{W}_2, \ldots, \mathsf{W}_k, \ldots, \mathsf{W}_{k+r}\}$ is pairwise orthogonal, and

$$\mathsf{V} = \mathsf{W}_1 \oplus \mathsf{W}_2 \oplus \cdots \oplus \mathsf{W}_k \oplus \cdots \oplus \mathsf{W}_{k+r}. \tag{27}$$

Moreover, if any β_i contains two vectors, then

$$\det(\mathsf{T}_{\mathsf{W}_{k+i}}) = \det([\mathsf{T}_{\mathsf{W}_{k+i}}]_{\beta_i}) = \det\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = 1.$$

So $\mathsf{T}_{\mathsf{W}_{k+i}}$ is a rotation, and hence $\mathsf{T}_{\mathsf{W}_j}$ is a rotation for $j < k + r$. If β_r consists of one vector, then $\dim(\mathsf{W}_{k+r}) = 1$ and

$$\det(\mathsf{T}_{\mathsf{W}_{k+r}}) = \det([\mathsf{T}_{\mathsf{W}_{k+r}}]_{\beta_r}) = \det(-1) = -1.$$

Thus $\mathsf{T}_{\mathsf{W}_{k+r}}$ is a reflection by Theorem 6.47, and we conclude that the decomposition in (27) satisfies (b). ∎

Example 2

Orthogonal Operators on a Three-Dimensional Real Inner Product Space

Let T be an orthogonal operator on a three-dimensional real inner product space V. Then, by Theorem 6.48(b), V can be decomposed into a direct sum of T-invariant orthogonal subspaces so that the restriction of T to each is either a rotation or a reflection, with at most one reflection. Let

$$\mathsf{V} = \mathsf{W}_1 \oplus \mathsf{W}_2 \oplus \cdots \oplus \mathsf{W}_m$$

be such a decomposition. Clearly, $m = 2$ or $m = 3$.

If $m = 2$, then $\mathsf{V} = \mathsf{W}_1 \oplus \mathsf{W}_2$. Without loss of generality, suppose that $\dim(\mathsf{W}_1) = 1$ and $\dim(\mathsf{W}_2) = 2$. Thus $\mathsf{T}_{\mathsf{W}_1}$ is a reflection or the identity on W_1, and $\mathsf{T}_{\mathsf{W}_2}$ is a rotation.

If $m = 3$, then $\mathsf{V} = \mathsf{W}_1 \oplus \mathsf{W}_2 \oplus \mathsf{W}_3$ and $\dim(\mathsf{W}_i) = 1$ for all i. If $\mathsf{T}_{\mathsf{W}_i}$ is not a reflection, then it is the identity on W_i. If no $\mathsf{T}_{\mathsf{W}_i}$ is a reflection, then T is the identity operator. ◆

EXERCISES

1. Label the following statements as true or false. Assume that the underlying vector spaces are one or two-dimensional real inner product spaces.
   (a) Any orthogonal operator is either a rotation or a reflection.
   (b) The composite of any two rotations is a rotation.
   (c) The identity operator is a rotation.

(d) The composite of two reflections is a reflection.
(e) Any orthogonal operator is a composite of rotations.
(f) For any orthogonal operator T, if $\det(\mathsf{T}) = -1$, then T is a reflection.
(g) Reflections always have eigenvalues.
(h) Rotations always have eigenvalues.
(i) If T is an operator on a 2-dimensional space V and W is a subspace of dimension 1 such that T is a reflection of V about $\mathsf{W}^{\perp}$, then W is the eigenspace of T corresponding to the eigenvalue $\lambda = -1$.
(j) The composite of an orthogonal operator and a translation is an orthogonal operator.

2. Prove that rotations, reflections, and composites of rotations and reflections are orthogonal operators.

3. Let

$$A = \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

(a) Prove that L_A is a reflection.
(b) Find the subspace of R^2 on which L_A acts as the identity.
(c) Prove that L_{AB} and L_{BA} are rotations.

4. For any real number ϕ, let

$$A = \begin{pmatrix} \cos\phi & \sin\phi \\ \sin\phi & -\cos\phi \end{pmatrix}.$$

(a) Prove that L_A is a reflection.
(b) Find the axis in R^2 about which L_A reflects.

5. For any real number ϕ, define $\mathsf{T}_\phi = \mathsf{L}_A$, where

$$A = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}.$$

(a) Prove that any rotation on R^2 is of the form T_ϕ for some ϕ.
(b) Prove that $\mathsf{T}_\phi \mathsf{T}_\psi = \mathsf{T}_{(\phi+\psi)}$ for any $\phi, \psi \in R$.
(c) Deduce that any two rotations on R^2 commute.

6. Prove that if T is a rotation on a 2-dimensional inner product space, then $-\mathsf{T}$ is also a rotation.

7. Prove that if T is a reflection on a 2-dimensional inner product space, then T^2 is the identity operator.

8. Prove Theorem 6.46 using the hints preceding the statement of the theorem.

9. Prove that no orthogonal operator can be both a rotation and a reflection.

10. Prove that if V is a two-dimensional real inner product space, then the composite of two reflections on V is a rotation of V.

11. Let V be a one- or a two-dimensional real inner product space. Define $\mathsf{T}\colon \mathsf{V} \to \mathsf{V}$ by $\mathsf{T}(x) = -x$. Prove that T is a rotation if and only if $\dim(\mathsf{V}) = 2$.

12. Complete the proof of the lemma to Theorem 6.47 by showing that $\mathsf{W} = \phi_\beta^{-1}(\mathsf{Z})$ satisfies the required conditions.

13. Let T be an orthogonal [unitary] operator on a finite-dimensional real [complex] inner product space V. If W is a T-invariant subspace of V, prove the following results.
    (a) $\mathsf{T_W}$ is an orthogonal [unitary] operator on W.
    (b) $\mathsf{W}^\perp$ is a T-invariant subspace of V. *Hint:* Use the fact that $\mathsf{T_W}$ is one-to-one and onto to conclude that, for any $y \in \mathsf{W}$, $\mathsf{T}^*(y) = \mathsf{T}^{-1}(y) \in \mathsf{W}$.
    (c) $\mathsf{T}_{\mathsf{W}^\perp}$ is an orthogonal [unitary] operator on W.

14. Let T be a linear operator on a finite-dimensional vector space V, where V is a direct sum of T-invariant subspaces, say, $\mathsf{V} = \mathsf{W}_1 \oplus \mathsf{W}_2 \oplus \cdots \oplus \mathsf{W}_k$. Prove that $\det(\mathsf{T}) = \det(\mathsf{T}_{\mathsf{W}_1}) \cdot \det(\mathsf{T}_{\mathsf{W}_2}) \cdot \cdots \cdot \det(\mathsf{T}_{\mathsf{W}_k})$.

15. Complete the proof of the corollary to Theorem 6.48.

$\boxed{16.}$ Let V be a real inner product space of dimension 2. For any $x, y \in \mathsf{V}$ such that $x \neq y$ and $\|x\| = \|y\| = 1$, show that there exists a unique rotation T on V such that $\mathsf{T}(x) = y$. Visit goo.gl/ahQT67 for a solution.

INDEX OF DEFINITIONS FOR CHAPTER 6

of a single cycle of length 3. If γ is such a cycle, then γ determines a single Jordan block

$$[\mathsf{T}]_\gamma = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{pmatrix},$$

which is a Jordan canonical form of T.

The end vector $h(x)$ of such a cycle must satisfy $(\mathsf{T}+\mathsf{I})^2(h(x)) \neq 0$. In any basis for K_λ, there must be a vector that satisfies this condition, or else no vector in K_λ satisfies this condition, contrary to our reasoning. Testing the vectors in β, we see that $h(x) = x^2$ is acceptable. Therefore

$$\gamma = \{(\mathsf{T}+\mathsf{I})^2(x^2), (\mathsf{T}+\mathsf{I})(x^2), x^2\} = \{2, -2x, x^2\}$$

is a Jordan canonical basis for T. ◆

In the next section, we develop a computational approach for finding a Jordan canonical form and a Jordan canonical basis. In the process, we prove that Jordan canonical forms are unique up to the order of the Jordan blocks.

Let T be a linear operator on a finite-dimensional vector space V, and suppose that the characteristic polynomial of T splits. By Theorem 5.10 (p. 277), T is diagonalizable if and only if V is the direct sum of the eigenspaces of T. If T is diagonalizable, then the eigenspaces and the generalized eigenspaces coincide.

For those familiar with the material on direct sums in Section 5.2, we conclude by stating a generalization of Theorem 5.10 for nondiagonalizable operators.

Theorem 7.8. *Let* T *be a linear operator on a finite-dimensional vector space* V *whose characteristic polynomial splits. Then* V *is the direct sum of the generalized eigenspaces of* T.

Proof. Exercise. ∎

EXERCISES

1. Label the following statements as true or false.
   (a) Eigenvectors of a linear operator T are also generalized eigenvectors of T.
   (b) It is possible for a generalized eigenvector of a linear operator T to correspond to a scalar that is not an eigenvalue of T.
   (c) Any linear operator on a finite-dimensional vector space has a Jordan canonical form.
   (d) A cycle of generalized eigenvectors is linearly independent.

(e) There is exactly one cycle of generalized eigenvectors corresponding to each eigenvalue of a linear operator on a finite-dimensional vector space.

(f) Let T be a linear operator on a finite-dimensional vector space whose characteristic polynomial splits, and let $\lambda_1, \lambda_2, \ldots, \lambda_k$ be the distinct eigenvalues of T. If, for each i, β_i is a basis for K_{λ_i}, then $\beta_1 \cup \beta_2 \cup \cdots \cup \beta_k$ is a Jordan canonical basis for T.

(g) For any Jordan block J, the operator L_J has Jordan canonical form J.

(h) Let T be a linear operator on an n-dimensional vector space whose characteristic polynomial splits. Then, for any eigenvalue λ of T, $\mathsf{K}_\lambda = \mathsf{N}((\mathsf{T} - \lambda\mathsf{I})^n)$.

2. For each matrix A, find a basis for each generalized eigenspace of L_A consisting of a union of disjoint cycles of generalized eigenvectors. Then find a Jordan canonical form J of A.

(a) $A = \begin{pmatrix} 1 & 1 \\ -1 & 3 \end{pmatrix}$ (b) $A = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}$

(c) $A = \begin{pmatrix} 11 & -4 & -5 \\ 21 & -8 & -11 \\ 3 & -1 & 0 \end{pmatrix}$ (d) $A = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 1 & -1 & 3 \end{pmatrix}$

3. For each linear operator T, find a basis for each generalized eigenspace of T consisting of a union of disjoint cycles of generalized eigenvectors. Then find a Jordan canonical form J of T.

(a) Define T on $\mathsf{P}_2(R)$ by $\mathsf{T}(f(x)) = 2f(x) - f'(x)$.

(b) V is the real vector space of functions spanned by the set of real-valued functions $\{1, t, t^2, e^t, te^t\}$, and T is the linear operator on V defined by $\mathsf{T}(f) = f'$.

(c) T is the linear operator on $\mathsf{M}_{2\times 2}(R)$ defined for all $A \in \mathsf{M}_{2\times 2}(R)$ by $\mathsf{T}(A) = BA$, where $B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$.

(d) $\mathsf{T}(A) = 2A + A^t$ for all $A \in \mathsf{M}_{2\times 2}(R)$.

$\boxed{\textbf{4.}}$ Let T be a linear operator on a vector space V, and let γ be a cycle of generalized eigenvectors that corresponds to the eigenvalue λ. Prove that $\text{span}(\gamma)$ is a T-invariant subspace of V. Visit `goo.gl/Lw4ahY` for a solution.

5. Let $\gamma_1, \gamma_2, \ldots, \gamma_p$ be cycles of generalized eigenvectors of a linear operator T corresponding to an eigenvalue λ. Prove that if the initial eigenvectors are distinct, then the cycles are disjoint.

6. Let $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$ be a linear transformation. Prove the following results.

(a) $\mathsf{N}(\mathsf{T}) = \mathsf{N}(-\mathsf{T})$.

(b) $\mathsf{N}(\mathsf{T}^k) = \mathsf{N}((-\mathsf{T})^k)$..

(c) If $\mathsf{V} = \mathsf{W}$ (so that T is a linear operator on V) and λ is an eigenvalue of T, then for any positive integer k

$$\mathsf{N}((\mathsf{T} - \lambda \mathsf{I}_\mathsf{V})^k) = \mathsf{N}((\lambda \mathsf{I}_\mathsf{V} - \mathsf{T})^k).$$

7. Let U be a linear operator on a finite-dimensional vector space V. Prove the following results.

(a) $\mathsf{N}(\mathsf{U}) \subseteq \mathsf{N}(\mathsf{U}^2) \subseteq \cdots \subseteq \mathsf{N}(\mathsf{U}^k) \subseteq \mathsf{N}(\mathsf{U}^{k+1}) \subseteq \cdots$.

(b) If $\operatorname{rank}(\mathsf{U}^m) = \operatorname{rank}(\mathsf{U}^{m+1})$ for some positive integer m, then $\operatorname{rank}(\mathsf{U}^m) = \operatorname{rank}(\mathsf{U}^k)$ for any positive integer $k \geq m$.

(c) If $\operatorname{rank}(\mathsf{U}^m) = \operatorname{rank}(\mathsf{U}^{m+1})$ for some positive integer m, then $\mathsf{N}(\mathsf{U}^m) = \mathsf{N}(\mathsf{U}^k)$ for any positive integer $k \geq m$.

(d) Let T be a linear operator on V, and let λ be an eigenvalue of T. Prove that if $\operatorname{rank}((\mathsf{T}-\lambda\mathsf{I})^m) = \operatorname{rank}((\mathsf{T}-\lambda\mathsf{I})^{m+1})$ for some integer m, then $\mathsf{K}_\lambda = \mathsf{N}((\mathsf{T} - \lambda\mathsf{I})^m)$.

(e) *Second Test for Diagonalizability.* Let T be a linear operator on V whose characteristic polynomial splits, and let $\lambda_1, \lambda_2, \ldots, \lambda_k$ be the distinct eigenvalues of T. Then T is diagonalizable if and only if $\operatorname{rank}(\mathsf{T} - \lambda_i\mathsf{I}) = \operatorname{rank}((\mathsf{T} - \lambda_i\mathsf{I})^2)$ for $1 \leq i \leq k$.

(f) Use (e) to obtain a simpler proof of Exercise 24 of Section 5.4: If T is a diagonalizable linear operator on a finite-dimensional vector space V and W is a T-invariant subspace of V, then T_W is diagonalizable.

8. Let T be a linear operator on a finite-dimensional vector space V whose characteristic polynomial splits and has Jordan canonical form J. Prove that for any nonzero scalar c, cJ is a Jordan canonical form for $c\mathsf{T}$.

9. Let T be a linear operator on a finite-dimensional vector space V whose characteristic polynomial splits.

(a) Prove Theorem 7.5(b).

(b) Suppose that β is a Jordan canonical basis for T, and let λ be an eigenvalue of T. Let $\beta' = \beta \cap \mathsf{K}_\lambda$. Prove that β' is a basis for K_λ.

10. Let T be a linear operator on a finite-dimensional vector space whose characteristic polynomial splits, and let λ be an eigenvalue of T.

(a) Suppose that γ is a basis for K_λ consisting of the union of q disjoint cycles of generalized eigenvectors. Prove that $q \leq \dim(\mathsf{E}_\lambda)$.

(b) Let β be a Jordan canonical basis for T, and suppose that $J = [\mathsf{T}]_\beta$ has q Jordan blocks with λ in the diagonal positions. Prove that $q \leq \dim(\mathsf{E}_\lambda)$.

11. Prove Corollary 2 to Theorem 7.7.

12. Let T be a linear operator on a finite-dimensional vector space V, and let λ be an eigenvalue of T with corresponding eigenspace and generalized eigenspace E_λ and K_λ, respectively. Let U be an invertible linear operator on V that commutes with T (i.e., TU = UT). Prove that $\mathsf{U}(\mathsf{E}_\lambda) = \mathsf{E}_\lambda$ and $\mathsf{U}(\mathsf{K}_\lambda) = \mathsf{K}_\lambda$.

Exercises 13 and 14 are concerned with direct sums of matrices, defined in Section 5.4 on page 318.

13. Prove Theorem 7.8.

14. Let T be a linear operator on a finite-dimensional vector space V such that the characteristic polynomial of T splits, and let $\lambda_1, \lambda_2, \ldots, \lambda_k$ be the distinct eigenvalues of T. For each i, let J_i be the Jordan canonical form of the restriction of T to K_{λ_i}. Prove that

$$J = J_1 \oplus J_2 \oplus \cdots \oplus J_k$$

is the Jordan canonical form of J.

7.2 THE JORDAN CANONICAL FORM II

For the purposes of this section, we fix a linear operator T on an n-dimensional vector space V such that the characteristic polynomial of T splits. Let $\lambda_1, \lambda_2, \ldots, \lambda_k$ be the distinct eigenvalues of T.

By Theorem 7.7 (p. 484), each generalized eigenspace K_{λ_i} contains an ordered basis β_i consisting of a union of disjoint cycles of generalized eigenvectors corresponding to λ_i. So by Theorems 7.4(b) (p. 480) and 7.5 (p. 482), the union $\beta = \displaystyle\bigcup_{i=1}^{k} \beta_i$ is a Jordan canonical basis for T. For each i, let T_i be the restriction of T to K_{λ_i}, and let $A_i = [\mathsf{T}_i]_{\beta_i}$. Then A_i is the Jordan canonical form of T_i, and

$$J = [\mathsf{T}]_\beta = \begin{pmatrix} A_1 & O & \cdots & O \\ O & A_2 & \cdots & O \\ \vdots & \vdots & & \vdots \\ O & O & \cdots & A_k \end{pmatrix}$$

is the Jordan canonical form of T. In this matrix, each O is a zero matrix of appropriate size.

In this section, we compute the matrices A_i and the bases β_i, thereby computing J and β as well. While developing a method for finding J, it becomes evident that in some sense the matrices A_i are unique.

$$C = \begin{pmatrix} 0 & -1 & -1 \\ -3 & -1 & -2 \\ 7 & 5 & 6 \end{pmatrix}, \quad \text{and} \quad D = \begin{pmatrix} 0 & 1 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{pmatrix},$$

are similar. Observe that A, B, and C have the same characteristic polynomial $-(t-1)(t-2)^2$, whereas D has $-t(t-1)(t-2)$ as its characteristic polynomial. Because similar matrices have the same characteristic polynomials, D cannot be similar to A, B, or C. Let J_A, J_B, and J_C be the Jordan canonical forms of A, B, and C, respectively, using the ordering 1, 2 for their common eigenvalues. Then (see Exercise 4)

$$J_A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}, \quad J_B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \quad \text{and} \quad J_C = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}.$$

Since $J_A = J_C$, A is similar to C. Since J_B is different from J_A and J_C, B is similar to neither A nor C. ◆

The reader should observe that any diagonal matrix is a Jordan canonical form. Thus a linear operator T on a finite-dimensional vector space V *is diagonalizable if and only if its Jordan canonical form is a diagonal matrix. Hence* T *is diagonalizable if and only if the Jordan canonical basis for* T *consists of eigenvectors of* T. Similar statements can be made about matrices. Thus, of the matrices A, B, and C in Example 5, A and C are not diagonalizable because their Jordan canonical forms are not diagonal matrices.

EXERCISES

1. Label the following statements as true or false. Assume that the characteristic polynomial of the matrix or linear operator splits.
   (a) The Jordan canonical form of a diagonal matrix is the matrix itself.
   (b) Let T be a linear operator on a finite-dimensional vector space V that has a Jordan canonical form J. If β is any basis for V, then the Jordan canonical form of $[\mathsf{T}]_\beta$ is J.
   (c) Linear operators having the same characteristic polynomial are similar.
   (d) Matrices having the same Jordan canonical form are similar.
   (e) Every matrix is similar to its Jordan canonical form.
   (f) Every linear operator with the characteristic polynomial $(-1)^n(t-\lambda)^n$ has the same Jordan canonical form.
   (g) Every linear operator on a finite-dimensional vector space has a unique Jordan canonical basis.
   (h) The dot diagrams of a linear operator on a finite-dimensional vector space are unique.

2. Let T be a linear operator on a finite-dimensional vector space V such that the characteristic polynomial of T splits. Suppose that $\lambda_1 = 2$, $\lambda_2 = 4$, and $\lambda_3 = -3$ are the distinct eigenvalues of T and that the dot diagrams for the restriction of T to K_{λ_i} $(i = 1, 2, 3)$ are as follows:

$$\begin{array}{ccc} \lambda_1 = 2 & \lambda_2 = 4 & \lambda_3 = -3 \\ \begin{array}{ccc} \bullet & \bullet & \bullet \\ \bullet & \bullet & \\ \bullet & & \end{array} & \begin{array}{cc} \bullet & \bullet \\ \bullet & \\ \bullet & \end{array} & \begin{array}{cc} \bullet & \bullet \\ & \\ & \end{array} \end{array}$$

Find the Jordan canonical form J of T.

3. Let T be a linear operator on a finite-dimensional vector space V with Jordan canonical form

$$\left(\begin{array}{ccc|cc|cc} 2 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 3 \end{array}\right).$$

(a) Find the characteristic polynomial of T.
(b) Find the dot diagram corresponding to each eigenvalue of T.
(c) For which eigenvalues λ_i, if any, does $\mathsf{E}_{\lambda_i} = \mathsf{K}_{\lambda_i}$?
(d) For each eigenvalue λ_i, find the smallest positive integer p_i for which $\mathsf{K}_{\lambda_i} = \mathsf{N}((\mathsf{T} - \lambda_i \mathsf{I})^{p_i})$.
(e) Compute the following numbers for each i, where U_i denotes the restriction of $\mathsf{T} - \lambda_i\mathsf{I}$ to K_{λ_i}.
   (i) $\operatorname{rank}(\mathsf{U}_i)$
   (ii) $\operatorname{rank}(\mathsf{U}_i^2)$
   (iii) $\operatorname{nullity}(\mathsf{U}_i)$
   (iv) $\operatorname{nullity}(\mathsf{U}_i^2)$

4. For each of the matrices A that follow, find a Jordan canonical form J and an invertible matrix Q such that $J = Q^{-1}AQ$. Notice that the matrices in (a), (b), and (c) are those used in Example 5.

(a) $A = \begin{pmatrix} -3 & 3 & -2 \\ -7 & 6 & -3 \\ 1 & -1 & 2 \end{pmatrix}$ (b) $A = \begin{pmatrix} 0 & 1 & -1 \\ -4 & 4 & -2 \\ -2 & 1 & 1 \end{pmatrix}$

(c) $A = \begin{pmatrix} 0 & -1 & -1 \\ -3 & -1 & -2 \\ 7 & 5 & 6 \end{pmatrix}$ (d) $A = \begin{pmatrix} 0 & -3 & 1 & 2 \\ -2 & 1 & -1 & 2 \\ -2 & 1 & -1 & 2 \\ -2 & -3 & 1 & 4 \end{pmatrix}$

5. For each linear operator T, find a Jordan canonical form J of T and a Jordan canonical basis β for T.

   (a) V is the real vector space of functions spanned by the set of real-valued functions $\{e^t, te^t, t^2e^t, e^{2t}\}$, and T is the linear operator on V defined by $\mathsf{T}(f) = f'$.
   (b) T is the linear operator on $\mathsf{P}_3(R)$ defined by $\mathsf{T}(f(x)) = xf''(x)$.
   (c) T is the linear operator on $\mathsf{P}_3(R)$ defined by $\mathsf{T}(f(x)) = f''(x) + 2f(x)$.
   (d) T is the linear operator on $\mathsf{M}_{2\times 2}(R)$ defined by

$$\mathsf{T}(A) = \begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix} \cdot A - A^t.$$

   (e) T is the linear operator on $\mathsf{M}_{2\times 2}(R)$ defined by

$$\mathsf{T}(A) = \begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix} \cdot (A - A^t).$$

   (f) V is the vector space of polynomial functions in two real variables x and y of degree at most 2, as defined in Example 4, and T is the linear operator on V defined by

$$\mathsf{T}(f(x,y)) = \frac{\partial}{\partial x} f(x,y) + \frac{\partial}{\partial y} f(x,y).$$

6. Let A be an $n \times n$ matrix whose characteristic polynomial splits. Prove that A and A^t have the same Jordan canonical form, and conclude that A and A^t are similar. *Hint:* For any eigenvalue λ of A and A^t and any positive integer r, show that $\text{rank}((A - \lambda I)^r) = \text{rank}((A^t - \lambda I)^r)$.

7. Let A be an $n \times n$ matrix whose characteristic polynomial splits, γ be a cycle of generalized eigenvectors corresponding to an eigenvalue λ, and W be the subspace spanned by γ. Define γ' to be the ordered set obtained from γ by reversing the order of the vectors in γ.

   (a) Prove that $[\mathsf{T}_\mathsf{W}]_{\gamma'} = ([\mathsf{T}_\mathsf{W}]_\gamma)^t$.
   (b) Let J be the Jordan canonical form of A. Use (a) to prove that J and J^t are similar.
   (c) Use (b) to prove that A and A^t are similar.

8. Let T be a linear operator on a finite-dimensional vector space, and suppose that the characteristic polynomial of T splits. Let β be a Jordan canonical basis for T.

   (a) Prove that for any nonzero scalar c, $\{cx : x \in \beta\}$ is a Jordan canonical basis for T.

(b) Suppose that γ is one of the cycles of generalized eigenvectors that forms β, and suppose that γ corresponds to the eigenvalue λ and has length greater than 1. Let x be the end vector of γ, and let y be a nonzero vector in E_λ. Let γ' be the ordered set obtained from γ by replacing x by $x + y$. Prove that γ' is a cycle of generalized eigenvectors corresponding to λ, and that if γ' replaces γ in the union that defines β, then the new union is also a Jordan canonical basis for T.

(c) Apply (b) to obtain a Jordan canonical basis for L_A, where A is the matrix given in Example 2, that is different from the basis given in the example.

9. Suppose that a dot diagram has k columns and m rows with p_j dots in column j and r_i dots in row i. Prove the following results.

(a) $m = p_1$ and $k = r_1$.

(b) $p_j = \max\{i\colon r_i \geq j\}$ for $1 \leq j \leq k$ and $r_i = \max\{j\colon p_j \geq i\}$ for $1 \leq i \leq m$. *Hint:* Use mathematical induction on m.

(c) $r_1 \geq r_2 \geq \cdots \geq r_m$.

(d) Deduce that the number of dots in each column of a dot diagram is completely determined by the number of dots in the rows.

10. Let T be a linear operator whose characteristic polynomial splits, and let λ be an eigenvalue of T.

(a) Prove that $\dim(\mathsf{K}_\lambda)$ is the sum of the lengths of all the blocks corresponding to λ in the Jordan canonical form of T.

(b) Deduce that $\mathsf{E}_\lambda = \mathsf{K}_\lambda$ if and only if all the Jordan blocks corresponding to λ are 1×1 matrices.

The following definitions are used in Exercises 11–19.

Definitions. *A linear operator* T *on a vector space* V *is called* **nilpotent** *if* $\mathsf{T}^p = \mathsf{T}_0$ *for some positive integer* p. *An* $n \times n$ *matrix* A *is called* **nilpotent** *if* $A^p = O$ *for some positive integer* p.

11. Let T be a linear operator on a finite-dimensional vector space V, and let β be an ordered basis for V. Prove that T is nilpotent if and only if $[\mathsf{T}]_\beta$ is nilpotent.

12. Prove that any square upper triangular matrix with each diagonal entry equal to zero is nilpotent.

13. Let T be a nilpotent operator on an n-dimensional vector space V, and suppose that p is the smallest positive integer for which $\mathsf{T}^p = \mathsf{T}_0$. Prove the following results.

(a) $\mathsf{N}(\mathsf{T}^i) \subseteq \mathsf{N}(\mathsf{T}^{i+1})$ for every positive integer i.

(b) There is a sequence of ordered bases $\beta_1, \beta_2, \ldots, \beta_p$ such that β_i is a basis for $\mathsf{N}(\mathsf{T}^i)$ and β_{i+1} contains β_i for $1 \leq i \leq p-1$.

(c) Let $\beta = \beta_p$ be the ordered basis for $\mathsf{N}(\mathsf{T}^p) = \mathsf{V}$ in (b). Then $[\mathsf{T}]_\beta$ is an upper triangular matrix with each diagonal entry equal to zero.

(d) The characteristic polynomial of T is $(-1)^n t^n$. Hence the characteristic polynomial of T splits, and 0 is the only eigenvalue of T.

14. Prove the converse of Exercise 13(d): If T is a linear operator on an n-dimensional vector space V and $(-1)^n t^n$ is the characteristic polynomial of T, then T is nilpotent.

15. Give an example of a linear operator T on a finite-dimensional vector space over the field of real numbers such that T is not nilpotent, but zero is the only eigenvalue of T. Characterize all such operators. Visit `goo.gl/nDjsWm` for a solution.

16. Let T be a nilpotent linear operator on a finite-dimensional vector space V. Recall from Exercise 13 that $\lambda = 0$ is the only eigenvalue of T, and hence $\mathsf{V} = \mathsf{K}_\lambda$. Let β be a Jordan canonical basis for T. Prove that for any positive integer i, if we delete from β the vectors corresponding to the last i dots in each column of a dot diagram of β, the resulting set is a basis for $\mathsf{R}(\mathsf{T}^i)$. (If a column of the dot diagram contains fewer than i dots, all the vectors associated with that column are removed from β.)

17. Let T be a linear operator on a finite-dimensional vector space V such that the characteristic polynomial of T splits, and let $\lambda_1, \lambda_2, \ldots, \lambda_k$ be the distinct eigenvalues of T. For each i, let v_i denote the unique vector in K_{λ_i} such that $x = v_1 + v_2 + \cdots + v_k$. (This unique representation is guaranteed by Theorem 7.3 (p. 479).) Define a mapping $\mathsf{S}\colon \mathsf{V} \to \mathsf{V}$ by

$$\mathsf{S}(x) = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_k v_k.$$

(a) Prove that S is a diagonalizable linear operator on V.

(b) Let $\mathsf{U} = \mathsf{T} - \mathsf{S}$. Prove that U is nilpotent and commutes with S, that is, $\mathsf{SU} = \mathsf{US}$.

18. Let T be a linear operator on a finite-dimensional vector space V over C, and let J be the Jordan canonical form of T. Let D be the diagonal matrix whose diagonal entries are the diagonal entries of J, and let $M = J - D$. Prove the following results.

(a) M is nilpotent.

(b) $MD = DM$.

(c) If p is the smallest positive integer for which $M^p = O$, then, for any positive integer $r < p$,

$$J^r = D^r + rD^{r-1}M + \frac{r(r-1)}{2!}D^{r-2}M^2 + \cdots + rDM^{r-1} + M^r,$$

and, for any positive integer $r \geq p$,

$$J^r = D^r + rD^{r-1}M + \frac{r(r-1)}{2!}D^{r-2}M^2 + \cdots + \frac{r!}{(r-p+1)!(p-1)!}D^{r-p+1}M^{p-1}.$$

19. Let $F = C$ and

$$J = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \cdots & 0 \\ 0 & 0 & \lambda & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & \lambda \end{pmatrix}$$

be the $m \times m$ Jordan block corresponding to λ, and let $N = J - \lambda I_m$. Prove the following results:

(a) $N^m = O$, and for $1 \leq r < m$,

$$N^r_{ij} = \begin{cases} 1 & \text{if } j = i + r \\ 0 & \text{otherwise.} \end{cases}$$

(b) For any integer $r \geq m$,

$$J^r = \begin{pmatrix} \lambda^r & r\lambda^{r-1} & \frac{r(r-1)}{2!}\lambda^{r-2} & \cdots & \frac{r(r-1)\cdots(r-m+2)}{(m-1)!}\lambda^{r-m+1} \\ 0 & \lambda^r & r\lambda^{r-1} & \cdots & \frac{r(r-1)\cdots(r-m+3)}{(m-2)!}\lambda^{r-m+2} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & \lambda^r \end{pmatrix}.$$

(c) $\lim\limits_{r\to\infty} J^r$ exists if and only if one of the following holds:

(i) $|\lambda| < 1$.

(ii) $\lambda = 1$ and $m = 1$.

(Note that $\lim\limits_{r\to\infty} \lambda^r$ exists under these conditions. See the discussion preceding Theorem 5.12 on page 284.) Furthermore, $\lim\limits_{r\to\infty} J^r$ is the zero matrix if condition (i) holds and is the 1×1 matrix (1) if condition (ii) holds.

(d) Prove Theorem 5.12 on page 284.

The following definition is used in Exercises 20 and 21.

Definition. *For any $A \in \mathsf{M}_{n\times n}(C)$, define the norm of A by*

$$\|A\|_m = \max\{|A_{ij}|: 1 \le i, j \le n\}.$$

20. Let $A, B \in \mathsf{M}_{n\times n}(C)$. Prove the following results.

(a) $\|A\|_m \ge 0$.
(b) $\|A\|_m = 0$ if and only if $A = O$.
(c) $\|cA\|_m = |c| \cdot \|A\|_m$ for any scalar c.
(d) $\|A + B\|_m \le \|A\|_m + \|B\|_m$.
(e) $\|AB\|_m \le n\|A\|_m\|B\|_m$.

21. Let $A \in \mathsf{M}_{n\times n}(C)$ be a transition matrix. (See Section 5.3.) Since C is an algebraically closed field, A has a Jordan canonical form J to which A is similar. Let P be an invertible matrix such that $P^{-1}AP = J$. Prove the following results.

(a) $\|A^k\|_m \le 1$ for every positive integer k.
(b) There exists a positive number c such that $\|J^k\|_m \le c$ for every positive integer k.
(c) Each Jordan block of J corresponding to the eigenvalue $\lambda = 1$ is a 1×1 matrix.
(d) $\lim_{k\to\infty} A^k$ exists if and only if 1 is the only eigenvalue of A with absolute value 1.
(e) Theorem 5.19(a), using (c) and Theorem 5.18.

The next exercise requires knowledge of absolutely convergent series as well as the definition of e^A for a matrix A. (See page 310.)

22. Use Exercise 20(d) to prove that e^A exists for every $A \in \mathsf{M}_{n\times n}(C)$.

23. Let $x' = Ax$ be a system of n linear differential equations, where x is an n-tuple of differentiable functions $x_1(t), x_2(t), \ldots, x_n(t)$ of the real variable t, and A is an $n \times n$ coefficient matrix as in Exercise 16 of Section 5.2. In contrast to that exercise, however, do not assume that A is diagonalizable, but assume that the characteristic polynomial of A splits. Let $\lambda_1, \lambda_2, \ldots, \lambda_k$ be the distinct eigenvalues of A.

(a) Prove that if u is the end vector of a cycle of generalized eigenvectors of L_A of length p and u corresponds to the eigenvalue λ_i, then for any polynomial $f(t)$ of degree less than p, the function

$$e^{\lambda_i t}[f(t)(A - \lambda_i I)^{p-1} + f'(t)(A - \lambda_i I)^{p-2} + \cdots + f^{(p-1)}(t)]u$$

is a solution to the system $x' = Ax$.

(b) Prove that the general solution to $x' = Ax$ is a sum of the functions of the form given in (a), where the vectors u are the end vectors of the distinct cycles that constitute a fixed Jordan canonical basis for L_A.

24. Use Exercise 23 to find the general solution to each of the following systems of linear equations, where x, y, and z are real-valued differentiable functions of the real variable t.

(a) $\begin{aligned} x' &= 2x + y \\ y' &= \phantom{2x+{}} 2y - z \\ z' &= \phantom{2x+2y-{}} 3z \end{aligned}$ (b) $\begin{aligned} x' &= 2x + y \\ y' &= \phantom{2x+{}} 2y + z \\ z' &= \phantom{2x+2y+{}} 2z \end{aligned}$

7.3 THE MINIMAL POLYNOMIAL

The Cayley–Hamilton theorem (Theorem 5.22 p. 315) tells us that for any linear operator T on an n-dimensional vector space, there is a polynomial $f(t)$ of degree n such that $f(\mathsf{T}) = \mathsf{T}_0$, namely, the characteristic polynomial of T. Hence there is a polynomial of least degree with this property, and this degree is at most n. If $g(t)$ is such a polynomial, we can divide $g(t)$ by its leading coefficient to obtain another polynomial $p(t)$ of the same degree with leading coefficient 1, that is, $p(t)$ is a *monic* polynomial. (See Appendix E.)

Definition. *Let T be a linear operator on a finite-dimensional vector space. A polynomial $p(t)$ is called a* **minimal polynomial** *of T if $p(t)$ is a monic polynomial of least positive degree for which $p(\mathsf{T}) = \mathsf{T}_0$.*

The preceding discussion shows that every linear operator on a finite-dimensional vector space has a minimal polynomial. The next result shows that it is unique, and hence we can speak of *the* minimal polynomial of T.

Theorem 7.12. *Let $p(t)$ be a minimal polynomial of a linear operator T on a finite-dimensional vector space V.*

(a) *For any polynomial $g(t)$, if $g(\mathsf{T}) = \mathsf{T}_0$, then $p(t)$ divides $g(t)$. In particular, $p(t)$ divides the characteristic polynomial of T.*
(b) *The minimal polynomial of T is unique.*

Proof. (a) Let $g(t)$ be a polynomial for which $g(\mathsf{T}) = \mathsf{T}_0$. By the division algorithm for polynomials (Theorem E.1 of Appendix E, p. 555), there exist polynomials $q(t)$ and $r(t)$ such that

$$g(t) = q(t)p(t) + r(t), \tag{1}$$

where $r(t)$ has degree less than the degree of $p(t)$. Substituting T into (1) and using that $g(\mathsf{T}) = p(\mathsf{T}) = \mathsf{T}_0$, we have $r(\mathsf{T}) = \mathsf{T}_0$. Since $r(t)$ has degree less than $p(t)$ and $p(t)$ is the minimal polynomial of T, $r(t)$ must be the zero polynomial. Thus (1) simplifies to $g(t) = q(t)p(t)$, proving (a).

EXERCISES

1. Label the following statements as true or false. Assume that all vector spaces are finite-dimensional.
   (a) Every linear operator T has a polynomial $p(t)$ of largest degree for which $p(\mathsf{T}) = \mathsf{T}_0$.
   (b) Every linear operator has a unique minimal polynomial.
   (c) The characteristic polynomial of a linear operator divides the minimal polynomial of that operator.
   (d) The minimal and the characteristic polynomials of any diagonalizable operator are equal.
   (e) Let T be a linear operator on an n-dimensional vector space V, $p(t)$ be the minimal polynomial of T, and $f(t)$ be the characteristic polynomial of T. Suppose that $f(t)$ splits. Then $f(t)$ divides $[p(t)]^n$.
   (f) The minimal polynomial of a linear operator always has the same degree as the characteristic polynomial of the operator.
   (g) A linear operator is diagonalizable if its minimal polynomial splits.
   (h) Let T be a linear operator on a vector space V such that V is a T-cyclic subspace of itself. Then the degree of the minimal polynomial of T equals dim(V).
   (i) Let T be a linear operator on a vector space V such that T has n distinct eigenvalues, where $n = \dim(\mathsf{V})$. Then the degree of the minimal polynomial of T equals n.

2. Find the minimal polynomial of each of the following matrices.

   (a) $\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ (b) $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$

   (c) $\begin{pmatrix} 4 & -14 & 5 \\ 1 & -4 & 2 \\ 1 & -6 & 4 \end{pmatrix}$ (d) $\begin{pmatrix} 3 & 0 & 1 \\ 2 & 2 & 2 \\ -1 & 0 & 1 \end{pmatrix}$

3. For each linear operator T on V, find the minimal polynomial of T.
   (a) $\mathsf{V} = \mathsf{R}^2$ and $\mathsf{T}(a, b) = (a + b, a - b)$
   (b) $\mathsf{V} = \mathsf{P}_2(R)$ and $\mathsf{T}(g(x)) = g'(x) + 2g(x)$
   (c) $\mathsf{V} = \mathsf{P}_2(R)$ and $\mathsf{T}(f(x)) = -xf''(x) + f'(x) + 2f(x)$
   (d) $\mathsf{V} = \mathsf{M}_{n\times n}(R)$ and $\mathsf{T}(A) = A^t$. *Hint:* Note that $\mathsf{T}^2 = \mathsf{I}$.

4. Determine which of the matrices and operators in Exercises 2 and 3 are diagonalizable.

5. Describe all linear operators T on R^2 such that T is diagonalizable and $\mathsf{T}^3 - 2\mathsf{T}^2 + \mathsf{T} = \mathsf{T}_0$.

6. Prove Theorem 7.13 and its corollary.

7. Prove the corollary to Theorem 7.14.

8. Let T be a linear operator on a finite-dimensional vector space, and let $p(t)$ be the minimal polynomial of T. Prove the following results.

   (a) T is invertible if and only if $p(0) \neq 0$.
   (b) If T is invertible and $p(t) = t^m + a_{m-1}t^{m-1} + \cdots + a_1 t + a_0$, then

$$\mathsf{T}^{-1} = -\frac{1}{a_0}\left(\mathsf{T}^{m-1} + a_{m-1}\mathsf{T}^{m-2} + \cdots + a_2\mathsf{T} + a_1\mathsf{I}\right).$$

9. Let T be a diagonalizable linear operator on a finite-dimensional vector space V. Prove that V is a T-cyclic subspace if and only if each of the eigenspaces of T is one-dimensional.

10. Let T be a linear operator on a finite-dimensional vector space V, and suppose that W is a T-invariant subspace of V. Prove that the minimal polynomial of T_W divides the minimal polynomial of T.

11. Let $g(t)$ be the auxiliary polynomial associated with a homogeneous linear differential equation with constant coefficients (as defined in Section 2.7), and let V denote the solution space of this differential equation. Prove the following results.

   (a) V is a D-invariant subspace, where D is the differentiation operator on C^∞.
   (b) The minimal polynomial of D_V (the restriction of D to V) is $g(t)$.
   (c) If the degree of $g(t)$ is n, then the characteristic polynomial of D_V is $(-1)^n g(t)$.

   Hint: Use Theorem 2.32 (p. 135) for (b) and (c).

12. Let D be the differentiation operator on $\mathsf{P}(R)$, the space of polynomials over R. Prove that there exists no polynomial $g(t)$ for which $g(\mathsf{D}) = \mathsf{T}_0$. Hence D has no minimal polynomial.

13. Let T be a linear operator on a finite-dimensional vector space, and suppose that the characteristic polynomial of T splits. Let $\lambda_1, \lambda_2, \ldots, \lambda_k$ be the distinct eigenvalues of T, and for each i let p_i be the number of rows in the largest Jordan block corresponding to λ_i in a Jordan canonical form of T. Prove that the minimal polynomial of T is

$$(t-\lambda_1)^{p_1}(t-\lambda_2)^{p_2}\cdots(t-\lambda_k)^{p_k}.$$

The following exercise requires knowledge of direct sums (see Section 5.2).

14. Let T be a linear operator on a finite-dimensional vector space V, and let W_1 and W_2 be T-invariant subspaces of V such that $V = W_1 \oplus W_2$. Suppose that $p_1(t)$ and $p_2(t)$ are the minimal polynomials of T_{W_1} and T_{W_2}, respectively. Either prove that the minimal polynomial $f(t)$ of T always equals $p_1(t)p_2(t)$ or give an example in which $f(t) \neq p_1(t)p_2(t)$.

Exercise 15 uses the following definition.

Definition. *Let* T *be a linear operator on a finite-dimensional vector space* V, *and let x be a nonzero vector in* V. *The polynomial $p(t)$ is called a* T-***annihilator*** *of x if $p(t)$ is a monic polynomial of least degree for which $p(\mathsf{T})(x) = 0$.*

$\boxed{\textbf{15.}^\dagger}$ Let T be a linear operator on a finite-dimensional vector space V, and let x be a nonzero vector in V. Prove the following results.

(a) The vector x has a unique T-annihilator.
(b) The T-annihilator of x divides any polynomial $g(t)$ for which $g(\mathsf{T}) = \mathsf{T}_0$.
(c) If $p(t)$ is the T-annihilator of x and W is the T-cyclic subspace generated by x, then $p(t)$ is the minimal polynomial of T_W, and $\dim(\mathsf{W})$ equals the degree of $p(t)$.
(d) The degree of the T-annihilator of x is 1 if and only if x is an eigenvector of T.

Visit goo.gl/8KD6Gw for a solution.

16. Let T be a linear operator on a finite-dimensional vector space V, and let W_1 be a T-invariant subspace of V. Let $x \in V$ such that $x \notin W_1$. Prove the following results.

(a) There exists a unique monic polynomial $g_1(t)$ of least positive degree such that $g_1(\mathsf{T})(x) \in W_1$.
(b) If $h(t)$ is a polynomial for which $h(\mathsf{T})(x) \in W_1$, then $g_1(t)$ divides $h(t)$.
(c) $g_1(t)$ divides the minimal and the characteristic polynomials of T.
(d) Let W_2 be a T-invariant subspace of V such that $W_2 \subseteq W_1$, and let $g_2(t)$ be the unique monic polynomial of least degree such that $g_2(\mathsf{T})(x) \in W_2$. Then $g_1(t)$ divides $g_2(t)$.

7.4* THE RATIONAL CANONICAL FORM

Until now we have used eigenvalues, eigenvectors, and generalized eigenvectors in our analysis of linear operators with characteristic polynomials that

Direct Sums*

The next theorem is a simple consequence of Theorem 7.23.

Theorem 7.25 (Primary Decomposition Theorem). *Let* T *be a linear operator on an n-dimensional vector space* V *with characteristic polynomial*

$$f(t) = (-1)^n(\phi_1(t))^{n_1}(\phi_2(t))^{n_2}\cdots(\phi_k(t))^{n_k},$$

where the $\phi_i(t)$'s $(1 \leq i \leq k)$ are distinct irreducible monic polynomials and the n_i's are positive integers. Then the following statements are true.

(a) $\mathsf{V} = \mathsf{K}_{\phi_1} \oplus \mathsf{K}_{\phi_2} \oplus \cdots \oplus \mathsf{K}_{\phi_k}$.

(b) *If T_i $(1 \leq i \leq k)$ is the restriction of* T *to K_{ϕ_i} and C_i is the rational canonical form of T_i, then $C_1 \oplus C_2 \oplus \cdots \oplus C_k$ is the rational canonical form of* T.

Proof. Exercise. ∎

The next theorem is a simple consequence of Theorem 7.17.

Theorem 7.26. *Let* T *be a linear operator on a finite-dimensional vector space* V. *Then* V *is a direct sum of* T*-cyclic subspaces C_{v_i}, where each v_i lies in K_ϕ for some irreducible monic divisor $\phi(t)$ of the characteristic polynomial of* T.

Proof. Exercise. ∎

EXERCISES

1. Label the following statements as true or false.

(a) Every rational canonical basis for a linear operator T is the union of T-cyclic bases.

(b) If a basis is the union of T-cyclic bases for a linear operator T, then it is a rational canonical basis for T.

(c) There exist square matrices having no rational canonical form.

(d) A square matrix is similar to its rational canonical form.

(e) For any linear operator T on a finite-dimensional vector space, any irreducible factor of the characteristic polynomial of T divides the minimal polynomial of T.

(f) Let $\phi(t)$ be an irreducible monic divisor of the characteristic polynomial of a linear operator T. The dots in the diagram used to compute the rational canonical form of the restriction of T to K_ϕ are in one-to-one correspondence with the vectors in a basis for K_ϕ.

(g) If a matrix has a Jordan canonical form, then its Jordan canonical form and rational canonical form are similar.

2. For each of the following matrices $A \in \mathsf{M}_{n\times n}(F)$, find the rational canonical form C of A and a matrix $Q \in \mathsf{M}_{n\times n}(F)$ such that $Q^{-1}AQ = C$.

(a) $A = \begin{pmatrix} 3 & 1 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{pmatrix} \quad F = R$ (b) $A = \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix} \quad F = R$

(c) $A = \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix} \quad F = C$

(d) $A = \begin{pmatrix} 0 & -7 & 14 & -6 \\ 1 & -4 & 6 & -3 \\ 0 & -4 & 9 & -4 \\ 0 & -4 & 11 & -5 \end{pmatrix} \quad F = R$

(e) $A = \begin{pmatrix} 0 & -4 & 12 & -7 \\ 1 & -1 & 3 & -3 \\ 0 & -1 & 6 & -4 \\ 0 & -1 & 8 & -5 \end{pmatrix} \quad F = R$

3. For each of the following linear operators T, find the elementary divisors, the rational canonical form C, and a rational canonical basis β.

(a) T is the linear operator on $\mathsf{P}_3(R)$ defined by

$$\mathsf{T}(f(x)) = f(0)x - f'(1).$$

(b) Let $S = \{\sin x, \cos x, x\sin x, x\cos x\}$, a subset of $\mathcal{F}(R, R)$, and let $\mathsf{V} = \text{span}(S)$. Define T to be the linear operator on V such that

$$\mathsf{T}(f) = f'.$$

(c) T is the linear operator on $\mathsf{M}_{2\times 2}(R)$ defined by

$$\mathsf{T}(A) = \begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix} \cdot A.$$

(d) Let $S = \{\sin x \sin y, \sin x \cos y, \cos x \sin y, \cos x \cos y\}$, a subset of $\mathcal{F}(R \times R, R)$, and let $\mathsf{V} = \text{span}(S)$. Define T to be the linear operator on V such that

$$\mathsf{T}(f)(x, y) = \frac{\partial f(x,y)}{\partial x} + \frac{\partial f(x,y)}{\partial y}.$$

4. Let T be a linear operator on a finite-dimensional vector space V with minimal polynomial $(\phi(t))^m$ for some positive integer m.

(a) Prove that $\mathsf{R}(\phi(\mathsf{T})) \subseteq \mathsf{N}((\phi(\mathsf{T}))^{m-1})$.

(b) Give an example to show that the subspaces in (a) need not be equal.

(c) Prove that the minimal polynomial of the restriction of T to $\mathsf{R}(\phi(\mathsf{T}))$ equals $(\phi(t))^{m-1}$.

5. Let T be a linear operator on a finite-dimensional vector space. Prove that the rational canonical form of T is a diagonal matrix if and only if T is diagonalizable. Visit goo.gl/tK8pru for a solution.

6. Let T be a linear operator on a finite-dimensional vector space V with characteristic polynomial $f(t) = (-1)^n \phi_1(t)\phi_2(t)$, where $\phi_1(t)$ and $\phi_2(t)$ are distinct irreducible monic polynomials and $n = \dim(\mathsf{V})$.

   (a) Prove that there exist $v_1, v_2 \in \mathsf{V}$ such that v_1 has T-annihilator $\phi_1(t)$, v_2 has T-annihilator $\phi_2(t)$, and $\beta_{v_1} \cup \beta_{v_2}$ is a basis for V.

   (b) Prove that there is a vector $v_3 \in \mathsf{V}$ with T-annihilator $\phi_1(t)\phi_2(t)$ such that β_{v_3} is a basis for V.

   (c) Describe the difference between the matrix representation of T with respect to $\beta_{v_1} \cup \beta_{v_2}$ and the matrix representation of T with respect to β_{v_3}.

   Thus, to assure the uniqueness of the rational canonical form, we require that the generators of the T-cyclic bases that constitute a rational canonical basis have T-annihilators equal to powers of irreducible monic factors of the characteristic polynomial of T.

7. Let T be a linear operator on a finite-dimensional vector space with minimal polynomial

$$f(t) = (\phi_1(t))^{m_1}(\phi_2(t))^{m_2} \cdots (\phi_k(t))^{m_k},$$

   where the $\phi_i(t)$'s are distinct irreducible monic factors of $f(t)$. Prove that for each i, m_i is the number of entries in the first column of the dot diagram for $\phi_i(t)$.

8. Let T be a linear operator on a finite-dimensional vector space V. Prove that for any irreducible polynomial $\phi(t)$, if $\phi(\mathsf{T})$ is not one-to-one, then $\phi(t)$ divides the characteristic polynomial of T. *Hint:* Apply Exercise 15 of Section 7.3.

9. Let V be a vector space and $\beta_1, \beta_2, \ldots, \beta_k$ be disjoint subsets of V whose union is a basis for V. Now suppose that $\gamma_1, \gamma_2, \ldots, \gamma_k$ are linearly independent subsets of V such that $\operatorname{span}(\gamma_i) = \operatorname{span}(\beta_i)$ for all i. Prove that $\gamma_1 \cup \gamma_2 \cup \cdots \cup \gamma_k$ is also a basis for V.

10. Let T be a linear operator on a finite-dimensional vector space, and suppose that $\phi(t)$ is an irreducible monic factor of the characteristic polynomial of T. Prove that if $\phi(t)$ is the T-annihilator of vectors x and y, then $x \in \mathsf{C}_y$ if and only if $\mathsf{C}_x = \mathsf{C}_y$.

Exercises 11 and 12 are concerned with direct sums.

11. Prove Theorem 7.25.

12. Prove Theorem 7.26.

INDEX OF DEFINITIONS FOR CHAPTER 7

Answers to Selected Exercises

CHAPTER 1

SECTION 1.1

1. Only the pairs in (b) and (c) are parallel.
2. (a) $x = (3, -2, 4) + t(-8, 9, -3)$ (c) $x = (3, 7, 2) + t(0, 0, -10)$
3. (a) $x = (2, -5, -1) + s(-2, 9, 7) + t(-5, 12, 2)$
   (c) $x = (-8, 2, 0) + s(9, 1, 0) + t(14, -7, 0)$

SECTION 1.2

1. (a) T (b) F (c) F (d) F (e) T (f) F
   (g) F (h) F (i) T (j) T (k) T
3. $M_{13} = 3$, $M_{21} = 4$, and $M_{22} = 5$
4. (a) $\begin{pmatrix} 6 & 3 & 2 \\ -4 & 3 & 9 \end{pmatrix}$ (c) $\begin{pmatrix} 8 & 20 & -12 \\ 4 & 0 & 28 \end{pmatrix}$
   (e) $2x^4 + x^3 + 2x^2 - 2x + 10$ (g) $10x^7 - 30x^4 + 40x^2 - 15x$
13. No, (VS 4) fails.
14. Yes
15. No
17. No, (VS 5) fails.
22. 2^{mn}

SECTION 1.3

1. (a) F (b) F (c) T (d) F (e) T (f) F (g) F
2. (a) $\begin{pmatrix} -4 & 5 \\ 2 & -1 \end{pmatrix}$; the trace is -5 (c) $\begin{pmatrix} -3 & 0 & 6 \\ 9 & -2 & 1 \end{pmatrix}$
   (e) $\begin{pmatrix} 1 \\ -1 \\ 3 \\ 5 \end{pmatrix}$ (g) $\begin{pmatrix} 5 & 6 & 7 \end{pmatrix}$
8. (a) Yes (c) Yes (e) No
11. No, the set is not closed under addition.
15. Yes

SECTION 1.4

1. (a) T (b) F (c) T (d) F (e) T (f) F
2. (a) $\{r(1,1,0,0)+s(-3,0,-2,1)+(5,0,4,0): r,s \in R\}$
   (c) There are no solutions.
   (e) $\{r(10,-3,1,0,0)+s(-3,2,0,1,0)+(-4,3,0,0,5): r,s \in R\}$
3. (a) Yes (c) No (e) No
4. (a) Yes (c) Yes (e) No
5. (a) Yes (c) No (e) Yes (g) Yes

SECTION 1.5

1. (a) F (b) T (c) F (d) F (e) T (f) T
2. (a) linearly dependent (c) linearly independent (e) linearly dependent
   (g) linearly dependent (i) linearly independent

7. $\left\{\begin{pmatrix}1 & 0\\0 & 0\end{pmatrix}, \begin{pmatrix}0 & 0\\0 & 1\end{pmatrix}\right\}$

11. 2^n

SECTION 1.6

1. (a) F (b) T (c) F (d) F (e) T (f) F
   (g) F (h) T (i) F (j) T (k) T (l) T
2. (a) Yes (c) Yes (e) No
3. (a) No (c) Yes (e) No
4. No
5. No

7. $\{u_1, u_2, u_5\}$

9. $(a_1,a_2,a_3,a_4) = a_1u_1 + (a_2-a_1)u_2 + (a_3-a_2)u_3 + (a_4-a_3)u_4$

10. (a) $-4x^2-x+8$ (c) $-x^3+2x^2+4x-5$

13. $\{(1,1,1)\}$

15. n^2-1

17. $\frac{1}{2}n(n-1)$

26. n

30. $\dim(\mathsf{W}_1)=3$, $\dim(\mathsf{W}_2)=2$, $\dim(\mathsf{W}_1+\mathsf{W}_2)=4$, and $\dim(\mathsf{W}_1\cap\mathsf{W}_2)=1$

SECTION 1.7

1. (a) F (b) F (c) F (d) T (e) T (f) T

CHAPTER 2

SECTION 2.1

1. (a) T (b) F (c) F (d) T (e) F (f) F (g) T (h) F

2. The nullity is 1, and the rank is 2. T is not one-to-one but is onto.

4. The nullity is 4, and the rank is 2. T is neither one-to-one nor onto.

5. The nullity is 0, and the rank is 3. T is one-to-one but not onto.

10. $\mathsf{T}(2,3) = (5,11)$. T is one-to-one. **12.** No.

SECTION 2.2

1. (a) T (b) T (c) F (d) T (e) T (f) F

2. (a) $\begin{pmatrix} 2 & -1 \\ 3 & 4 \\ 1 & 0 \end{pmatrix}$ (c) $\begin{pmatrix} 2 & 1 & -3 \end{pmatrix}$ (d) $\begin{pmatrix} 0 & 2 & 1 \\ -1 & 4 & 5 \\ 1 & 0 & 1 \end{pmatrix}$

(f) $\begin{pmatrix} 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 1 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 1 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \end{pmatrix}$ (g) $\begin{pmatrix} 1 & 0 & \cdots & 0 & 1 \end{pmatrix}$

3. $[\mathsf{T}]_\beta^\gamma = \begin{pmatrix} -\frac{1}{3} & -1 \\ 0 & 1 \\ \frac{2}{3} & 0 \end{pmatrix}$ and $[\mathsf{T}]_\alpha^\gamma = \begin{pmatrix} -\frac{7}{3} & -\frac{11}{3} \\ 2 & 3 \\ \frac{2}{3} & \frac{4}{3} \end{pmatrix}$

5. (a) $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$ (b) $\begin{pmatrix} 0 & 1 & 0 \\ 2 & 2 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix}$ (e) $\begin{pmatrix} 1 \\ -2 \\ 0 \\ 4 \end{pmatrix}$

10. $\begin{pmatrix} 1 & 1 & 0 & \cdots & 0 \\ 0 & 1 & 1 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}$

SECTION 2.3

1. (a) F (b) T (c) F (d) T (e) F (f) F
(g) F (h) F (i) T (j) T

2. (a) $A(2B+3C) = \begin{pmatrix} 20 & -9 & 18 \\ 5 & 10 & 8 \end{pmatrix}$ and $A(BD) = \begin{pmatrix} 29 \\ -26 \end{pmatrix}$

(b) $A^tB = \begin{pmatrix} 23 & 19 & 0 \\ 26 & -1 & 10 \end{pmatrix}$ and $CB = \begin{pmatrix} 27 & 7 & 9 \end{pmatrix}$

3. (a) $[\mathsf{T}]_\beta = \begin{pmatrix} 2 & 3 & 0 \\ 0 & 3 & 6 \\ 0 & 0 & 4 \end{pmatrix}$, $[\mathsf{U}]_\beta^\gamma = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & -1 & 0 \end{pmatrix}$, and $[\mathsf{UT}]_\beta^\gamma = \begin{pmatrix} 2 & 6 & 6 \\ 0 & 0 & 4 \\ 2 & 0 & -6 \end{pmatrix}$

4. (a) $\begin{pmatrix} 1 \\ -1 \\ 4 \\ 6 \end{pmatrix}$ (c) (5)

12. (a) No. (b) No.

SECTION 2.4

1. (a) F (b) T (c) F (d) F (e) T (f) F
(g) T (h) T (i) T

2. (a) No (b) No (c) Yes (d) No (e) No (f) Yes

3. (a) No (b) Yes (c) Yes (d) No

19. (a) $[\mathsf{T}]_\beta = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$

SECTION 2.5

1. (a) F (b) T (c) T (d) F (e) T

2. (a) $\begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}$ (c) $\begin{pmatrix} 3 & -1 \\ 5 & -2 \end{pmatrix}$

3. (a) $\begin{pmatrix} a_2 & b_2 & c_2 \\ a_1 & b_1 & c_1 \\ a_0 & b_0 & c_0 \end{pmatrix}$ (c) $\begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ -3 & 2 & 1 \end{pmatrix}$ (e) $\begin{pmatrix} 5 & -6 & 3 \\ 0 & 4 & -1 \\ 3 & -1 & 2 \end{pmatrix}$

4. $[\mathsf{T}]_{\beta'} = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & -3 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 8 & 13 \\ -5 & -9 \end{pmatrix}$

5. $[\mathsf{T}]_{\beta'} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix}$

6. (a) $Q = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$, $[\mathsf{L}_A]_\beta = \begin{pmatrix} 6 & 11 \\ -2 & -4 \end{pmatrix}$

(c) $Q = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 2 \end{pmatrix}$, $[\mathsf{L}_A]_\beta = \begin{pmatrix} 2 & 2 & 2 \\ -2 & -3 & -4 \\ 1 & 1 & 2 \end{pmatrix}$

7. (a) $\mathsf{T}(x, y) = \dfrac{1}{1+m^2}((1-m^2)x + 2my, 2mx + (m^2-1)y)$

SECTION 2.6

1. (a) F (b) T (c) T (d) T (e) F (f) T (g) T (h) F

2. The functions in (a), (c), (e), and (f) are linear functionals.

3. (a) $\mathsf{f}_1(x, y, z) = x - \frac{1}{2}y$, $\mathsf{f}_2(x, y, z) = \frac{1}{2}y$, and $\mathsf{f}_3(x, y, z) = -x + z$

5. The basis for V is $\{p_1(x), p_2(x)\}$, where $p_1(x) = 2 - 2x$ and $p_2(x) = -\frac{1}{2} + x$.

7. (a) $\mathsf{T}^t(\mathsf{f}) = \mathsf{g}$, where $\mathsf{g}(a + bx) = -3a - 4b$

(b) $[\mathsf{T}^t]_{\gamma^*}^{\beta^*} = \begin{pmatrix} -1 & 1 \\ -2 & 1 \end{pmatrix}$ (c) $[\mathsf{T}]_\beta^\gamma = \begin{pmatrix} -1 & -2 \\ 1 & 1 \end{pmatrix}$

SECTION 2.7

1. (a) T (b) T (c) F (d) F (e) T (f) F (g) T
2. (a) F (b) F (c) T (d) T (e) F
3. (a) $\{e^{-t}, te^{-t}\}$ (c) $\{e^{-t}, te^{-t}, e^{t}, te^{t}\}$ (e) $\{e^{-t}, e^{t}\cos 2t, e^{t}\sin 2t\}$
4. (a) $\{e^{(1+\sqrt{5})t/2}, e^{(1-\sqrt{5})t/2}\}$ (c) $\{1, e^{-4t}, e^{-2t}\}$

CHAPTER 3

SECTION 3.1

1. (a) T (b) F (c) T (d) F (e) T (f) F
   (g) T (h) F (i) T
2. Adding -2 times column 1 to column 2 transforms A into B.
3. (a) $\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$ (c) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}$

SECTION 3.2

1. (a) F (b) F (c) T (d) T (e) F (f) T
   (g) T (h) T (i) T
2. (a) 2 (c) 2 (e) 3 (g) 1
4. (a) $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$; the rank is 2.
5. (a) The rank is 2, and the inverse is $\begin{pmatrix} -1 & 2 \\ 1 & -1 \end{pmatrix}$.

   (c) The rank is 2, and so no inverse exists.

   (e) The rank is 3, and the inverse is $\begin{pmatrix} \frac{1}{6} & -\frac{1}{3} & \frac{1}{2} \\ \frac{1}{2} & 0 & -\frac{1}{2} \\ -\frac{1}{6} & \frac{1}{3} & \frac{1}{2} \end{pmatrix}$.

   (g) The rank is 4, and the inverse is $\begin{pmatrix} -51 & 15 & 7 & 12 \\ 31 & -9 & -4 & -7 \\ -10 & 3 & 1 & 2 \\ -3 & 1 & 1 & 1 \end{pmatrix}$.
6. (a) $\mathsf{T}^{-1}(ax^2 + bx + c) = -ax^2 - (4a + b)x - (10a + 2b + c)$
   (c) $\mathsf{T}^{-1}(a, b, c) = \left(\frac{1}{6}a - \frac{1}{3}b + \frac{1}{2}c, \frac{1}{2}a - \frac{1}{2}c, -\frac{1}{6} + \frac{1}{3}b + \frac{1}{2}c\right)$
   (e) $\mathsf{T}^{-1}(a, b, c) = \left(\frac{1}{2}a - b + \frac{1}{2}c\right)x^2 + \left(-\frac{1}{2}a + \frac{1}{2}c\right)x + b$
7. $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

20. (a) $\begin{pmatrix} 1 & 3 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{pmatrix}$

SECTION 3.3

1. (a) F (b) F (c) T (d) F (e) F (f) F (g) T (h) F

2. (a) $\left\{\begin{pmatrix} -3 \\ 1 \end{pmatrix}\right\}$ (c) $\left\{\begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}\right\}$

(e) $\left\{\begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 0 \\ 1 \end{pmatrix}\right\}$ (g) $\left\{\begin{pmatrix} -3 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \\ 1 \end{pmatrix}\right\}$

3. (a) $\left\{\begin{pmatrix} 5 \\ 0 \end{pmatrix} + t\begin{pmatrix} -3 \\ 1 \end{pmatrix} : t \in R\right\}$ (c) $\left\{\begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} + t\begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} : t \in R\right\}$

(e) $\left\{\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + r\begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix} + s\begin{pmatrix} 3 \\ 0 \\ 1 \\ 0 \end{pmatrix} + t\begin{pmatrix} -1 \\ 0 \\ 0 \\ 1 \end{pmatrix} : r, s, t \in R\right\}$

(g) $\left\{\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} + r\begin{pmatrix} -3 \\ 1 \\ 1 \\ 0 \end{pmatrix} + s\begin{pmatrix} 1 \\ -1 \\ 0 \\ 1 \end{pmatrix} : r, s \in R\right\}$

4. (b) (1) $A^{-1} = \begin{pmatrix} \frac{1}{3} & 0 & \frac{1}{3} \\ \frac{1}{9} & \frac{1}{3} & -\frac{2}{9} \\ -\frac{4}{9} & \frac{2}{3} & -\frac{1}{9} \end{pmatrix}$ (2) $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 3 \\ 0 \\ -2 \end{pmatrix}$

6. $\mathsf{T}^{-1}\{(1, 11)\} = \left\{\begin{pmatrix} \frac{11}{2} \\ -\frac{9}{2} \\ 0 \end{pmatrix} + t\begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} : t \in R\right\}$

7. The systems in parts (b), (c), and (d) have solutions.

11. The farmer, tailor, and carpenter must have incomes in the proportions 4 : 3 : 4.

13. There must be 7.8 units of the first commodity and 9.5 units of the second.

SECTION 3.4

1. (a) F (b) T (c) T (d) T (e) F (f) T (g) T

2. (a) $\begin{pmatrix} 4 \\ -3 \\ -1 \end{pmatrix}$ (c) $\begin{pmatrix} 2 \\ 3 \\ -2 \\ -1 \end{pmatrix}$ (e) $\left\{ \begin{pmatrix} 4 \\ 0 \\ 1 \\ 0 \end{pmatrix} + r \begin{pmatrix} 4 \\ 1 \\ 0 \\ 0 \end{pmatrix} + s \begin{pmatrix} 1 \\ 0 \\ 2 \\ 1 \end{pmatrix} : r, s \in R \right\}$

(g) $\left\{ \begin{pmatrix} -23 \\ 0 \\ 7 \\ 9 \\ 0 \end{pmatrix} + r \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + s \begin{pmatrix} -23 \\ 0 \\ 6 \\ 9 \\ 1 \end{pmatrix} : r, s \in R \right\}$

(i) $\left\{ \begin{pmatrix} 2 \\ 0 \\ 0 \\ -1 \\ 0 \end{pmatrix} + r \begin{pmatrix} 0 \\ 2 \\ 1 \\ 0 \\ 0 \end{pmatrix} + s \begin{pmatrix} 1 \\ -4 \\ 0 \\ -2 \\ 1 \end{pmatrix} : r, s \in R \right\}$

4. (a) $\left\{ \begin{pmatrix} \frac{4}{3} \\ \frac{1}{3} \\ 0 \\ 0 \end{pmatrix} + t \begin{pmatrix} 1 \\ -1 \\ 1 \\ 2 \end{pmatrix} : t \in R \right\}$ (c) There are no solutions.

5. $\begin{pmatrix} 1 & 0 & 2 & 1 & 4 \\ -1 & -1 & 3 & -2 & -7 \\ 3 & 1 & 1 & 0 & -9 \end{pmatrix}$

7. $\{u_1, u_2, u_5\}$

11. (b) $\{(1,2,1,0,0), (2,1,0,0,0), (1,0,0,1,0), (-2,0,0,0,1)\}$

13. (b) $\{(1,0,1,1,1,0), (0,2,1,1,0,0), (1,1,1,0,0,0), (-3,-2,0,0,0,1)\}$

CHAPTER 4

SECTION 4.1

1. (a) F (b) T (c) F (d) F (e) T

2. (a) 30 (c) -8

3. (a) $-10 + 15i$ (c) -24

4. (a) 19 (c) 14

SECTION 4.2

1. (a) F (b) T (c) T (d) T (e) F (f) F (g) F (h) T

3. 42 **5.** -12 **7.** -12 **9.** 22 **11.** -3

13. -8 **15.** 0 **17.** -49 **19.** $-28 - i$ **21.** 95

SECTION 4.3

1. (a) F (b) T (c) F (d) T (e) F (f) T (g) F (h) F
3. $(4, -3, 0)$ 5. $(-20, -48, -8)$ 7. $(0, -12, 16)$
24. $t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0$
26. (a) $\begin{pmatrix} A_{22} & -A_{12} \\ -A_{21} & A_{11} \end{pmatrix}$ (c) $\begin{pmatrix} 10 & 0 & 0 \\ 0 & -20 & 0 \\ 0 & 0 & -8 \end{pmatrix}$

(e) $\begin{pmatrix} -3i & 0 & 0 \\ 4 & -1+i & 0 \\ 10+16i & -5-3i & 3+3i \end{pmatrix}$ (g) $\begin{pmatrix} 18 & 28 & -6 \\ -20 & -21 & 37 \\ 48 & 14 & -16 \end{pmatrix}$

SECTION 4.4

1. (a) T (b) T (c) T (d) F (e) F (f) T
(g) T (h) F (i) T (j) T (k) T
2. (a) 22 (c) $2 - 4i$
3. (a) -12 (c) -12 (e) 22 (g) -3
4. (a) 0 (c) -49 (e) $-28 - i$ (g) 95

SECTION 4.5

1. (a) F (b) T (c) T (d) F (e) F (f) T
3. No 5. Yes 7. Yes 9. No

CHAPTER 5

SECTION 5.1

1. (a) F (b) T (c) T (d) F (e) F (f) F
(g) F (h) T (i) T (j) F (k) F
2. (a) $11,\ t^2 - 5t + 11$ (c) $-2,\ t^4 - 2t^3 + t^2 - 2$
3. (a) $[\mathsf{T}]_\beta = \begin{pmatrix} 0 & 2 \\ -1 & 0 \end{pmatrix}$, no (c) $[\mathsf{T}]_\beta = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$, yes

(e) $[\mathsf{T}]_\beta = \begin{pmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$, no

4. (a) The eigenvalues are 4 and -1, a basis of eigenvectors is
$\left\{ \begin{pmatrix} 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}$, $Q = \begin{pmatrix} 2 & 1 \\ 3 & -1 \end{pmatrix}$, and $D = \begin{pmatrix} 4 & 0 \\ 0 & -1 \end{pmatrix}$.
(c) The eigenvalues are 1 and -1, a basis of eigenvectors is
$\left\{ \begin{pmatrix} 1 \\ 1-i \end{pmatrix}, \begin{pmatrix} 1 \\ -1-i \end{pmatrix} \right\}$, $Q = \begin{pmatrix} 1 & 1 \\ 1-i & -1-i \end{pmatrix}$, and $D = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

5. (a) $\lambda = 3, 4$ $\quad \beta = \{(3,5), (1,2)\}$
(b) $\lambda = -1, 1, 2$ $\quad \beta = \{(1,2,0), (1,-1,-1), (2,0,-1)\}$
(f) $\lambda = 1, 3$ $\quad \beta = \{-2+x, -4+x^2, -8+x^3, x\}$
(h) $\lambda = -1, 1, 1, 1$ $\quad \beta = \left\{\begin{pmatrix}-1 & 0\\ 0 & 1\end{pmatrix}, \begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}, \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}, \begin{pmatrix}0 & 0\\ 1 & 0\end{pmatrix}\right\}$
(i) $\lambda = 1, 1, -1, -1$ $\quad \beta = \left\{\begin{pmatrix}1 & 0\\ 1 & 0\end{pmatrix}, \begin{pmatrix}0 & 1\\ 0 & 1\end{pmatrix}, \begin{pmatrix}-1 & 0\\ 1 & 0\end{pmatrix}, \begin{pmatrix}0 & -1\\ 0 & 1\end{pmatrix}\right\}$
(j) $\lambda = -1, 1, 5$ $\quad \beta = \left\{\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}, \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}, \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}, \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\right\}$

25. 4

SECTION 5.2

1. (a) F (b) F (c) F (d) T (e) T (f) F
(g) T (h) T (i) F

2. (a) Not diagonalizable (c) $Q = \begin{pmatrix}1 & 4\\ 1 & -3\end{pmatrix}$

(e) Not diagonalizable (g) $Q = \begin{pmatrix}1 & 1 & 1\\ 2 & -1 & 0\\ -1 & 0 & -1\end{pmatrix}$

3. (a) Not diagonalizable (c) Not diagonalizable

(d) $\beta = \{x - x^2, 1 - x - x^2, x + x^2\}$ (e) $\beta = \{(1,1), (1,-1)\}$

7. $A^n = \dfrac{1}{3}\begin{pmatrix}5^n + 2(-1)^n & 2(5^n) - 2(-1)^n\\ 5^n - (-1)^n & 2(5^n) + (-1)^n\end{pmatrix}$

15. (b) $x(t) = c_1 e^{3t}\begin{pmatrix}-2\\ 1\end{pmatrix} + c_2 e^{-2t}\begin{pmatrix}1\\ -1\end{pmatrix}$

(c) $x(t) = e^t\left[c_1\begin{pmatrix}1\\ 0\\ 0\end{pmatrix} + c_2\begin{pmatrix}0\\ 1\\ 0\end{pmatrix}\right] + c_3 e^{2t}\begin{pmatrix}1\\ 1\\ 1\end{pmatrix}$

SECTION 5.3

1. (a) T (b) T (c) F (d) F (e) T (f) T
(g) T (h) F (i) F (j) T

2. (a) $\begin{pmatrix}0 & 0\\ 0 & 0\end{pmatrix}$ (c) $\begin{pmatrix}\frac{7}{13} & \frac{7}{13}\\ \frac{6}{13} & \frac{6}{13}\end{pmatrix}$ (e) No limit exists.

(g) $\begin{pmatrix}-1 & 0 & -1\\ -4 & 1 & -2\\ 2 & 0 & 2\end{pmatrix}$ (i) No limit exists.

6. For those who arrived during the week of June 1, after one week 25% of the patients recovered, 20% were ambulatory, 41% were bedridden, and 14% died. Eventually $\frac{59}{90}$ recover and $\frac{31}{90}$ die.

7. $\frac{3}{7}$.

8. Only the matrices in (a) and (b) are regular transition matrices.

9. (a) $\begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$ (c) No limit exists.

(e) $\begin{pmatrix} 0 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ \frac{1}{2} & 0 & 1 \end{pmatrix}$ (g) $\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 1 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 1 \end{pmatrix}$

10. (a) $\begin{pmatrix} 0.225 \\ 0.441 \\ 0.334 \end{pmatrix}$ after two stages and $\begin{pmatrix} 0.20 \\ 0.60 \\ 0.20 \end{pmatrix}$ eventually

(c) $\begin{pmatrix} 0.372 \\ 0.225 \\ 0.403 \end{pmatrix}$ after two stages and $\begin{pmatrix} 0.50 \\ 0.20 \\ 0.30 \end{pmatrix}$ eventually

(e) $\begin{pmatrix} 0.329 \\ 0.334 \\ 0.337 \end{pmatrix}$ after two stages and $\begin{pmatrix} \frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{pmatrix}$ eventually

12. $\frac{9}{19}$ new, $\frac{6}{19}$ once-used, and $\frac{4}{19}$ twice-used

13. In 1995, 24% will own large cars, 34% will own intermediate-sized cars, and 42% will own small cars; the corresponding eventual percentages are 10%, 30%, and 60%.

20. $e^O = I$ and $e^I = eI$.

SECTION 5.4

1. (a) F (b) T (c) F (d) F (e) T (f) T (g) T

2. The subspaces in (a), (c), and (d) are T-invariant.

6. (a) $\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 2 \\ 2 \end{pmatrix} \right\}$ (c) $\left\{ \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right\}$

9. (a) $-t(t^2 - 3t + 3)$ (c) $1 - t$

10. (a) $t(t-1)(t^2 - 3t + 3)$ (c) $(t-1)^3(t+1)$

18. (c) $A^{-1} = \frac{1}{2}\begin{pmatrix} 2 & -2 & -4 \\ 0 & 1 & 3 \\ 0 & 0 & -2 \end{pmatrix}$

31. (a) $t^2 - 6t + 6$ (c) $-(t+1)(t^2 - 6t + 6)$

CHAPTER 6

SECTION 6.1

1. (a) T (b) T (c) F (d) F (e) F (f) F (g) F (h) T

2. $\langle x, y\rangle = 8 + 5i$, $\|x\| = \sqrt{7}$, $\|y\| = \sqrt{14}$, and $\|x + y\| = \sqrt{37}$.

3. $\langle f, g\rangle = 1$, $\|f\| = \frac{\sqrt{3}}{3}$, $\|g\| = \sqrt{\dfrac{e^2 - 1}{2}}$, and $\|f + g\| = \sqrt{\dfrac{11 + 3e^2}{6}}$.

16. (b) No

SECTION 6.2

1. (a) F (b) T (c) T (d) F (e) T (f) F (g) T

2. For each part the orthonormal basis and the Fourier coefficients are given.

(b) $\left\{\frac{\sqrt{3}}{3}(1,1,1), \frac{\sqrt{6}}{6}(-2,1,1), \frac{\sqrt{2}}{2}(0,-1,1)\right\}$; $\frac{2\sqrt{3}}{3}, -\frac{\sqrt{6}}{6}, \frac{\sqrt{2}}{2}$.

(c) $\{1, 2\sqrt{3}(x - \frac{1}{2}), 6\sqrt{5}(x^2 - x + \frac{1}{6})\}$; $\frac{3}{2}, \frac{\sqrt{3}}{6}, 0$.

(e) $\left\{\frac{1}{5}(2,-1,-2,4), \frac{1}{\sqrt{30}}(-4,2,-3,1), \frac{1}{\sqrt{155}}(-3,4,9,7)\right\}$; $10, 3\sqrt{30}, \sqrt{155}$

(g) $\left\{\frac{1}{6}\begin{pmatrix} 3 & 5 \\ -1 & 1 \end{pmatrix}, \frac{1}{6\sqrt{2}}\begin{pmatrix} -4 & 4 \\ 6 & -2 \end{pmatrix}, \frac{1}{9\sqrt{2}}\begin{pmatrix} 9 & -3 \\ 6 & -6 \end{pmatrix}\right\}$; $24, 6\sqrt{2}, -9\sqrt{2}$

(i) $\left\{\sqrt{\frac{2}{\pi}}\sin t, \sqrt{\frac{2}{\pi}}\cos t, \sqrt{\frac{\pi}{\pi^2-8}}(1 - \frac{4}{\pi}\sin t), \sqrt{\frac{12\pi}{\pi^4-96}}(t + \frac{4}{\pi}\cos t - \frac{\pi}{2})\right\}$;

$\sqrt{\frac{2}{\pi}}(2\pi + 2), -4\sqrt{\frac{2}{\pi}}, \sqrt{\frac{\pi^2-8}{\pi}}(1+\pi), \sqrt{\frac{\pi^4-96}{3\pi}}$

(k) $\left\{\frac{1}{\sqrt{47}}(-4, 3-2i, i, 1-4i), \frac{1}{\sqrt{60}}(3-i, -5i, -2+4i, 2+i),\right.$

$\left.\frac{1}{\sqrt{1160}}(-17-i, -9+8i, -18+16i, -9+8i)\right\}$;

$\sqrt{47}(-1-i), \sqrt{60}(-1+2i), \sqrt{1160}(1+i)$

(m) $\left\{\frac{1}{\sqrt{18}}\begin{pmatrix} -1+i & -i \\ 2-i & 1+3i \end{pmatrix}, \frac{1}{\sqrt{246}}\begin{pmatrix} -4i & -11-9i \\ 1+5i & 1-i \end{pmatrix},\right.$

$\left.\frac{1}{\sqrt{39063}}\begin{pmatrix} -5-118i & -7-26i \\ -145i & -58 \end{pmatrix}\right\}$; $\sqrt{18}(2+i), \sqrt{246}(-1-i), 0$

4. $S^{\perp} = \text{span}(\{(i, -\frac{1}{2}(1+i), 1)\})$

5. $S_0^{\perp}$ is the plane through the origin that is perpendicular to x_0; $S^{\perp}$ is the line through the origin that is perpendicular to the plane containing x_1 and x_2.

19. (a) $\dfrac{1}{17}\begin{pmatrix} 26 \\ 104 \end{pmatrix}$ (b) $\dfrac{1}{14}\begin{pmatrix} 29 \\ 17 \\ 40 \end{pmatrix}$

20. (b) $\dfrac{1}{\sqrt{14}}$

SECTION 6.3

1. (a) T (b) F (c) F (d) T (e) F (f) T (g) T
2. (a) $y = (1, -2, 4)$ (c) $y = 210x^2 - 204x + 33$
3. (a) $\mathsf{T}^*(x) = (11, -12)$ (c) $\mathsf{T}^*(f(t)) = 12 + 6t$
14. $\mathsf{T}^*(x) = \langle x, z\rangle\, y$
20. (a) The linear function is $y = -2t + 5/2$ with $E = 1$, and the quadratic function is $y = t^2/3 - 4t/3 + 2$ with $E = 0$.
    (b) The linear function is $y = 1.25t + 0.55$ with $E = 0.3$, and the quadratic function is $t^2/56 + 15t/14 + 239/280$ with $E = 0.22857$ (approximation).
21. The spring constant is approximately 2.1.
22. (a) $x = \frac{2}{7},\ y = \frac{3}{7},\ z = \frac{1}{7}$ (d) $x = \frac{7}{12},\ y = \frac{1}{12},\ z = \frac{1}{4},\ w = -\frac{1}{12}$

SECTION 6.4

1. (a) T (b) F (c) F (d) T (e) T (f) T (g) F (h) T
2. (a) T is self-adjoint. An orthonormal basis of eigenvectors is
   $$\left\{\frac{1}{\sqrt{5}}(1,-2),\ \frac{1}{\sqrt{5}}(2,1)\right\},$$ with corresponding eigenvalues 6 and 1.
   (c) T is normal, but not self-adjoint. An orthonormal basis of eigenvectors is
   $$\left\{\frac{1}{2}(1+i, \sqrt{2}),\ \frac{1}{2}(1+i, -\sqrt{2})\right\}$$ with corresponding eigenvalues
   $2 + \dfrac{1+i}{\sqrt{2}}$ and $2 - \dfrac{1+i}{\sqrt{2}}$.
   (e) T is self-adjoint. An orthonormal basis of eigenvectors is
   $$\left\{\frac{1}{\sqrt{2}}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix}-1 & 0\\ 0 & 1\end{pmatrix}\right\}$$
   with corresponding eigenvalues 1, 1, -1, 1.

SECTION 6.5

1. (a) T (b) F (c) F (d) T (e) F (f) T
   (g) F (h) F (i) F
2. (a) $P = \dfrac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix}$ and $D = \begin{pmatrix}3 & 0\\ 0 & -1\end{pmatrix}$

   (d) $P = \begin{pmatrix}\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}}\\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}}\\ 0 & -\frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}}\end{pmatrix}$ and $D = \begin{pmatrix}-2 & 0 & 0\\ 0 & -2 & 0\\ 0 & 0 & 4\end{pmatrix}$
4. T_z is normal for all $z \in C$, T_z is self-adjoint if and only if $z \in R$, and T_z is unitary if and only if $|z| = 1$.
5. Only the pair of matrices in (d) are unitarily equivalent.

25. $2(\psi - \phi)$

26. (a) $\psi - \dfrac{\phi}{2}$ (b) $\psi + \dfrac{\phi}{2}$

27. (a) $x = \dfrac{1}{\sqrt{2}}x' + \dfrac{1}{\sqrt{2}}y'$ and $y = \dfrac{1}{\sqrt{2}}x' - \dfrac{1}{\sqrt{2}}y'$

The new quadratic form is $3(x')^2 - (y')^2$.

(c) $x = \dfrac{3}{\sqrt{13}}x' + \dfrac{2}{\sqrt{13}}y'$ and $y = \dfrac{-2}{\sqrt{13}}x' + \dfrac{2}{\sqrt{13}}y'$

The new quadratic form is $5(x')^2 - 8(y')^2$.

29. (c) $Q = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & -\frac{\sqrt{6}}{6} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} & \frac{\sqrt{6}}{6} \\ 0 & \frac{1}{\sqrt{3}} & \frac{\sqrt{6}}{3} \end{pmatrix}$ and $R = \begin{pmatrix} \sqrt{2} & \sqrt{2} & 2\sqrt{2} \\ 0 & \sqrt{3} & \frac{\sqrt{3}}{3} \\ 0 & 0 & \frac{\sqrt{6}}{3} \end{pmatrix}$

(e) $x_1 = 3,\ x_2 = -5,\ x_3 = 4$

SECTION 6.6

1. (a) F (b) T (c) T (d) F (e) F

2. For $\mathsf{W} = \operatorname{span}(\{(1,2)\})$, $[\mathsf{T}]_\beta = \begin{pmatrix} \frac{1}{5} & \frac{2}{5} \\ \frac{2}{5} & \frac{4}{5} \end{pmatrix}$.

3. (2) (a) $\mathsf{T}_1(a,b) = \frac{1}{2}(a+b, a+b)$ and $\mathsf{T}_2(a,b) = \frac{1}{2}(a-b, -a+b)$

(d) $\mathsf{T}_1(a,b,c) = \frac{1}{3}(2a-b-c, -a+2b-c, -a-b+2c)$ and

$\mathsf{T}_2(a,b,c) = \frac{1}{3}(a+b+c, a+b+c, a+b+c)$

SECTION 6.7

1. (a) F (b) F (c) T (d) T (e) F (f) F (g) T

2. (a) $v_1 = \begin{pmatrix}1\\0\end{pmatrix}$, $v_2 = \begin{pmatrix}0\\1\end{pmatrix}$, $u_1 = \dfrac{1}{\sqrt{3}}\begin{pmatrix}1\\1\\1\end{pmatrix}$, $u_2 = \dfrac{1}{\sqrt{2}}\begin{pmatrix}0\\1\\-1\end{pmatrix}$, $u_3 = \dfrac{1}{\sqrt{6}}\begin{pmatrix}2\\-1\\-1\end{pmatrix}$

$\sigma_1 = \sqrt{3},\ \sigma_2 = \sqrt{2}$

(c) $v_1 = \dfrac{1}{\sqrt{\pi}}\sin x,\ v_2 = \dfrac{1}{\sqrt{\pi}}\cos x,\ v_3 = \dfrac{1}{\sqrt{2\pi}}$

$u_1 = \dfrac{\cos x + 2\sin x}{\sqrt{5\pi}},\ u_2 = \dfrac{2\cos x - \sin x}{\sqrt{5\pi}},\ u_3 = \dfrac{1}{\sqrt{2\pi}},$

$\sigma_1 = \sqrt{5},\ \sigma_2 = \sqrt{5},\ \sigma_3 = 2$

3. (a) $\begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \end{pmatrix} \begin{pmatrix} \sqrt{6} & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}^*$

(c) $\begin{pmatrix} \frac{2}{\sqrt{10}} & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} & -\frac{1}{\sqrt{2}} & 0 & -\frac{2}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} & \frac{1}{\sqrt{2}} & 0 & -\frac{2}{\sqrt{10}} \\ \frac{2}{\sqrt{10}} & 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{10}} \end{pmatrix} \begin{pmatrix} \sqrt{5} & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}^*$

(e) $\begin{pmatrix} \frac{1+i}{2} & \frac{1+i}{2} \\ \frac{1-i}{2} & \frac{-1+i}{2} \end{pmatrix} \begin{pmatrix} \sqrt{6} & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{2}{\sqrt{6}} & \frac{1-i}{\sqrt{6}} \\ \frac{1+i}{\sqrt{6}} & -\frac{2}{\sqrt{6}} \end{pmatrix}^*$

4. (a) $WP = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \frac{\sqrt{8}+\sqrt{2}}{2} & \frac{-\sqrt{8}+\sqrt{2}}{2} \\ \frac{-\sqrt{8}+\sqrt{2}}{2} & \frac{\sqrt{8}+\sqrt{2}}{2} \end{pmatrix}$

5. (a) $\mathsf{T}^\dagger(x,y,z) = \left(\frac{x+y+z}{3}, \frac{y-z}{2}\right)$

(c) $\mathsf{T}^\dagger(a+b\sin x+c\cos x) = \mathsf{T}^{-1}(a+b\sin x+c\cos x) =$
$\frac{a}{2} + \frac{(2b+c)\sin x + (-b+2c)\cos x}{5}$

6. (a) $\frac{1}{6}\begin{pmatrix} 1 & 1 & -1 \\ 1 & 1 & -1 \end{pmatrix}$ (c) $\frac{1}{5}\begin{pmatrix} 1 & -2 & 3 & 1 \\ 1 & 3 & -2 & 1 \end{pmatrix}$ (e) $\frac{1}{6}\begin{pmatrix} 1-i & 1+i \\ 1 & i \end{pmatrix}$

7. (a) $\mathsf{Z}_1 = \mathsf{N(T)}^\perp = \mathsf{R}^2$ and $\mathsf{Z}_2 = \mathsf{R(T)} = \text{span}\{(1,1,1),(0,1,-1)\}$
(c) $\mathsf{Z}_1 = \mathsf{N(T)}^\perp = \mathsf{V}$ and $\mathsf{Z}_2 = \mathsf{R(T)} = \mathsf{V}$

8. (a) No solution $\frac{1}{2}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$

SECTION 6.8

1. (a) F (b) F (c) T (d) F (e) T (f) F
(g) F (h) F (i) T (j) F

4. (a) Yes (b) No (c) No (d) Yes (e) Yes (f) No

5. (a) $\begin{pmatrix} 0 & 2 & -2 \\ 2 & 0 & -2 \\ 1 & 1 & 0 \end{pmatrix}$ (b) $\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}$ (c) $\begin{pmatrix} 0 & 0 & 0 & 0 \\ -1 & 0 & -4 & 0 \\ 0 & 0 & 0 & 0 \\ -2 & 0 & -8 & 0 \end{pmatrix}$

17. (a) and (b) $\left\{ \begin{pmatrix} \frac{2}{\sqrt{5}} \\ -\frac{1}{\sqrt{5}} \end{pmatrix}, \begin{pmatrix} \frac{1}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} \end{pmatrix} \right\}$ (c) $\left\{ \begin{pmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ -\frac{1}{\sqrt{2}} \end{pmatrix} \right\}$

18. Same as Exercise 17(c)

22. (a) $Q = \begin{pmatrix} 1 & -3 \\ 0 & 1 \end{pmatrix}$ and $D = \begin{pmatrix} 1 & 0 \\ 0 & -7 \end{pmatrix}$

(b) $Q = \begin{pmatrix} 1 & -\frac{1}{2} \\ 1 & \frac{1}{2} \end{pmatrix}$ and $D = \begin{pmatrix} 2 & 0 \\ 0 & -\frac{1}{2} \end{pmatrix}$

(c) $Q = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & -0.25 \\ 1 & 0 & 2 \end{pmatrix}$ and $D = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 6.75 \end{pmatrix}$

SECTION 6.9

5. $(B_v)^{-1} = \begin{pmatrix} \dfrac{1}{\sqrt{1-v^2}} & 0 & \dfrac{v}{\sqrt{1-v^2}} \\ 0 & 1 & 0 \\ \dfrac{v}{\sqrt{1-v^2}} & 0 & \dfrac{1}{\sqrt{1-v^2}} \end{pmatrix}$

SECTION 6.10

1. (a) F (b) F (c) T (d) F (e) F
2. (a) $\sqrt{18}$ (c) approximately 2.34
4. (a) $\|A\| \approx 84.74$, $\|A^{-1}\| \approx 17.01$, and $\text{cond}(A) \approx 1441$
   (b) $\|\tilde{x} - A^{-1}b\| \le \|A^{-1}\| \cdot \|A\tilde{x} - b\| \approx 0.17$ and
   $$\frac{\|\tilde{x} - A^{-1}b\|}{\|A^{-1}b\|} \le \text{cond}(A)\frac{\|b - A\tilde{x}\|}{\|b\|} \approx \frac{14.41}{\|b\|}$$
5. $0.001 \le \dfrac{\|x - \tilde{x}\|}{\|x\|} \le 10$
6. $R\begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix} = \dfrac{9}{7}$, $\|B\|_E = 2$, and $\text{cond}(B) = 2$.

SECTION 6.11

1. (a) T (b) T (c) T (d) F (e) F (f) T
   (g) T (h) F (i) T (j) F
3. (b) $\left\{t\begin{pmatrix} \sqrt{3} \\ 1 \end{pmatrix} : t \in R\right\}$
4. (b) $\left\{t\begin{pmatrix} 1 \\ 0 \end{pmatrix} : t \in R\right\}$ if $\phi = 0$ and $\left\{t\begin{pmatrix} \cos\phi + 1 \\ \sin\phi \end{pmatrix} : t \in R\right\}$ if $\phi \ne 0$

CHAPTER 7

SECTION 7.1

1. (a) T (b) F (c) F (d) T (e) F (f) F (g) T (h) T
2. (a) For $\lambda = 2$, $\left\{\begin{pmatrix} -1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}\right\}$ $J = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$

   (c) For $\lambda = -1$, $\left\{\begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix}\right\}$ For $\lambda = 2$, $\left\{\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}\right\}$ $J = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$
3. (a) For $\lambda = 2$, $\{2, -2x, x^2\}$ $J = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$

   (c) For $\lambda = 1$, $\left\{\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right\}$ $J = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$

SECTION 7.2

1. (a) T (b) T (c) F (d) T (e) T (f) F (g) F (h) T

2. $J = \begin{pmatrix} A_1 & O & O \\ O & A_2 & O \\ O & O & A_3 \end{pmatrix}$ where $A_1 = \begin{pmatrix} 2 & 1 & 0 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 \end{pmatrix}$,

$$A_2 = \begin{pmatrix} 4 & 1 & 0 & 0 \\ 0 & 4 & 1 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix} \quad \text{and} \quad A_3 = \begin{pmatrix} -3 & 0 \\ 0 & -3 \end{pmatrix}$$

3. (a) $-(t-2)^5(t-3)^2$ (b)

$$\begin{array}{cc} \lambda_1 = 2 & \lambda_2 = 3 \\ \begin{array}{cc} \bullet & \bullet \\ \bullet & \bullet \\ \bullet & \end{array} & \begin{array}{cc} \bullet & \bullet \end{array} \end{array}$$

(c) $\lambda_2 = 3$ (d) $p_1 = 3$ and $p_2 = 1$
(e) (i) $\operatorname{rank}(\mathsf{U}_1) = 3$ and $\operatorname{rank}(\mathsf{U}_2) = 0$
(ii) $\operatorname{rank}(\mathsf{U}_1^2) = 1$ and $\operatorname{rank}(\mathsf{U}_2^2) = 0$
(iii) $\operatorname{nullity}(\mathsf{U}_1) = 2$ and $\operatorname{nullity}(\mathsf{U}_2) = 2$
(iv) $\operatorname{nullity}(\mathsf{U}_1^2) = 4$ and $\operatorname{nullity}(\mathsf{U}_2^2) = 2$

4. (a) $J = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$ and $Q = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 1 & 2 \\ 1 & -1 & 0 \end{pmatrix}$

(d) $J = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}$ and $Q = \begin{pmatrix} 1 & 0 & 1 & -1 \\ 1 & -1 & 0 & 1 \\ 1 & -2 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix}$

5. (a) $J = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}$ and $\beta = \{2e^t, 2te^t, t^2e^t, e^{2t}\}$

(c) $J = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix}$ and $\beta = \{6x, x^3, 2, x^2\}$

(d) $J = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix}$ and

$$\beta = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 0 & 2 \end{pmatrix}, \begin{pmatrix} 1 & -2 \\ 2 & 0 \end{pmatrix} \right\}$$

24. (a) $\begin{pmatrix} x \\ y \\ z \end{pmatrix} = e^{2t}\left[(c_1 + c_2t)\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + c_2\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}\right] + c_3e^{3t}\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$

(b) $\begin{pmatrix} x \\ y \\ z \end{pmatrix} = e^{2t}\left[(c_1 + c_2t + c_3t^2)\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + (c_2 + 2c_3t)\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + 2c_3\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right]$

SECTION 7.3

1. (a) F (b) T (c) F (d) F (e) T (f) F
(g) F (h) T (i) T

2. (a) $(t-1)(t-3)$ (c) $(t-1)^2(t-2)$ (d) $(t-2)^2$

3. (a) $t^2 - 2$ (c) $(t-2)^2$ (d) $(t-1)(t+1)$

4. For (2), (a); for (3), (a) and (d)

5. The operators are T_0, I, and all operators having both 0 and 1 as eigenvalues.

SECTION 7.4

1. (a) T (b) F (c) F (d) T (e) T (f) F (g) T

2. (a) $\begin{pmatrix} 0 & 0 & 27 \\ 1 & 0 & -27 \\ 0 & 1 & 9 \end{pmatrix}$ (b) $\begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix}$

(c) $\begin{pmatrix} \frac{1}{2}(-1 + i\sqrt{3}) & 0 \\ 0 & \frac{1}{2}(-1 - i\sqrt{3}) \end{pmatrix}$ (e) $\begin{pmatrix} 0 & -2 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -3 \\ 0 & 0 & 1 & 0 \end{pmatrix}$

3. (a) $t^2 + 1$ and t^2 $\quad C = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$; $\quad \beta = \{1, x, -2x + x^2, -3x + x^3\}$

(c) $t^2 - t + 1$ $\quad C = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$

$$\beta = \left\{\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix}\right\}$$