Example 2 To what diagonal matrix is the following matrix A similar?

$$A = \begin{bmatrix} 3 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 4 \end{bmatrix}$$

Since A is symmetric, we know that A is similar to a diagonal matrix. To find which diagonal matrix A is similar to, we need only compute its characteristic polynomial.

$$C_A(x) = \begin{vmatrix} x-3 & -2 & -2 \\ -2 & x-2 & 0 \\ -2 & 0 & x-4 \end{vmatrix} = x^3 - 9x + 18x = x(x-3)(x-6)$$

Thus, A is similar to a diagonal matrix having diagonal entries 0, 3, and 6.

Example 3 For the following matrix A, find a matrix U such that $U^{-1}AU$ is diagonal.

$$A = \begin{bmatrix} -1 & 2 & 2 \\ 2 & -1 & 2 \\ 2 & 2 & -1 \end{bmatrix}$$

First, we compute the characteristic polynomial

$$C_A(x) = \begin{vmatrix} x+1 & -2 & -2 \\ -2 & x+1 & -2 \\ -2 & -2 & x+1 \end{vmatrix} = x^3 + 3x^2 - 9x - 27 = (x-3)(x+3)^2$$

Next, the eigenspaces corresponding to the eigenvalues 3 and -3 can be found by solving the equations

$$\begin{aligned} -4x + 2y + 2z &= 0 \\ 2x - 4y + 2z &= 0 \\ 2x + 2y - 4z &= 0 \end{aligned} \quad \text{and} \quad \begin{aligned} 2x + 2y + 2z &= 0 \\ 2x + 2y + 2z &= 0 \\ 2x + 2y + 2z &= 0 \end{aligned}$$

The solution to the first system is $x = y = z = c$, with c arbitrary. For the second system, any vector satisfying $x + y + z = 0$ is a solution. Thus, a unit eigenvector corresponding to the eigenvalue 3 is

$$\begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}$$

For the eigenvalue -3, we must choose two orthogonal unit vectors satisfying $x + y + z = 0$. We write one unit vector satisfying this equation. A second is then determined by the equation and the fact that it is orthogonal to the first.

We take as our first vector

$$\begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}$$

Then the second satisfies $x + y + z = 0$ and $x - y = 0$. It thus has the form $x = c, y = c, z = -2c$. To insure that the vector is a unit vector, we take $c = \frac{1}{\sqrt{6}}$. Thus,

$$U = \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & -\frac{2}{\sqrt{6}} \end{bmatrix}$$

EXERCISES

1. For each of the following matrices A determine an orthogonal matrix U such that $U^{-1}AU$ is diagonal.

   (a) $\begin{bmatrix} 2 & 3 \\ 3 & 2 \end{bmatrix}$ (b) $\begin{bmatrix} 5 & 2 \\ 2 & 2 \end{bmatrix}$ (c) $\begin{bmatrix} 7 & 6 \\ 6 & 2 \end{bmatrix}$ (d) $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

   (e) $\begin{bmatrix} 2 & -4 & 2 \\ -4 & 2 & -2 \\ 2 & -2 & -1 \end{bmatrix}$ (f) $\begin{bmatrix} -3 & -6 & 0 \\ -6 & 0 & 6 \\ 0 & 6 & 3 \end{bmatrix}$ (g) $\begin{bmatrix} -1 & -2 & 1 \\ -2 & 2 & -2 \\ 1 & -2 & -1 \end{bmatrix}$

2. Find a 3×3 matrix A having eigenvalues $1, 0, -1$ and corresponding eigenvectors

   $$\tfrac{1}{3}\begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix}, \quad \tfrac{1}{3}\begin{bmatrix} 2 \\ -1 \\ 2 \end{bmatrix}, \quad \tfrac{1}{3}\begin{bmatrix} 2 \\ 2 \\ -1 \end{bmatrix}$$

3. For each of the following matrices A determine a unitary matrix U such that $U^{-1}AU$ is diagonal.

   (a) $\begin{bmatrix} 0 & i \\ -i & 0 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & 1+i \\ 1-i & 0 \end{bmatrix}$

4. Let B be a real 3×3 matrix that is similar to a diagonal matrix having diagonal entries x_1, x_2, x_3. Show that $aI_3 + bB$ is similar to a diagonal matrix having diagonal entries $a + x_1 b, a + x_2 b, a + x_3 b$. To which diagonal matrices are the following matrices similar?

   (a) $\begin{bmatrix} a & b \\ b & a \end{bmatrix}$ (b) $\begin{bmatrix} a & b & b \\ b & a & b \\ b & b & a \end{bmatrix}$ (c) $\begin{bmatrix} a & 0 & b \\ 0 & a & 0 \\ b & 0 & a \end{bmatrix}$

5. To what diagonal matrix is the matrix $\begin{bmatrix} a & c \\ c & b \end{bmatrix}$ similar?

Since

$$(T(\boldsymbol{x}), \boldsymbol{x}) = (\boldsymbol{x}, T(\boldsymbol{x})), \text{ as } T \text{ is symmetric}$$

$$\lambda(\boldsymbol{x}, \boldsymbol{x}) = \bar{\lambda}(\boldsymbol{x}, \boldsymbol{x})$$

Since $\boldsymbol{x} \neq \mathbf{0}$, $(\boldsymbol{x}, \boldsymbol{x}) \neq 0$, and $\lambda = \bar{\lambda}$. Thus, λ is real. ■

Note that the same result holds if A is an $n \times n$ Hermitian matrix, that is, $A = A^*$. Indeed,

$$(Ax, y) = (Ax)^* y = x^* A^* y = x^* A y = (x, Ay)$$

If $Ax = \lambda x$, the proof above yields $\lambda = \bar{\lambda}$.

If T is a symmetric operator on an inner product space, by observing that the matrix representation of T with respect to an orthonormal basis is Hermitian, it follows that all eigenvalues of T are real.

We say that a matrix is **orthogonally diagonalizable** if it is orthogonally similar to a diagonal matrix. In §7.1 we saw that a matrix A is diagonalizable over the reals if and only if $\boldsymbol{R}^n$ admits a basis of eigenvectors of A. In the same way, it may be shown that a matrix A is orthogonally diagonalizable if and only if $\boldsymbol{R}^n$ admits an orthonormal basis of eigenvectors of A. With the same idea in mind we say that a linear operator T on an inner product space V is orthogonally diagonalizable if V admits an orthonormal basis of T eigenvectors.

We now prove the main theorem of this section.

Theorem 2 Let T be a symmetric linear operator on a finite-dimensional inner product space V. Then V has an orthonormal basis of eigenvectors of T.

PROOF The proof is by induction on dim V. The theorem is obvious if dim $V = 1$.

We know that T has a (possibly complex) eigenvalue. (Take any root of the characteristic polynomial.) By Theorem 1, this root is real. Thus, T has a real eigenvalue. Let $\boldsymbol{x}$ be an eigenvector corresponding to this eigenvalue. So we have $T(\boldsymbol{x}) = \lambda \boldsymbol{x}$, with λ real. We may assume $|\boldsymbol{x}| = 1$.

Let H be the orthogonal complement of $\boldsymbol{x}$ in V. We claim: If $\boldsymbol{h}$ belongs to H, so does $T(\boldsymbol{h})$.

Indeed,

$$\begin{aligned}(T(\boldsymbol{h}), \boldsymbol{x}) &= (\boldsymbol{h}, T(\boldsymbol{x})) && \text{(by symmetry of } T) \\ &= (\boldsymbol{h}, \lambda \boldsymbol{x}) \\ &= \lambda(\boldsymbol{h}, \boldsymbol{x}) \\ &= 0 && \text{(since } \boldsymbol{h} \text{ belongs to } H)\end{aligned}$$

Thus, $T(\boldsymbol{h})$ belongs to H, as well.

It follows that if T is restricted to H, T is a linear operator on H. Since dim H = dim $V - 1$, by induction hypothesis, H has an orthonormal basis of T eigenvectors, say $x_2, \boldsymbol{x}_3, \ldots, \boldsymbol{x}_n$. Thus, $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n$ is an orthonormal basis for V of eigenvectors of T. ■

As a consequence of this theorem, we have

(1) If A is a real symmetric matrix, A is orthogonally similar to a real diagonal matrix.
(2) If A is a Hermitian matrix, A is unitarily similar to a real diagonal matrix.

It is interesting to note that the only matrices that are orthogonally similar to diagonal matrices are symmetric. Indeed, let A be a matrix that is orthogonally similar to the diagonal matrix D. Then, there is an orthogonal matrix U such that $U^{-1}AU = D$. It follows that $A = UDU^{-1}$. Since U is orthogonal, $U^{-1} = U^T$. Thus, $A = UDU^T$. Therefore, $A^T = (UDU^T)^T = (U^T)^T D^T U^T = UD^T U^T$. Now since D is diagonal, $D^T = D$. Thus, $A^T = UDU^T$. It follows that $A^T = A$, and so A is symmetric. Thus, a real matrix is orthogonally diagonalizable if and only if it is symmetric.

Example 1 If $A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$, find an orthogonal matrix U such that $U^{-1}AU$ is diagonal.

In this case, $C_A(x) = x^2 - 4x + 3 = (x-1)(x-3)$. Thus, the eigenvalues of A are 1 and 3. To find the corresponding eigenvectors, we must solve the systems

$$\begin{aligned} 2x + y &= x \\ x + 2y &= y \end{aligned} \quad \text{and} \quad \begin{aligned} 2x + y &= 3x \\ x + 2y &= 3y \end{aligned}$$

Doing this, we obtain the eigenvectors $\begin{bmatrix} c \\ -c \end{bmatrix}$ and $\begin{bmatrix} d \\ d \end{bmatrix}$, with c and d arbitrary. Next, we choose c and d so that the eigenvectors have length 1. Thus, take $c = d = \frac{1}{\sqrt{2}}$. The matrix U can be found by writing the eigenvectors in its successive columns. We find that

$$U = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$

Note that U is an orthogonal matrix. By the way U has been chosen, it follows that $U^{-1}AU$ is the diagonal matrix with diagonal entries 1 and 3. We verify this although it is not really necessary. Note

$$U^{-1}AU = U^T AU = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \cdot \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$

$$= \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \cdot \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 3 \\ -1 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}$$

(d) Prove that

$$\frac{1}{30}\begin{bmatrix}16 & 16 & 4\\ 12 & 12 & 3\\ 8 & 8 & 2\end{bmatrix}+\tfrac{1}{30}\left(-\tfrac{1}{2}\right)^n\begin{bmatrix}20 & -20 & -10\\ -30 & 30 & 15\\ 40 & -40 & -20\end{bmatrix}+\tfrac{1}{30}\left(-\tfrac{1}{4}\right)^n\begin{bmatrix}-6 & 4 & 6\\ 18 & -12 & -18\\ -48 & 32 & 48\end{bmatrix}$$

(e) If the initial population is $\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}$, show that the population distribution approaches $\dfrac{(4x_1+4x_2+x_3)}{30}\begin{bmatrix}4\\ 3\\ 2\end{bmatrix}$ after enough time has elapsed. (In fact, after 10 years the distribution is correct to within 1% of the total population.)

17. Let $A_1, A_2, A_3, \ldots$ be a sequence of real $n\times n$ matrices. The sequence is said to approach the limit L if for all i and j, the sequence of entries in position (i,j) approaches the entry of L in position (i,j). This is written as $\lim A_n = L$.
 (a) If $\lim A_n = L$ and B is an invertible $n\times n$ matrix, show that $\lim B^{-1}A_nB = B^{-1}LB$.
 (b) Let A be a 2×2 real matrix having positive entries. Show that the sequence $A, A^2, A^3, \ldots$ of powers of A has a limit if and only if the eigenvalues λ_1 and λ_2 of A satisfy $-1<\lambda_1\leqslant 1$ and $-1<\lambda_2\leqslant 1$. [Hint: Use exercise 15.]
 (c) If in (b) neither eigenvalue is 1, show that the limit of the sequence is 0.

18. On the vector space P_n, show that the linear operator $T(f)=xf'$ is diagonalizable.

19. For each of the following matrices A find a complex matrix B such that $B^{-1}AB$ is diagonal.

$$\text{(a) }\begin{bmatrix}3 & -4\\ 4 & 3\end{bmatrix}\qquad \text{(b) }\begin{bmatrix}0 & 1\\ -1 & 0\end{bmatrix}\qquad \text{(c) }\begin{bmatrix}0 & 0 & 1\\ 1 & 0 & 0\\ 0 & 1 & 0\end{bmatrix}$$

20. Let a, b, and c be arbitrary numbers. A sequence is defined according to the following rule: $x_1=a$, $x_2=b$, $x_3=c$, and for all $k>3$, x_k is the average of the three elements of the sequence that precedes it. Thus, $x_4=\frac{1}{3}(a+b+c)$, $x_5=\frac{1}{9}(a+4b+4c)$, and so on. Show that the limit of the sequence is $\frac{1}{6}(a+2b+3c)$. [Hint: Let v_k be the vector with first component x_{k+2}, second component x_{k+1}, and third component x_k. Find a matrix A such that $v_{k+1}=Av_k$ for all k, and compute the powers of A.]

21. Let A and B be two $n\times n$ matrices (real or complex), and let x be a fixed scalar.
 (a) Show that if λ is sufficiently large,

$$\det\left[xI_n-(A-\lambda I_n)B\right]=\det\left[xI_n-B(A-\lambda I_n)\right]$$

 [Hint: If λ is big enough, $A-\lambda I_n$ is invertible.]
 (b) Using part (a) show that for all λ

$$\det\left[xI_n-(A-\lambda I_n)B\right]=\det\left[xI_n-B(A-\lambda I_n)\right]$$

 [Hint: Two polynomials that are equal at infinitely many points are equal everywhere.]
 (c) Show that $C_{AB}(x)=C_{BA}(x)$.

22. Let A be an $n\times n$ real matrix with n real distinct eigenvalues. If B is real and commutes with A, show that B is diagonalizable.

23. If

$$A=\begin{bmatrix}a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{bmatrix}$$

show that the coefficient of x in $C_A(x)$ is

$$\begin{vmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{vmatrix}+\begin{vmatrix}a_{11} & a_{13}\\ a_{31} & a_{33}\end{vmatrix}+\begin{vmatrix}a_{22} & a_{23}\\ a_{32} & a_{33}\end{vmatrix}$$

2 SYMMETRIC MATRICES

If A and B are two $n\times n$ matrices and there is an orthogonal matrix U such that $A=U^{-1}BU$, we say that A and B are **orthogonally similar.** Since U is orthogonal, $U^{-1}=U^T$. Thus, $A=U^TBU$. If U is unitary and $A=U^{-1}BU$, A and B are said to be **unitarily similar.** In this section, we show that any symmetric matrix is orthogonally similar to a diagonal matrix. In particular, it follows that any symmetric matrix is diagonalizable.

In this section, we often use the standard inner product on $\boldsymbol{R}^n$ defined by $(x, y)=x^Ty$. On $\boldsymbol{C}^n$ we use the inner product $(x, y)=x^*y$ of §6.2.

Suppose that A is a real symmetric matrix and $T(\boldsymbol{x})=A\boldsymbol{x}$ is the linear operator induced by A. We observe that

$$(T(x), y)=(T(x))^Ty=(Ax)^Ty=x^TA^Ty=x^TAy=(x, T(y))$$

Thus, T has the property that $(T(x), y)=(x, T(y))$. Any linear operator on an inner product space that has this property is said to be **symmetric.** It is not hard to show (using exercise 14 of §6.2) that the matrix representation of a symmetric operator, with respect to the standard basis for $\boldsymbol{R}^n$ (or any other orthonormal basis) is symmetric. In proving things about symmetric matrices, it is generally useful to keep the property $(T(x), y)=(x, T(y))$ in mind.

Our first objective is to prove that all eigenvalues of a symmetric matrix are real numbers.

Theorem 1 Let A be an $n\times n$ real symmetric matrix. Then any eigenvalue of A is real.

PROOF Let T be the linear operator on $\boldsymbol{C}^n$ defined by $T(\boldsymbol{x})=A\boldsymbol{x}$. Let $(\ ,\ )$ be the standard inner product on $\boldsymbol{C}^n$, i.e., $(\boldsymbol{x}, \boldsymbol{y})=\boldsymbol{x}^*\boldsymbol{y}$, for all $\boldsymbol{x}$ and $\boldsymbol{y}$ in $\boldsymbol{C}^n$. As above, $(T(x), y)=(x, T(y))$.

If $T(x)=\lambda x$,

$$(T(\boldsymbol{x}), \boldsymbol{x})=(\lambda\boldsymbol{x}, \boldsymbol{x})=\lambda(\boldsymbol{x}, \boldsymbol{x})$$

and

$$(\boldsymbol{x}, T(\boldsymbol{x}))=(\boldsymbol{x}, \lambda\boldsymbol{x})=\bar{\lambda}(\boldsymbol{x}, \boldsymbol{x})$$

right, it follows that $A = 0$, which is a contradiction. Therefore, A is not diagonalizable.

EXERCISES

1. Determine the characteristic polynomial and eigenvalues of the following matrices.

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

(e) $\begin{bmatrix} 7 & 8 & 8 \\ -10 & -11 & -10 \\ 4 & 4 & 3 \end{bmatrix}$ (f) $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$ (g) $\begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$

2. Determine the eigenvalues and corresponding eigenspaces of the following matrices.

(a) $\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$ (b) $\begin{bmatrix} -5 & 3 & 0 \\ -6 & 4 & 2 \\ 2 & -1 & 1 \end{bmatrix}$ (c) $\begin{bmatrix} 0 & 1 & 0 \\ -1 & 2 & 0 \\ 1 & 1 & -1 \end{bmatrix}$

(d) $\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$

3. Show that the following matrices are diagonalizable, and find the diagonal matrix to which each is similar.

(a) $\begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \end{bmatrix}$ (b) $\begin{bmatrix} 2 & 2 & 0 \\ 1 & 1 & 2 \\ 1 & 1 & 2 \end{bmatrix}$ (c) $\begin{bmatrix} -2 & -4 & -5 \\ 1 & 3 & 1 \\ 2 & 2 & 5 \end{bmatrix}$

4. Determine the axis of rotation of the following rotation matrices.

(a) $\frac{1}{7}\begin{bmatrix} 3 & -2 & 6 \\ -6 & -3 & 2 \\ 2 & -6 & -3 \end{bmatrix}$ (b) $\frac{1}{3}\begin{bmatrix} 2 & -1 & 2 \\ -2 & -2 & 1 \\ 1 & -2 & -2 \end{bmatrix}$ (c) $\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$

5. Determine the characteristic polynomial and eigenvalues of the following operator on P_3.

(a) $D(f) = f'$ (b) $(S(f))(x) = f(x+1)$ (c) $T(f) = (1-x^2)f'' - 2xf'$

6. If $A = [a_{ij}]$ is an $n \times n$ upper triangular matrix, show that $C_A(x) = (x - a_{11})(x - a_{22}) \ldots (x - a_{nn})$.

7. Let A be an $n \times n$ matrix having characteristic polynomial $(x-a)^n$. If A is diagonalizable, show that $A = aI_n$.

8. Let A be an $n \times n$ matrix. If the sum of the entries on all columns of A is equal to s, show that s is an eigenvalue of A.

9. As in example 4 of §5.1, three companies control the market for a certain product. The following stochastic matrices give the retention and loss of customers, as in example 4. Compute the stable states of the market.

(a) $\begin{bmatrix} \frac{3}{4} & \frac{1}{8} & \frac{1}{4} \\ \frac{1}{8} & \frac{3}{4} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{5}{8} \end{bmatrix}$ (b) $\begin{bmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{bmatrix}$ (c) $\begin{bmatrix} \frac{1}{3} & \frac{1}{2} & \frac{4}{9} \\ \frac{1}{3} & \frac{1}{4} & \frac{4}{9} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{9} \end{bmatrix}$

10. For each of the following matrices A, find a matrix B such that $B^{-1}AB$ is diagonal.

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

11. Show that two diagonalizable matrices are similar if and only if they have the same characteristic polynomial.

12. (a) If A is an $n \times n$ matrix and $A^T = \lambda A$, show that $\lambda = \pm 1$, if $A \neq 0$.
(b) Show that the only eigenvalues of the linear transformation of the space of $n \times n$ matrices into itself defined by $T(A) = A^T$ are ± 1.
(c) Determine the eigenspaces of each eigenvalue.

13. Find a general formula for the nth power of the following matrices.

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

14. Let T be a linear operator on a finite-dimensional vector space V. Show that T is invertible if and only if 0 is not an eigenvalue of T.

15. Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ be a 2×2 real matrix, and suppose that all of the entries of A are positive numbers.
(a) Show that the eigenvalues of A are given by the formula

$$\tfrac{1}{2}\left((a+d) \pm \sqrt{(a-d)^2 + 4bc}\right)$$

(b) Show that the eigenvalues of A are real and distinct.
(c) Show that at least one eigenvalue of A is positive.
(d) Show that there is an eigenvector corresponding to the largest eigenvalue such that both components of the eigenvector are positive.

16. As in example 3 of §2.5, we study the population changes of a certain animal species. Here we suppose that there are only three age groups. Let v be the 3-vector whose components give the number of members of each age group. Let $T(v)$ give the population distribution after a year has elapsed. We suppose that $T(v)$ is given by

$$T(v) = Av = \begin{bmatrix} \frac{1}{4} & \frac{5}{6} & \frac{1}{4} \\ \frac{3}{4} & 0 & 0 \\ 0 & \frac{2}{3} & 0 \end{bmatrix} v$$

(a) Interpret the entries of the matrix A.
(b) Find the characteristic polynomial and eigenvalues of A.
(c) Find a matrix B such that $B^{-1}AB$ is diagonal.

In this case we may also provide an explicit means for calculating the matrix B such that $B^{-1}AB = D$ is diagonal. Indeed, let $x_1, \ldots, x_n$ be the eigenvectors corresponding to $\lambda_1, \ldots, \lambda_n$. These vectors may be found by solving the system of linear equations $(A - \lambda_i)x = \mathbf{0}$.

Let $B = [x_1, x_2, \ldots, x_n]$, that is, B is the matrix whose ith column is x_i. Then $Be_i = x_i$ and $B^{-1}x_i = e_i$. Consequently,

$$\begin{aligned}(B^{-1}AB)e_i &= B^{-1}(Ax_i)\\ &= B^{-1}(\lambda_i x_i)\\ &= \lambda_i e_i\end{aligned}$$

Thus, $B^{-1}AB$ is a diagonal matrix with successive diagonal entries $\lambda_1, \lambda_2, \ldots, \lambda_n$.

Example 8 Let

$$A = \begin{bmatrix} -2 & 3 & 1 \\ 0 & 1 & 1 \\ -3 & 4 & 1 \end{bmatrix}$$

Then

$$C_A(x) = \begin{vmatrix} x+2 & -3 & -1 \\ 0 & x-1 & -1 \\ 3 & -4 & x-1 \end{vmatrix} = x^3 - 4x = x(x-2)(x+2)$$

Thus, the eigenvalues of A are 0, 2, and -2, which are real and distinct. It follows that there is a real invertible 3×3 matrix B such that

$$B^{-1}AB = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -2 \end{bmatrix}$$

B may be determined explicitly by finding the eigenvectors corresponding to 0, 2, and -2 and writing these as successive columns of the matrix B.

Example 9 Let

$$B_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

Earlier, we saw that B_θ has the complex eigenvalues

$$\cos\theta + i\sin\theta, \quad \cos\theta - i\sin\theta$$

Thus, by Theorem 4, there is a complex matrix C such that

$$CB_\theta C^{-1} = \begin{bmatrix} \cos\theta + i\sin\theta & 0 \\ 0 & \cos\theta - i\sin\theta \end{bmatrix}$$

Note that though B_θ has no real eigenvalues, it is diagonalizable over the complexes.

Many calculations involving diagonal matrices are easily performed. It is, for example, quite simple to compute high powers of diagonal matrices. Such calculations can also be performed with diagonalizable matrices. For example, if C is diagonalizable and $B^{-1}CB = D$ is diagonal, to compute C^n, we need only compute D^n. Then, $C^n = BD^nB^{-1}$. In example 4 of §5.8, this procedure was used to compute high powers of a stochastic matrix. We present another example of this type.

Example 10 Compute C^n if $C = \begin{bmatrix} 4 & -2 \\ 1 & 1 \end{bmatrix}$.

In this case the characteristic polynomial of C is $x^2 - 5x + 6 = (x-2)(x-3)$. Thus, C has eigenvalues 2 and 3, and by Theorem 4, C is diagonalizable. The eigenvectors corresponding to the eigenvalues 3 and 2, respectively, can be obtained by solving the systems

$$\begin{aligned} -x + 2y &= 0 \\ -x + 2y &= 0 \end{aligned} \quad \text{and} \quad \begin{aligned} -2x + 2y &= 0 \\ x - y &= 0 \end{aligned}$$

We find that $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ are eigenvectors corresponding to the eigenvalues 3 and 2, respectively. We set $B = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$. It follows that $B^{-1} = \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}$ and that $B^{-1}CB = \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}$. Therefore,

$$\begin{aligned} C^n = B\begin{bmatrix} 3^n & 0 \\ 0 & 2^n \end{bmatrix} B^{-1} &= \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 3^n & 0 \\ 0 & 2^n \end{bmatrix}\begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix} \\ &= \begin{bmatrix} 2 \cdot 3^n - 2^n & -2 \cdot 3^n + 2^{n+1} \\ 3^n - 2^n & -3^n + 2^{n+1} \end{bmatrix} \end{aligned}$$

A similar procedure shows that if f is any polynomial

$$f(C) = \begin{bmatrix} 2f(3) - f(2) & -2f(3) + 2f(2) \\ f(3) - f(2) & -f(3) + 2f(2) \end{bmatrix}$$

Diagonalizable matrices are easy to compute with. However, it is not true that all matrices are diagonalizable.

Example 11 Let $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$. Then $C_A(x) = x^2$.

Suppose that A is diagonalizable. Then there is an invertible matrix C such that $C^{-1}AC = \begin{bmatrix} d_1 & 0 \\ 0 & d_2 \end{bmatrix}$. Since similar matrices have the same characteristic polynomial, $(x - d_1)(x - d_2) = x^2$. It follows that $d_1 = d_2 = 0$. Thus, $C^{-1}AC = 0$. If we multiply this equation by C on the left and C^{-1} on the

Let A be an $n \times n$ matrix all of whose column sums are 1. Let $\boldsymbol{u}$ be the column n-vector having all entries 1. Note that in the matrix A^T all row sums are 1. It follows from this that $A^T\boldsymbol{u} = \boldsymbol{u}$. Since $\boldsymbol{u} \neq 0$, 1 is an eigenvalue of A^T. Thus, $C_{A^T}(1) = 0$. Since $C_{A^T}(x) = C_A(x)$, it follows that $C_A(1) = 0$. Thus, 1 is an eigenvalue of A.

It is important to note that a real matrix may have no real eigenvalues. For example, consider the matrix

$$B_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

We have seen that B_θ induces a rotation of the plane by θ degrees. If $0 < \theta < \pi$, it is clear that no vector in $\boldsymbol{R}^2$ is carried into a scalar multiple of itself. Thus, B_θ has no real eigenvalues.

In this case, the characteristic polynomial is $x^2 - (2\cos\theta)x + 1$. Its roots are $\cos\theta + i\sin\theta$ and $\cos\theta - i\sin\theta$. If $0 < \theta < \pi$, $\sin\theta \neq 0$, and neither root is real, a conclusion that was to be expected from the above geometric considerations.

It is known that any polynomial with complex coefficients has a complex root. Thus, any $n \times n$ matrix with complex entries has some complex eigenvalue. Consequently, any real matrix has some complex eigenvalue. Sometimes this fact, coupled with other information about a given matrix, can be used to show that the matrix has a real eigenvalue.

There is an important relationship between eigenvectors and linear independence that is brought forth in the following theorem.

Theorem 3 Let T be a linear operator on a vector space V. Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be distinct eigenvalues of T and let $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n$ be eigenvectors (necessarily nonzero vectors) corresponding to the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$, respectively. Then $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n$ are linearly independent.

PROOF Suppose $\alpha_1\boldsymbol{x}_1 + \alpha_2\boldsymbol{x}_2 + \cdots + \alpha_n\boldsymbol{x}_n = \boldsymbol{0}$. Applying T to this equation and using $T(\boldsymbol{x}_i) = \lambda_i\boldsymbol{x}_i$, we get

$$\alpha_1\lambda_1\boldsymbol{x}_1 + \alpha_2\lambda_2\boldsymbol{x}_2 + \cdots + \alpha_n\lambda_n\boldsymbol{x}_n = \boldsymbol{0}$$

Multiplying the first of the above equations by λ_1 and subtracting from the second, we obtain

$$\alpha_2(\lambda_2 - \lambda_1)\boldsymbol{x}_2 + \alpha_3(\lambda_3 - \lambda_1)\boldsymbol{x}_3 + \cdots + \alpha_n(\lambda_n - \lambda_1)\boldsymbol{x}_n = \boldsymbol{0}$$

Applying T to this equation,

$$\alpha_2(\lambda_2 - \lambda_1)\lambda_2\boldsymbol{x}_2 + \alpha_3(\lambda_3 - \lambda_1)\lambda_3\boldsymbol{x}_3 + \cdots + \alpha_n(\lambda_n - \lambda_1)\lambda_n\boldsymbol{x}_n = \boldsymbol{0}$$

Multiplying the first of the preceding two equations by λ_2 and subtracting from the second, we obtain

$$\alpha_3(\lambda_3 - \lambda_1)(\lambda_3 - \lambda_2)\boldsymbol{x}_3 + \cdots + \alpha_n(\lambda_n - \lambda_1)(\lambda_n - \lambda_2)\boldsymbol{x}_n = \boldsymbol{0}$$

By repeating this process, we eventually obtain

$$\alpha_n(\lambda_n - \lambda_1)(\lambda_n - \lambda_2)\cdots(\lambda_n - \lambda_{n-1})\boldsymbol{x}_n = \boldsymbol{0}$$

Since $\lambda_1, \lambda_2, \ldots, \lambda_n$ are distinct, $(\lambda_n - \lambda_1)(\lambda_n - \lambda_2)\cdots(\lambda_n - \lambda_{n-1}) \neq 0$, and since it is also true that $\boldsymbol{x}_n \neq \boldsymbol{0}$, it follows that $\alpha_n = 0$.

Thus, $\alpha_1\boldsymbol{x}_1 + \cdots + \alpha_{n-1}\boldsymbol{x}_{n-1} = \boldsymbol{0}$. Using the same procedure on this equation repeatedly yields

$$\alpha_1 = \alpha_2 = \cdots = \alpha_{n-1} = 0$$

Thus, we have shown that whenever $\alpha_1\boldsymbol{x}_1 + \alpha_2\boldsymbol{x}_2 + \cdots + \alpha_n\boldsymbol{x}_n = \boldsymbol{0}$, $\alpha_1 = \alpha_2 = \cdots \alpha_n = 0$, and so $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n$ are linearly independent. ■

Recall that two real matrices C and D are said to be similar if there is a real invertible matrix B such that $B^{-1}CB = D$. There is an analogous definition for similarity of complex matrices. We say that a real matrix is **diagonalizable** over the reals if it is similar over the reals to some real diagonal matrix. There is an analogous definition for diagonalizability over the complexes.

A linear operator T on a vector space V is said to be diagonalizable if there is some basis for V such that the matrix representation of T, with respect to some basis for V, is diagonal. Since the matrix representations of T, with respect to the various bases of V, are all similar, T is diagonalizable if and only if one (and hence all) of its matrix representations are diagonalizable.

Suppose that the linear operator T is represented, with respect to the basis $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$, by the diagonal matrix D having diagonal entries $d_1, \ldots, d_n$. Then $T(\boldsymbol{x}_i) = d_i\boldsymbol{x}_i$. Thus, the basis vectors $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ are eigenvectors of T. Thus, if T is diagonalizable, there is a basis consisting entirely of eigenvectors of T. Conversely, if there is a basis consisting entirely of eigenvectors of T, then the matrix representation of T, with respect to this basis, is diagonal, and so T is diagonalizable.

In terms of matrices this says: An $n \times n$ matrix A is diagonalizable over the reals if and only if $\boldsymbol{R}^n$ has a basis consisting of eigenvectors of A.

An analogous statement holds for the diagonalization of complex matrices.

We give a condition on the characteristic polynomial that guarantees that a matrix is diagonalizable.

Theorem 4 Let A be an $n \times n$ matrix with real entries and suppose $C_A(x) = (x - \lambda_1)(x - \lambda_2)\cdots(x - \lambda_n)$, where $\lambda_1, \ldots, \lambda_n$ are distinct real numbers. Then A is diagonalizable over the reals.

PROOF Take $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ eigenvectors of A corresponding to the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$, respectively. By Theorem 3 above and by virtue of the fact that $\lambda_1, \lambda_2, \ldots, \lambda_n$ are distinct numbers, $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n$ are linearly independent. But by Theorem 2, §4.8, $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n$ also span $\boldsymbol{R}^n$, and thus $\boldsymbol{R}^n$ has a basis of eigenvectors of A. Consequently, A is diagonalizable. ■

Then

$$\begin{aligned}\det[xI_n - A'] &= \det[xI_n - BAB^{-1}] \\ &= \det[B(xI_n - A)B^{-1}] \\ &= \det B \det[xI_n - A] \det B^{-1} \\ &= \det[xI_n - A]\end{aligned}$$

Consequently, the characteristic polynomial of T is independent of the basis used to compute it.

In a similar vein, we denote the characteristic polynomial of the matrix A by C_A. The previous paragraph implies that similar matrices have the same characteristic polynomial.

From our definition of the characteristic polynomial and from Theorem 1, it follows that the roots of the characteristic polynomial are precisely the eigenvalues of the linear operator.

If $A = [a_{ij}]_{(nn)}$ is an $n \times n$ matrix,

$$C_A(x) = \begin{vmatrix} x - a_{11} & -a_{12} & -a_{13} & \cdots & -a_{1n} \\ -a_{21} & x - a_{22} & -a_{23} & \cdots & -a_{2n} \\ -a_{31} & -a_{32} & x - a_{33} & \cdots & -a_{3n} \\ & & & \vdots & \\ -a_{n1} & -a_{n2} & -a_{n3} & \cdots & x - a_{nn} \end{vmatrix}$$

It is clear that $C_A(x)$ is a polynomial of degree n with leading coefficient 1:

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

Expanding the above determinant for $C_A(x)$ yields that the coefficient of x^{n-1} is

$$-(a_{11} + a_{22} + \cdots + a_{nn}) = -\operatorname{tr} A$$

The constant term a_0 in the above polynomial may be calculated by setting $x = 0$. Then $a_0 = C_A(0) = \det(-A)$, or $a_0 = (-1)^n \det A$.

Example 4 Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ be a 2×2 matrix. Then

$$C_A(x) = \begin{vmatrix} x - a & -b \\ -c & x - d \end{vmatrix} = x^2 - (a + d)x + ad - bc = x^2 - (\operatorname{tr} A)x + \det A$$

Thus, in the two-dimensional case, the characteristic polynomial of A is completely determined by the determinant and trace of A. The eigenvalues of A may be found by solving the equation $x^2 - (\operatorname{tr} A)x + \det A = 0$.

For example, if $A = \begin{bmatrix} 2 & 3 \\ -1 & -2 \end{bmatrix}$, $\operatorname{tr} A = 0$, $\det A = -1$, and $C_A(x) = x^2 - 1$. Since the roots of the equation $x^2 - 1 = 0$ are $x = 1, -1$, the eigenvalues of A are 1 and -1. By solving linear equations, we find that corresponding eigenvectors are $\begin{bmatrix} 3 \\ -1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$.

Example 5 Let $A = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & -1 \\ 1 & 5 & 3 \end{bmatrix}$. Find the eigenvalues of A. $C_A(x) = \begin{vmatrix} x & -1 & -1 \\ -1 & x & 1 \\ -1 & -5 & x - 3 \end{vmatrix} = x(x^2 - 3x + 5) + (-x + 3 + 1) - (5 + x) = x^3 - 3x^2 + 3x - 1 = (x - 1)^3$. From this, it follows that the only eigenvalue of A is 1. The eigenvectors of A corresponding to the eigenvalue 1 can be found by determining the nullspace of $I_3 - A$.

The characteristic polynomials of certain matrices are quite easy to compute. For example, the characteristic polynomial of I_n is simply $(x - 1)^n$. More generally, if D is an $n \times n$ diagonal matrix with diagonal entries $d_1, d_2, \ldots, d_n$, the characteristic polynomial of D is $(x - d_1)(x - d_2) \cdots (x - d_n)$.

We list a few additional properties of the characteristic polynomial.

Theorem 2 Let A be an $n \times n$ matrix. Then, $C_{A^T}(x) = C_A(x)$. Moreover, if A is invertible, $C_{A^{-1}}(x) = (\det A)^{-1}(-x)^n C_A(\frac{1}{x})$.

PROOF We prove the first statement. We have $C_{A^T}(x) = \det[xI_n - A^T] = \det[(xI_n - A)^T] = \det[xI_n - A] = C_A(x)$.

Next, suppose that A is invertible. Then $C_{A^{-1}}(x) = \det[xI_n - A^{-1}] = \det[A^{-1}(xA - I_n)] = (\det A)^{-1} \det[xA - I_n] = (\det A)^{-1}(-x)^n \det[\frac{1}{x} I_n - A] = (\det A)^{-1}(-x)^n C_A(\frac{1}{x})$. ■

Example 6 A 3×3 orthogonal matrix U is said to be a rotation matrix if $\det U = 1$. We will show that 1 is an eigenvalue of a rotation matrix. Geometrically this means that every rotation of $\boldsymbol{R}^3$ fixes a vector, namely the eigenvector corresponding to the eigenvalue 1. In other words, every rotation of $\boldsymbol{R}^3$ has an axis.

We wish to show that $C_U(1) = 0$. By Theorem 2, $C_U(1) = C_{U^T}(1)$ and $C_{U^{-1}}(1) = 1(-1)^3 C_U(1) = -C_U(1)$. Since $UU^T = I_3$, $U^T = U^{-1}$. Thus, $C_U(1) = -C_U(1)$. It follows that $C_U(1) = 0$.

A similar argument shows that a 3×3 orthogonal matrix with determinant -1 has -1 as an eigenvalue.

Example 7 Recall that a square matrix A is stochastic if all its entries are nonnegative and its column sums are 1. In example 4 of both §5.1 and §5.8, we saw instances of such matrices. In many applications of such matrices, we are concerned with those vectors $\boldsymbol{v}$ such that $A\boldsymbol{v} = \boldsymbol{v}$, i.e., the eigenvectors corresponding to the eigenvalue 1. In example 2 above, we found such eigenvectors. We will show that a stochastic matrix always has 1 as an eigenvalue. Thus, there always exist nonzero vectors $\boldsymbol{v}$ such that $A\boldsymbol{v} = \boldsymbol{v}$.

operator T fixes the lines through the origin determined by scalar multiples of the eigenvectors. (See Figure 7-1.)

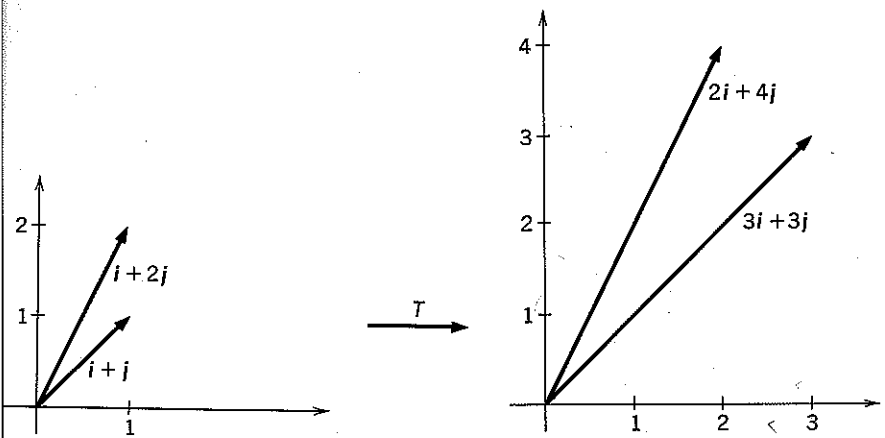

FIGURE 7-1

Earlier in the text, we encountered several examples of eigenvalues and eigenvectors, although at the time we did not regard them as such. For instance, in example 3 of §2.5 we considered a model for population growth in which the 4-vector $\boldsymbol{p}$ exhibits the population of each age group. Multiplying $\boldsymbol{p}$ by the matrix T gives the population distribution $T\boldsymbol{p}$ one year later. At the end of that example, we noted that there was a vector $\boldsymbol{p}_0$ such that $T\boldsymbol{p}_0 = \boldsymbol{p}_0$. Thus, T has eigenvalues 1 and $\boldsymbol{p}_0$ is the corresponding eigenvector. In this example, determining all the eigenvectors corresponding to the eigenvalue 1 will give the population distributions that are stable from year to year.

Example 2 In example 4 of §5.1, we considered a situation in which three companies controlled the market for a certain product. A matrix A was given that indicated retention and loss of customers. If the distribution of the market is given by the 3-vector $\boldsymbol{v}$, then the distribution one year later is given by $A\boldsymbol{v}$.

What are the stable distributions of the market?

Here the problem is to determine those vectors $\boldsymbol{v}$ such that $A\boldsymbol{v} = \boldsymbol{v}$, i.e., all eigenvectors corresponding to the eigenvalue 1. Equivalently, we wish to determine those vectors $\boldsymbol{v}$ for which $(A - I_3)\boldsymbol{v} = \boldsymbol{0}$. In other words, we must solve the linear equations

$$-\tfrac{1}{2}x + \tfrac{1}{3}y + \tfrac{1}{6}z = 0$$

$$\tfrac{1}{4}x - \tfrac{1}{2}y + \tfrac{1}{6}z = 0$$

$$\tfrac{1}{4}x + \tfrac{1}{6}y - \tfrac{1}{3}z = 0$$

for x, y, and z. Solving as usual, we find that $x = 10c$, $y = 9c$, and $z = 12c$, with c arbitrary. In this case, it is natural to choose c so that $x + y + z = 1$, or $c = \frac{1}{31}$. Thus, in a stable market the shares of the companies are $\frac{10}{31}$, $\frac{9}{31}$, and $\frac{13}{31}$, respectively. The vector $\boldsymbol{v}_0$ with these components is an eigenvector corresponding to the eigenvalue 1.

Since $T(\boldsymbol{x}) = \lambda\boldsymbol{x}$ is equivalent to $(T - \lambda I)(\boldsymbol{x}) = \boldsymbol{0}$, it follows that $\boldsymbol{x}$ is an eigenvector of T corresponding to the eigenvalue λ if and only if $\boldsymbol{x}$ belongs to the nullspace of $\lambda I - T$. We call the collection of all eigenvectors of T corresponding to the eigenvalue λ the **eigenspace** of T corresponding to the eigenvalue λ. It follows that the eigenspace of T corresponding to the eigenvalue λ is just the nullspace of $\lambda I - T$. In particular, it is clear that an eigenspace is a subspace of the vector space upon which T operates.

Example 3 Let T be the linear operator on P_n defined by $T(f)(x) = f(-x)$. In other words, T carries the polynomial $f(x)$ into the polynomial $f(-x)$. Since $T(1) = 1$, 1 is an eigenvalue of T. Since $T(x) = -x$, -1 is an eigenvalue of T. We determine the eigenspaces of T corresponding to the eigenvalues 1 and -1.

The eigenspace corresponding to the eigenvalue 1 consists of those polynomials f that satisfy $f(x) = f(-x)$, i.e., the even polynomials. The polynomials $1, x^2, x^4, \ldots$ form a basis for this space. The eigenspace corresponding to the eigenvalue -1 consists of those polynomials f that satisfy $f(-x) = -f(x)$, i.e., the odd polynomials. The polynomials $x, x^3, x^5, \ldots$ form a basis for this space.

In example 2, we saw that determining the eigenvectors corresponding to a given eigenvalue is simply a question of solving simultaneous linear equations. Next, we shall develop a method for determining which scalars are eigenvalues of a given linear operator.

Theorem 1 Let T be a linear operator on a real vector space V. Let λ be a real number. There is a nonzero vector $\boldsymbol{x}$ in V such that $T(\boldsymbol{x}) = \lambda\boldsymbol{x}$ if and only if $\det[T - \lambda I] = 0$.

PROOF Note that the nullspace of $T - \lambda I$ is nonzero if and only if $\det[T - \lambda I] = 0$. ■

An analogous result holds if the word "real" in the above theorem is replaced throughout by "complex."

The polynomial $f(x) = \det[xI - T]$ is called the **characteristic polynomial** of T and is denoted by C_T. It is calculated by choosing a basis for the vector space V and determining the matrix representation of the operator T relative to this basis. If A represents T relative to the chosen basis, $C_T(x) = \det[xI_n - A]$. If A' represents T with respect to another basis, in §5.8 we saw that there is an invertible matrix B such that $A' = BAB^{-1}$.

8. Let x_1, x_2, x_3, x_4 be four distinct points on the real line. If f and g lie in P_3, define $(f, g) = f(x_1)g(x_1) + f(x_2)g(x_2) + f(x_3)g(x_3) + f(x_4)g(x_4)$. Show that this is an inner product. If three points had been used instead of four, would it still be an inner product?

9. Let $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n$ be a basis for $\boldsymbol{R}^n$. Show that there is an inner product on $\boldsymbol{R}^n$ with respect to which $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n$ is an orthonormal basis.

10. Regarding the complex numbers as a vector space over the reals, define $(z_1, z_2) = \frac{1}{2}(z_1\bar{z}_2 + z_2\bar{z}_1)$.
   (a) Show that $(\ ,\ )$ is an inner product.
   (b) If $z_1 = \alpha_1 + \alpha_2 i$ and $z_2 = \beta_1 + \beta_2 i$, show that
   $$(z_1, z_2) = \alpha_1\beta_1 + \alpha_2\beta_2$$
   Show that (z, z) is the square of the absolute value of the complex number z in the usual sense.
   (c) Let M_a be the linear transformation of $\boldsymbol{C}^n$ into itself, defined by $M_a(z) = az$. Show that $(M_a z_1, M_a z_2) = a\bar{a}(z_1, z_2)$.
   (d) With M_a defined as in (c) show that M_a is an isometry if and only if $|a| = 1$.
   (e) Letting $T(z) = \bar{z}$, show that T is an isometry.

11. In the space of $n \times n$ matrices with the inner product of example 3 in the text, let E_{ij} be the $n \times n$ matrix with 0 in all entries except (i, j) and 1 in entry (i, j). Show that the matrices E_{ij}, $1 \leqslant i \leqslant n$ and $1 \leqslant j \leqslant n$, form an orthonormal basis for M_{nn}.

12. In the space of $n \times n$ matrices with the inner product of example 3 in the text, let T be the linear operator defined by $T(B) = AB$, where A is an $n \times n$ orthogonal matrix. Show that T is an isometry.

13. Let $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n$ be an orthonormal basis for $\boldsymbol{R}^n$ with the standard inner product. Show that the matrices $A_{ij} = x_i x_j^T$ form an orthonormal basis for the space of $n \times n$ matrices with the inner product of example 3 in the text.

14. In the space P_n of polynomials of degree not exceeding n with real coefficients in a variable x, consider the inner products
$$(f, g)_1 = \int_a^b f(x)g(x)\,dx$$
$$(f, g)_2 = \int_c^d f(x)g(x)\,dx$$
where $a < b$ and $c < d$. Let T be the function from P_n (regarded as an inner product space with the first inner product) into P_n (regarded as an inner product space with the second inner product), defined by
$$(T(f))(x) = \sqrt{\frac{b-a}{d-c}}\; f\left(\frac{b-a}{d-c}x + \frac{ad-bc}{d-c}\right)$$
Show that T is an isometry.

15. Consider the space of $n \times n$ matrices with the inner product of example 3 in the text.
   (a) Show that the orthogonal complement of the subspace of diagonal matrices is the subspaces of matrices all of whose diagonal entries are 0.
   (b) Show that the orthogonal complement of the subspace of symmetric matrices is the subspace of skew-symmetric matrices.

7

eigenvalues and canonical forms

1 EIGENVALUES AND EIGENVECTORS

In many of the applications of the theory of linear transformations, the following problem arises. Given a linear operator T on a vector space V, determine those scalars λ and those vectors $\boldsymbol{x}$ in V that satisfy the equation $T(\boldsymbol{x}) = \lambda\boldsymbol{x}$. This is known as the eigenvalue problem.

If T is a linear operator on a vector space V and $T(\boldsymbol{x}) = \lambda\boldsymbol{x}$ for some nonzero $\boldsymbol{x}$ in V and some scalar λ, λ is said to be an **eigenvalue** of T. The vector $\boldsymbol{x}$ is said to be an **eigenvector** of T corresponding to the eigenvalue λ. Often eigenvalues are called **characteristic values** and eigenvectors are called **characteristic vectors**. If A is an $n \times n$ matrix with real entries, a real number λ for which there is some nonzero real column n-vector $\boldsymbol{x}$ such that $A\boldsymbol{x} = \lambda\boldsymbol{x}$ is said to be an eigenvalue of A, and $\boldsymbol{x}$ is said to be an eigenvector of A. Complex eigenvalues and eigenvectors are defined similarly.

Intuitively, eigenvectors correspond to the fixed directions of the linear operator T.

Example 1 Let T be the linear operator on $\boldsymbol{R}^2$ defined by $T(\boldsymbol{x}) = \begin{bmatrix} 4 & -1 \\ 2 & 1 \end{bmatrix}\boldsymbol{x}$. Since $T\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 4 & -1 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = 3\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $T\left(\begin{bmatrix} 1 \\ 2 \end{bmatrix}\right) = \begin{bmatrix} 4 & -1 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = 2\begin{bmatrix} 1 \\ 2 \end{bmatrix}$, it follows that 3 and 2 are eigenvalues of T with corresponding eigenvectors $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$, respectively. In this case, the linear

linear algebra

SECOND EDITION

MICHAEL O'NAN

Rutgers University

HARCOURT BRACE JOVANOVICH, INC.

NEW YORK CHICAGO SAN FRANCISCO ATLANTA