CHAPTER 9

Diagonalization: Eigenvalues and Eigenvectors

9.1 Introduction

The ideas in this chapter can be discussed from two points of view.

Matrix Point of View

Suppose an n-square matrix A is given. The matrix A is said to be *diagonalizable* if there exists a nonsingular matrix P such that

$$B = P^{-1}AP$$

is diagonal. This chapter discusses the diagonalization of a matrix A. In particular, an algorithm is given to find the matrix P when it exists.

Linear Operator Point of View

Suppose a linear operator $T: V \to V$ is given. The linear operator T is said to be *diagonalizable* if there exists a basis S of V such that the matrix representation of T relative to the basis S is a diagonal matrix D. This chapter discusses conditions under which the linear operator T is diagonalizable.

Equivalence of the Two Points of View

The above two concepts are essentially the same. Specifically, a square matrix A may be viewed as a linear operator F defined by

$$F(X) = AX$$

where X is a column vector, and $B = P^{-1}AP$ represents F relative to a new coordinate system (basis) S whose elements are the columns of P. On the other hand, any linear operator T can be represented by a matrix A relative to one basis and, when a second basis is chosen, T is represented by the matrix

$$B = P^{-1}AP$$

where P is the change-of-basis matrix.

Most theorems will be stated in two ways: one in terms of matrices A and again in terms of linear mappings T.

Role of Underlying Field K

The underlying number field K did not play any special role in our previous discussions on vector spaces and linear mappings. However, the diagonalization of a matrix A or a linear operator T will depend on the roots of a polynomial $\Delta(t)$ over K, and these roots do depend on K. For example, suppose $\Delta(t) = t^2 + 1$. Then $\Delta(t)$ has no roots if $K = \mathbf{R}$, the real field; but $\Delta(t)$ has roots $\pm i$ if $K = \mathbf{C}$, the complex field. Furthermore, finding the roots of a polynomial with degree greater than two is a subject unto itself (frequently discussed in numerical analysis courses). Accordingly, our examples will usually lead to those polynomials $\Delta(t)$ whose roots can be easily determined.

9.2 Polynomials of Matrices

Consider a polynomial $f(t) = a_n t^n + \cdots + a_1 t + a_0$ over a field K. Recall (Section 2.8) that if A is any square matrix, then we define

$$f(A) = a_n A^n + \cdots + a_1 A + a_0 I$$

where I is the identity matrix. In particular, we say that A is a *root* of $f(t)$ if $f(A) = 0$, the zero matrix.

EXAMPLE 9.1 Let $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$. Then $A^2 = \begin{bmatrix} 7 & 10 \\ 15 & 22 \end{bmatrix}$. Let

$$f(t) = 2t^2 - 3t + 5 \qquad \text{and} \qquad g(t) = t^2 - 5t - 2$$

Then

$$f(A) = 2A^2 - 3A + 5I = \begin{bmatrix} 14 & 20 \\ 30 & 44 \end{bmatrix} + \begin{bmatrix} -3 & -6 \\ -9 & -12 \end{bmatrix} + \begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix} = \begin{bmatrix} 16 & 14 \\ 21 & 37 \end{bmatrix}$$

and

$$g(A) = A^2 - 5A - 2I = \begin{bmatrix} 7 & 10 \\ 15 & 22 \end{bmatrix} + \begin{bmatrix} -5 & -10 \\ -15 & -20 \end{bmatrix} + \begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Thus, A is a zero of $g(t)$.

The following theorem (proved in Problem 9.7) applies.

THEOREM 9.1: Let f and g be polynomials. For any square matrix A and scalar k,

(i) $(f + g)(A) = f(A) + g(A)$ (iii) $(kf)(A) = kf(A)$
(ii) $(fg)(A) = f(A)g(A)$ (iv) $f(A)g(A) = g(A)f(A)$.

Observe that (iv) tells us that any two polynomials in A commute.

Matrices and Linear Operators

Now suppose that $T: V \to V$ is a linear operator on a vector space V. Powers of T are defined by the composition operation:

$$T^2 = T \circ T, \qquad T^3 = T^2 \circ T, \qquad \ldots$$

Also, for any polynomial $f(t) = a_n t^n + \cdots + a_1 t + a_0$, we define $f(T)$ in the same way as we did for matrices:

$$f(T) = a_n T^n + \cdots + a_1 T + a_0 I$$

where I is now the identity mapping. We also say that T is a *zero* or *root* of $f(t)$ if $f(T) = 0$, the zero mapping. We note that the relations in Theorem 9.1 hold for linear operators as they do for matrices.

Remark: Suppose A is a matrix representation of a linear operator T. Then $f(A)$ is the matrix representation of $f(T)$, and, in particular, $f(T) = 0$ if and only if $f(A) = 0$.

9.3 Characteristic Polynomial, Cayley–Hamilton Theorem

Let $A = [a_{ij}]$ be an n-square matrix. The matrix $M = A - tI_n$, where I_n is the n-square identity matrix and t is an indeterminate, may be obtained by subtracting t down the diagonal of A. The negative of M is the matrix $tI_n - A$, and its determinant

$$\Delta(t) = \det(tI_n - A) = (-1)^n \det(A - tI_n)$$

which is a polynomial in t of degree n and is called the *characteristic polynomial* of A.

We state an important theorem in linear algebra (proved in Problem 9.8).

THEOREM 9.2: (Cayley–Hamilton) Every matrix A is a root of its characteristic polynomial.

Remark: Suppose $A = [a_{ij}]$ is a triangular matrix. Then $tI - A$ is a triangular matrix with diagonal entries $t - a_{ii}$; hence,

$$\Delta(t) = \det(tI - A) = (t - a_{11})(t - a_{22}) \cdots (t - a_{nn})$$

Observe that the roots of $\Delta(t)$ are the diagonal elements of A.

EXAMPLE 9.2 Let $A = \begin{bmatrix} 1 & 3 \\ 4 & 5 \end{bmatrix}$. Its characteristic polynomial is

$$\Delta(t) = |tI - A| = \begin{vmatrix} t-1 & -3 \\ -4 & t-5 \end{vmatrix} = (t-1)(t-5) - 12 = t^2 - 6t - 7$$

As expected from the Cayley–Hamilton theorem, A is a root of $\Delta(t)$; that is,

$$\Delta(A) = A^2 - 6A - 7I = \begin{bmatrix} 13 & 18 \\ 24 & 37 \end{bmatrix} + \begin{bmatrix} -6 & -18 \\ -24 & -30 \end{bmatrix} + \begin{bmatrix} -7 & 0 \\ 0 & -7 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Now suppose A and B are similar matrices, say $B = P^{-1}AP$, where P is invertible. We show that A and B have the same characteristic polynomial. Using $tI = P^{-1}tIP$, we have

$$\begin{aligned} \Delta_B(t) &= \det(tI - B) = \det(tI - P^{-1}AP) = \det(P^{-1}tIP - P^{-1}AP) \\ &= \det[P^{-1}(tI - A)P] = \det(P^{-1}) \det(tI - A) \det(P) \end{aligned}$$

Using the fact that determinants are scalars and commute and that $\det(P^{-1}) \det(P) = 1$, we finally obtain

$$\Delta_B(t) = \det(tI - A) = \Delta_A(t)$$

Thus, we have proved the following theorem.

THEOREM 9.3: Similar matrices have the same characteristic polynomial.

Characteristic Polynomials of Degrees 2 and 3

There are simple formulas for the characteristic polynomials of matrices of orders 2 and 3.

(a) Suppose $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$. Then

$$\Delta(t) = t^2 - (a_{11} + a_{22})t + \det(A) = t^2 - \operatorname{tr}(A)\, t + \det(A)$$

Here $\operatorname{tr}(A)$ denotes the trace of A—that is, the sum of the diagonal elements of A.

(b) Suppose $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$. Then

$$\Delta(t) = t^3 - \operatorname{tr}(A)\, t^2 + (A_{11} + A_{22} + A_{33})t - \det(A)$$

(Here A_{11}, A_{22}, A_{33} denote, respectively, the cofactors of a_{11}, a_{22}, a_{33}.)

EXAMPLE 9.3 Find the characteristic polynomial of each of the following matrices:

(a) $A = \begin{bmatrix} 5 & 3 \\ 2 & 10 \end{bmatrix}$, (b) $B = \begin{bmatrix} 7 & -1 \\ 6 & 2 \end{bmatrix}$, (c) $C = \begin{bmatrix} 5 & -2 \\ 4 & -4 \end{bmatrix}$.

(a) We have $\text{tr}(A) = 5 + 10 = 15$ and $|A| = 50 - 6 = 44$; hence, $\Delta(t) + t^2 - 15t + 44$.

(b) We have $\text{tr}(B) = 7 + 2 = 9$ and $|B| = 14 + 6 = 20$; hence, $\Delta(t) = t^2 - 9t + 20$.

(c) We have $\text{tr}(C) = 5 - 4 = 1$ and $|C| = -20 + 8 = -12$; hence, $\Delta(t) = t^2 - t - 12$.

EXAMPLE 9.4 Find the characteristic polynomial of $A = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 3 & 2 \\ 1 & 3 & 9 \end{bmatrix}$.

We have $\text{tr}(A) = 1 + 3 + 9 = 13$. The cofactors of the diagonal elements are as follows:

$$A_{11} = \begin{vmatrix} 3 & 2 \\ 3 & 9 \end{vmatrix} = 21, \qquad A_{22} = \begin{vmatrix} 1 & 2 \\ 1 & 9 \end{vmatrix} = 7, \qquad A_{33} = \begin{vmatrix} 1 & 1 \\ 0 & 3 \end{vmatrix} = 3$$

Thus, $A_{11} + A_{22} + A_{33} = 31$. Also, $|A| = 27 + 2 + 0 - 6 - 6 - 0 = 17$. Accordingly,

$$\Delta(t) = t^3 - 13t^2 + 31t - 17$$

Remark: The coefficients of the characteristic polynomial $\Delta(t)$ of the 3-square matrix A are, with alternating signs, as follows:

$$S_1 = \text{tr}(A), \qquad S_2 = A_{11} + A_{22} + A_{33}, \qquad S_3 = \det(A)$$

We note that each S_k is the sum of all principal minors of A of order k.

The next theorem, whose proof lies beyond the scope of this text, tells us that this result is true in general.

THEOREM 9.4: Let A be an n-square matrix. Then its characteristic polynomial is

$$\Delta(t) = t^n - S_1 t^{n-1} + S_2 t^{n-2} + \cdots + (-1)^n S_n$$

where S_k is the sum of the principal minors of order k.

Characteristic Polynomial of a Linear Operator

Now suppose $T: V \to V$ is a linear operator on a vector space V of finite dimension. We define the *characteristic polynomial* $\Delta(t)$ of T to be the characteristic polynomial of any matrix representation of T. Recall that if A and B are matrix representations of T, then $B = P^{-1}AP$, where P is a change-of-basis matrix. Thus, A and B are similar, and by Theorem 9.3, A and B have the same characteristic polynomial. Accordingly, the characteristic polynomial of T is independent of the particular basis in which the matrix representation of T is computed.

Because $f(T) = 0$ if and only if $f(A) = 0$, where $f(t)$ is any polynomial and A is any matrix representation of T, we have the following analogous theorem for linear operators.

THEOREM 9.2′: (Cayley–Hamilton) A linear operator T is a zero of its characteristic polynomial.

9.4 Diagonalization, Eigenvalues and Eigenvectors

Let A be any n-square matrix. Then A can be represented by (or is similar to) a diagonal matrix $D = \text{diag}(k_1, k_2, \ldots, k_n)$ if and only if there exists a basis S consisting of (column) vectors $u_1, u_2, \ldots, u_n$ such that

$$\begin{aligned} Au_1 &= k_1u_1 \\ Au_2 &= \qquad k_2u_2 \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ Au_n &= \qquad\qquad k_nu_n \end{aligned}$$

In such a case, A is said to be *diagonizable*. Furthermore, $D = P^{-1}AP$, where P is the nonsingular matrix whose columns are, respectively, the basis vectors $u_1, u_2, \ldots, u_n$.

The above observation leads us to the following definition.

DEFINITION: Let A be any square matrix. A scalar λ is called an *eigenvalue* of A if there exists a nonzero (column) vector v such that

$$Av = \lambda v$$

Any vector satisfying this relation is called an *eigenvector* of A *belonging* to the eigenvalue λ.

We note that each scalar multiple kv of an eigenvector v belonging to λ is also such an eigenvector, because

$$A(kv) = k(Av) = k(\lambda v) = \lambda(kv)$$

The set E_λ of all such eigenvectors is a subspace of V (Problem 9.19), called the *eigenspace* of λ. (If $\dim E_\lambda = 1$, then E_λ is called an *eigenline* and λ is called a *scaling factor*.)

The terms *characteristic value* and *characteristic vector* (or *proper value* and *proper vector*) are sometimes used instead of eigenvalue and eigenvector.

The above observation and definitions give us the following theorem.

THEOREM 9.5: An n-square matrix A is similar to a diagonal matrix D if and only if A has n linearly independent eigenvectors. In this case, the diagonal elements of D are the corresponding eigenvalues and $D = P^{-1}AP$, where P is the matrix whose columns are the eigenvectors.

Suppose a matrix A can be diagonalized as above, say $P^{-1}AP = D$, where D is diagonal. Then A has the extremely useful *diagonal factorization*:

$$A = PDP^{-1}$$

Using this factorization, the algebra of A reduces to the algebra of the diagonal matrix D, which can be easily calculated. Specifically, suppose $D = \text{diag}(k_1, k_2, \ldots, k_n)$. Then

$$A^m = (PDP^{-1})^m = PD^mP^{-1} = P\,\text{diag}(k_1^m, \ldots, k_n^m)P^{-1}$$

More generally, for any polynomial $f(t)$,

$$f(A) = f(PDP^{-1}) = Pf(D)P^{-1} = P\,\text{diag}(f(k_1), f(k_2), \ldots, f(k_n))P^{-1}$$

Furthermore, if the diagonal entries of D are nonnegative, let

$$B = P\,\text{diag}(\sqrt{k_1}, \sqrt{k_2}, \ldots, \sqrt{k_n})\,P^{-1}$$

Then B is a *nonnegative square root* of A; that is, $B^2 = A$ and the eigenvalues of B are nonnegative.

EXAMPLE 9.5 Let $A = \begin{bmatrix} 3 & 1 \\ 2 & 2 \end{bmatrix}$ and let $v_1 = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$ and $v_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Then

$$Av_1 = \begin{bmatrix} 3 & 1 \\ 2 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ -2 \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \end{bmatrix} = v_1 \qquad \text{and} \qquad Av_2 = \begin{bmatrix} 3 & 1 \\ 2 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \end{bmatrix} = 4v_2$$

Thus, v_1 and v_2 are eigenvectors of A belonging, respectively, to the eigenvalues $\lambda_1 = 1$ and $\lambda_2 = 4$. Observe that v_1 and v_2 are linearly independent and hence form a basis of $\mathbf{R}^2$. Accordingly, A is diagonalizable. Furthermore, let P be the matrix whose columns are the eigenvectors v_1 and v_2. That is, let

$$P = \begin{bmatrix} 1 & 1 \\ -2 & 1 \end{bmatrix}, \qquad \text{and so} \qquad P^{-1} = \begin{bmatrix} \frac{1}{3} & -\frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix}$$

Then A is similar to the diagonal matrix

$$D = P^{-1}AP = \begin{bmatrix} \frac{1}{3} & -\frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix}\begin{bmatrix} 3 & 1 \\ 2 & 2 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ -2 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}$$

As expected, the diagonal elements 1 and 4 in D are the eigenvalues corresponding, respectively, to the eigenvectors v_1 and v_2, which are the columns of P. In particular, A has the factorization

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

Accordingly,

$$A^4 = \begin{bmatrix} 1 & 1 \\ -2 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 256 \end{bmatrix}\begin{bmatrix} \frac{1}{3} & -\frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix} = \begin{bmatrix} 171 & 85 \\ 170 & 86 \end{bmatrix}$$

Moreover, suppose $f(t) = t^3 - 5t^2 + 3t + 6$; hence, $f(1) = 5$ and $f(4) = 2$. Then

$$f(A) = Pf(D)P^{-1} = \begin{bmatrix} 1 & 1 \\ -2 & 1 \end{bmatrix}\begin{bmatrix} 5 & 0 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} \frac{1}{3} & -\frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix} = \begin{bmatrix} 3 & -1 \\ -2 & 4 \end{bmatrix}$$

Last, we obtain a ''positive square root'' of A. Specifically, using $\sqrt{1} = 1$ and $\sqrt{4} = 2$, we obtain the matrix

$$B = P\sqrt{D}P^{-1} = \begin{bmatrix} 1 & 1 \\ -2 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} \frac{1}{3} & -\frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix} = \begin{bmatrix} \frac{5}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{4}{3} \end{bmatrix}$$

where $B^2 = A$ and where B has positive eigenvalues 1 and 2.

Remark: Throughout this chapter, we use the following fact:

$$\text{If } P = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \text{then } P^{-1} = \begin{bmatrix} d/|P| & -b/|P| \\ -c/|P| & a/|P| \end{bmatrix}.$$

That is, P^{-1} is obtained by interchanging the diagonal elements a and d of P, taking the negatives of the nondiagonal elements b and c, and dividing each element by the determinant $|P|$.

Properties of Eigenvalues and Eigenvectors

Example 9.5 indicates the advantages of a diagonal representation (factorization) of a square matrix. In the following theorem (proved in Problem 9.20), we list properties that help us to find such a representation.

THEOREM 9.6: Let A be a square matrix. Then the following are equivalent.

(i) A scalar λ is an eigenvalue of A.

(ii) The matrix $M = A - \lambda I$ is singular.

(iii) The scalar λ is a root of the characteristic polynomial $\Delta(t)$ of A.

The eigenspace E_λ of an eigenvalue λ is the solution space of the homogeneous system $MX = 0$, where $M = A - \lambda I$; that is, M is obtained by subtracting λ down the diagonal of A.

Some matrices have no eigenvalues and hence no eigenvectors. However, using Theorem 9.6 and the Fundamental Theorem of Algebra (every polynomial over the complex field **C** has a root), we obtain the following result.

THEOREM 9.7: Let A be a square matrix over the complex field **C**. Then A has at least one eigenvalue.

The following theorems will be used subsequently. (The theorem equivalent to Theorem 9.8 for linear operators is proved in Problem 9.21, and Theorem 9.9 is proved in Problem 9.22.)

THEOREM 9.8: Suppose $v_1, v_2, \ldots, v_n$ are nonzero eigenvectors of a matrix A belonging to distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$. Then $v_1, v_2, \ldots, v_n$ are linearly independent.

THEOREM 9.9: Suppose the characteristic polynomial $\Delta(t)$ of an n-square matrix A is a product of n distinct factors, say, $\Delta(t) = (t - a_1)(t - a_2) \cdots (t - a_n)$. Then A is similar to the diagonal matrix $D = \text{diag}(a_1, a_2, \ldots, a_n)$.

If λ is an eigenvalue of a matrix A, then the *algebraic multiplicity* of λ is defined to be the multiplicity of λ as a root of the characteristic polynomial of A, and the *geometric multiplicity* of λ is defined to be the dimension of its eigenspace, dim E_λ. The following theorem (whose equivalent for linear operators is proved in Problem 9.23) holds.

THEOREM 9.10: The geometric multiplicity of an eigenvalue λ of a matrix A does not exceed its algebraic multiplicity.

Diagonalization of Linear Operators

Consider a linear operator $T: V \to V$. Then T is said to be *diagonalizable* if it can be represented by a diagonal matrix D. Thus, T is diagonalizable if and only if there exists a basis $S = \{u_1, u_2, \ldots, u_n\}$ of V for which

$$\begin{aligned} T(u_1) &= k_1 u_1 \\ T(u_2) &= \qquad k_2 u_2 \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ T(u_n) &= \qquad\qquad k_n u_n \end{aligned}$$

In such a case, T is represented by the diagonal matrix

$$D = \text{diag}(k_1, k_2, \ldots, k_n)$$

relative to the basis S.

The above observation leads us to the following definitions and theorems, which are analogous to the definitions and theorems for matrices discussed above.

DEFINITION: Let T be a linear operator. A scalar λ is called an *eigenvalue* of T if there exists a nonzero vector v such that $T(v) = \lambda v$.

Every vector satisfying this relation is called an *eigenvector* of T *belonging* to the eigenvalue λ.

The set E_λ of all eigenvectors belonging to an eigenvalue λ is a subspace of V, called the *eigenspace* of λ. (Alternatively, λ is an eigenvalue of T if $\lambda I - T$ is singular, and, in this case, E_λ is the kernel of $\lambda I - T$.) The *algebraic* and *geometric multiplicities* of an eigenvalue λ of a linear operator T are defined in the same way as those of an eigenvalue of a matrix A.

The following theorems apply to a linear operator T on a vector space V of finite dimension.

THEOREM 9.5′: T can be represented by a diagonal matrix D if and only if there exists a basis S of V consisting of eigenvectors of T. In this case, the diagonal elements of D are the corresponding eigenvalues.

THEOREM 9.6′: Let T be a linear operator. Then the following are equivalent:

(i) A scalar λ is an eigenvalue of T.

(ii) The linear operator $\lambda I - T$ is singular.

(iii) The scalar λ is a root of the characteristic polynomial $\Delta(t)$ of T.

THEOREM 9.7′: Suppose V is a complex vector space. Then T has at least one eigenvalue.

THEOREM 9.8′: Suppose $v_1, v_2, \ldots, v_n$ are nonzero eigenvectors of a linear operator T belonging to distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$. Then $v_1, v_2, \ldots, v_n$ are linearly independent.

THEOREM 9.9′: Suppose the characteristic polynomial $\Delta(t)$ of T is a product of n distinct factors, say, $\Delta(t) = (t - a_1)(t - a_2) \cdots (t - a_n)$. Then T can be represented by the diagonal matrix $D = \mathrm{diag}(a_1, a_2, \ldots, a_n)$.

THEOREM 9.10′: The geometric multiplicity of an eigenvalue λ of T does not exceed its algebraic multiplicity.

Remark: The following theorem reduces the investigation of the diagonalization of a linear operator T to the diagonalization of a matrix A.

THEOREM 9.11: Suppose A is a matrix representation of T. Then T is diagonalizable if and only if A is diagonalizable.

9.5 Computing Eigenvalues and Eigenvectors, Diagonalizing Matrices

This section gives an algorithm for computing eigenvalues and eigenvectors for a given square matrix A and for determining whether or not a nonsingular matrix P exists such that $P^{-1}AP$ is diagonal.

ALGORITHM 9.1: (Diagonalization Algorithm) The input is an n-square matrix A.

Step 1. Find the characteristic polynomial $\Delta(t)$ of A.

Step 2. Find the roots of $\Delta(t)$ to obtain the eigenvalues of A.

Step 3. Repeat (a) and (b) for each eigenvalue λ of A.

(a) Form the matrix $M = A - \lambda I$ by subtracting λ down the diagonal of A.

(b) Find a basis for the solution space of the homogeneous system $MX = 0$. (These basis vectors are linearly independent eigenvectors of A belonging to λ.)

Step 4. Consider the collection $S = \{v_1, v_2, \ldots, v_m\}$ of all eigenvectors obtained in Step 3.

(a) If $m \neq n$, then A is not diagonalizable.

(b) If $m = n$, then A is diagonalizable. Specifically, let P be the matrix whose columns are the eigenvectors $v_1, v_2, \ldots, v_n$. Then

$$D = P^{-1}AP = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$$

where λ_i is the eigenvalue corresponding to the eigenvector v_i.

EXAMPLE 9.6 The diagonalizable algorithm is applied to $A = \begin{bmatrix} 4 & 2 \\ 3 & -1 \end{bmatrix}$.

(1) The characteristic polynomial $\Delta(t)$ of A is computed. We have

$$\operatorname{tr}(A) = 4 - 1 = -3, \qquad |A| = -4 - 6 = -10;$$

hence,

$$\Delta(t) = t^2 - 3t - 10 = (t-5)(t+2)$$

(2) Set $\Delta(t) = (t-5)(t+2) = 0$. The roots $\lambda_1 = 5$ and $\lambda_2 = -2$ are the eigenvalues of A.

(3) (i) We find an eigenvector v_1 of A belonging to the eigenvalue $\lambda_1 = 5$. Subtract $\lambda_1 = 5$ down the diagonal of A to obtain the matrix $M = \begin{bmatrix} -1 & 2 \\ 3 & -6 \end{bmatrix}$. The eigenvectors belonging to $\lambda_1 = 5$ form the solution of the homogeneous system $MX = 0$; that is,

$$\begin{bmatrix} -1 & 2 \\ 3 & -6 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad \text{or} \qquad \begin{aligned} -x + 2y &= 0 \\ 3x - 6y &= 0 \end{aligned} \qquad \text{or} \qquad -x + 2y = 0$$

The system has only one free variable. Thus, a nonzero solution, for example, $v_1 = (2, 1)$, is an eigenvector that spans the eigenspace of $\lambda_1 = 5$.

(ii) We find an eigenvector v_2 of A belonging to the eigenvalue $\lambda_2 = -2$. Subtract -2 (or add 2) down the diagonal of A to obtain the matrix

$$M = \begin{bmatrix} 6 & 2 \\ 3 & 1 \end{bmatrix} \text{ and the homogenous system } \begin{aligned} 6x + 2y &= 0 \\ 3x + y &= 0 \end{aligned} \qquad \text{or} \qquad 3x + y = 0.$$

The system has only one independent solution. Thus, a nonzero solution, say $v_2 = (-1, 3)$, is an eigenvector that spans the eigenspace of $\lambda_2 = -2$.

(4) Let P be the matrix whose columns are the eigenvectors v_1 and v_2. Then

$$P = \begin{bmatrix} 2 & -1 \\ 1 & 3 \end{bmatrix}, \qquad \text{and so} \qquad P^{-1} = \begin{bmatrix} \frac{3}{7} & \frac{1}{7} \\ -\frac{1}{7} & \frac{2}{7} \end{bmatrix}$$

Accordingly, $D = P^{-1}AP$ is the diagonal matrix whose diagonal entries are the corresponding eigenvalues; that is,

$$D = P^{-1}AP = \begin{bmatrix} \frac{3}{7} & \frac{1}{7} \\ -\frac{1}{7} & \frac{2}{7} \end{bmatrix}\begin{bmatrix} 4 & 2 \\ 3 & -1 \end{bmatrix}\begin{bmatrix} 2 & -1 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 5 & 0 \\ 0 & -2 \end{bmatrix}$$

EXAMPLE 9.7 Consider the matrix $B = \begin{bmatrix} 5 & -1 \\ 1 & 3 \end{bmatrix}$. We have

$$\operatorname{tr}(B) = 5 + 3 = 8, \qquad |B| = 15 + 1 = 16; \qquad \text{so} \qquad \Delta(t) = t^2 - 8t + 16 = (t-4)^2$$

Accordingly, $\lambda = 4$ is the only eigenvalue of B.

Subtract $\lambda = 4$ down the diagonal of B to obtain the matrix

$$M = \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix} \text{ and the homogeneous system } \quad \begin{aligned} x - y &= 0 \\ x - y &= 0 \end{aligned} \qquad \text{or} \qquad x - y = 0$$

The system has only one independent solution; for example, $x = 1, y = 1$. Thus, $v = (1, 1)$ and its multiples are the only eigenvectors of B. Accordingly, B is not diagonalizable, because there does not exist a basis consisting of eigenvectors of B.

EXAMPLE 9.8 Consider the matrix $A = \begin{bmatrix} 3 & -5 \\ 2 & -3 \end{bmatrix}$. Here $\operatorname{tr}(A) = 3 - 3 = 0$ and $|A| = -9 + 10 = 1$. Thus, $\Delta(t) = t^2 + 1$ is the characteristic polynomial of A. We consider two cases:

(a) A is a matrix over the real field $\mathbf{R}$. Then $\Delta(t)$ has no (real) roots. Thus, A has no eigenvalues and no eigenvectors, and so A is not diagonalizable.

(b) A is a matrix over the complex field $\mathbf{C}$. Then $\Delta(t) = (t - i)(t + i)$ has two roots, i and $-i$. Thus, A has two distinct eigenvalues i and $-i$, and hence, A has two independent eigenvectors. Accordingly there exists a nonsingular matrix P over the complex field $\mathbf{C}$ for which

$$P^{-1}AP = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}$$

Therefore, A is diagonalizable (over $\mathbf{C}$).

9.6 Diagonalizing Real Symmetric Matrices and Quadratic Forms

There are many real matrices A that are not diagonalizable. In fact, some real matrices may not have any (real) eigenvalues. However, if A is a real symmetric matrix, then these problems do not exist. Namely, we have the following theorems.

THEOREM 9.12: Let A be a real symmetric matrix. Then each root λ of its characteristic polynomial is real.

THEOREM 9.13: Let A be a real symmetric matrix. Suppose u and v are eigenvectors of A belonging to distinct eigenvalues λ_1 and λ_2. Then u and v are orthogonal, that; is, $\langle u, v \rangle = 0$.

The above two theorems give us the following fundamental result.

THEOREM 9.14: Let A be a real symmetric matrix. Then there exists an orthogonal matrix P such that $D = P^{-1}AP$ is diagonal.

The orthogonal matrix P is obtained by normalizing a basis of orthogonal eigenvectors of A as illustrated below. In such a case, we say that A is ''orthogonally diagonalizable.''

EXAMPLE 9.9 Let $A = \begin{bmatrix} 2 & -2 \\ -2 & 5 \end{bmatrix}$, a real symmetric matrix. Find an orthogonal matrix P such that $P^{-1}AP$ is diagonal.

First we find the characteristic polynomial $\Delta(t)$ of A. We have

$$\operatorname{tr}(A) = 2 + 5 = 7, \qquad |A| = 10 - 4 = 6; \qquad \text{so} \qquad \Delta(t) = t^2 - 7t + 6 = (t - 6)(t - 1)$$

Accordingly, $\lambda_1 = 6$ and $\lambda_2 = 1$ are the eigenvalues of A.

(a) Subtracting $\lambda_1 = 6$ down the diagonal of A yields the matrix

$$M = \begin{bmatrix} -4 & -2 \\ -2 & -1 \end{bmatrix} \qquad \text{and the homogeneous system} \qquad \begin{aligned} -4x - 2y &= 0 \\ -2x - y &= 0 \end{aligned} \qquad \text{or} \qquad 2x + y = 0$$

A nonzero solution is $u_1 = (1, -2)$.

(b) Subtracting $\lambda_2 = 1$ down the diagonal of A yields the matrix

$$M = \begin{bmatrix} 1 & -2 \\ -2 & 4 \end{bmatrix} \qquad \text{and the homogeneous system} \qquad x - 2y = 0$$

(The second equation drops out, because it is a multiple of the first equation.) A nonzero solution is $u_2 = (2, 1)$.

As expected from Theorem 9.13, u_1 and u_2 are orthogonal. Normalizing u_1 and u_2 yields the orthonormal vectors

$$\hat{u}_1 = (1/\sqrt{5}, -2/\sqrt{5}) \qquad \text{and} \qquad \hat{u}_2 = (2/\sqrt{5}, 1/\sqrt{5})$$

Finally, let P be the matrix whose columns are $\hat{u}_1$ and $\hat{u}_2$, respectively. Then

$$P = \begin{bmatrix} 1/\sqrt{5} & 2/\sqrt{5} \\ -2/\sqrt{5} & 1/\sqrt{5} \end{bmatrix} \qquad \text{and} \qquad P^{-1}AP = \begin{bmatrix} 6 & 0 \\ 0 & 1 \end{bmatrix}$$

As expected, the diagonal entries of $P^{-1}AP$ are the eigenvalues corresponding to the columns of P.

The procedure in the above Example 9.9 is formalized in the following algorithm, which finds an orthogonal matrix P such that $P^{-1}AP$ is diagonal.

ALGORITHM 9.2: (Orthogonal Diagonalization Algorithm) The input is a real symmetric matrix A.

Step 1. Find the characteristic polynomial $\Delta(t)$ of A.

Step 2. Find the eigenvalues of A, which are the roots of $\Delta(t)$.

Step 3. For each eigenvalue λ of A in Step 2, find an orthogonal basis of its eigenspace.

Step 4. Normalize all eigenvectors in Step 3, which then forms an orthonormal basis of $\mathbf{R}^n$.

Step 5. Let P be the matrix whose columns are the normalized eigenvectors in Step 4.

Application to Quadratic Forms

Let q be a real polynomial in variables $x_1, x_2, \ldots, x_n$ such that every term in q has degree two; that is,

$$q(x_1, x_2, \ldots, x_n) = \sum_i c_i x_i^2 + \sum_{i<j} d_{ij} x_i x_j, \qquad \text{where} \qquad c_i, d_{ij} \in \mathbf{R}$$

Then q is called a *quadratic form*. If there are no cross-product terms $x_i x_j$ (i.e., all $d_{ij} = 0$), then q is said to be *diagonal*.

The above quadratic form q determines a real symmetric matrix $A = [a_{ij}]$, where $a_{ii} = c_i$ and $a_{ij} = a_{ji} = \frac{1}{2} d_{ij}$. Namely, q can be written in the matrix form

$$q(X) = X^T A X$$

where $X = [x_1, x_2, \ldots, x_n]^T$ is the column vector of the variables. Furthermore, suppose $X = PY$ is a linear substitution of the variables. Then substitution in the quadratic form yields

$$q(Y) = (PY)^T A (PY) = Y^T (P^T A P) Y$$

Thus, $P^T AP$ is the matrix representation of q in the new variables.

We seek an orthogonal matrix P such that the *orthogonal substitution* $X = PY$ yields a diagonal quadratic form for which $P^T AP$ is diagonal. Because P is orthogonal, $P^T = P^{-1}$, and hence, $P^T AP = P^{-1}AP$. The above theory yields such an orthogonal matrix P.

EXAMPLE 9.10 Consider the quadratic form

$$q(x,y) = 2x^2 - 4xy + 5y^2 = X^T AX, \qquad \text{where} \qquad A = \begin{bmatrix} 2 & -2 \\ -2 & 5 \end{bmatrix} \qquad \text{and} \qquad X = \begin{bmatrix} x \\ y \end{bmatrix}$$

By Example 9.9,

$$P^{-1}AP = \begin{bmatrix} 6 & 0 \\ 0 & 1 \end{bmatrix} = P^T AP, \qquad \text{where} \qquad P = \begin{bmatrix} 1/\sqrt{5} & 2/\sqrt{5} \\ -2/\sqrt{5} & 1/\sqrt{5} \end{bmatrix}$$

Let $Y = [s,t]^T$. Then matrix P corresponds to the following linear orthogonal substitution $x = PY$ of the variables x and y in terms of the variables s and t:

$$x = \frac{1}{\sqrt{5}}s + \frac{2}{\sqrt{5}}t, \qquad y = -\frac{2}{\sqrt{5}}s + \frac{1}{\sqrt{5}}t$$

This substitution in $q(x,y)$ yields the diagonal quadratic form $q(s,t) = 6s^2 + t^2$.

9.7 Minimal Polynomial

Let A be any square matrix. Let $J(A)$ denote the collection of all polynomials $f(t)$ for which A is a root—that is, for which $f(A) = 0$. The set $J(A)$ is not empty, because the Cayley–Hamilton Theorem 9.1 tells us that the characteristic polynomial $\Delta_A(t)$ of A belongs to $J(A)$. Let $m(t)$ denote the monic polynomial of lowest degree in $J(A)$. (Such a polynomial $m(t)$ exists and is unique.) We call $m(t)$ the *minimal polynomial* of the matrix A.

Remark: A polynomial $f(t) \neq 0$ is *monic* if its leading coefficient equals one.

The following theorem (proved in Problem 9.33) holds.

THEOREM 9.15: The minimal polynomial $m(t)$ of a matrix (linear operator) A divides every polynomial that has A as a zero. In particular, $m(t)$ divides the characteristic polynomial $\Delta(t)$ of A.

There is an even stronger relationship between $m(t)$ and $\Delta(t)$.

THEOREM 9.16: The characteristic polynomial $\Delta(t)$ and the minimal polynomial $m(t)$ of a matrix A have the same irreducible factors.

This theorem (proved in Problem 9.35) does not say that $m(t) = \Delta(t)$, only that any irreducible factor of one must divide the other. In particular, because a linear factor is irreducible, $m(t)$ and $\Delta(t)$ have the same linear factors. Hence, they have the same roots. Thus, we have the following theorem.

THEOREM 9.17: A scalar λ is an eigenvalue of the matrix A if and only if λ is a root of the minimal polynomial of A.

EXAMPLE 9.11 Find the minimal polynomial $m(t)$ of $A = \begin{bmatrix} 2 & 2 & -5 \\ 3 & 7 & -15 \\ 1 & 2 & -4 \end{bmatrix}$.

First find the characteristic polynomial $\Delta(t)$ of A. We have

$$\operatorname{tr}(A) = 5, \qquad A_{11} + A_{22} + A_{33} = 2 - 3 + 8 = 7, \qquad \text{and} \qquad |A| = 3$$

Hence,

$$\Delta(t) = t^3 - 5t^2 + 7t - 3 = (t-1)^2(t-3)$$

The minimal polynomial $m(t)$ must divide $\Delta(t)$. Also, each irreducible factor of $\Delta(t)$ (i.e., $t-1$ and $t-3$) must also be a factor of $m(t)$. Thus, $m(t)$ is exactly one of the following:

$$f(t) = (t-3)(t-1) \qquad \text{or} \qquad g(t) = (t-3)(t-1)^2$$

We know, by the Cayley–Hamilton theorem, that $g(A) = \Delta(A) = 0$. Hence, we need only test $f(t)$. We have

$$f(A) = (A-I)(A-3I) = \begin{bmatrix} 1 & 2 & -5 \\ 3 & 6 & -15 \\ 1 & 2 & -5 \end{bmatrix} \begin{bmatrix} -1 & 2 & -5 \\ 3 & 4 & -15 \\ 1 & 2 & -7 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Thus, $f(t) = m(t) = (t-1)(t-3) = t^2 - 4t + 3$ is the minimal polynomial of A.

EXAMPLE 9.12

(a) Consider the following two r-square matrices, where $a \neq 0$:

$$J(\lambda, r) = \begin{bmatrix} \lambda & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda & 1 & \dots & 0 & 0 \\ \multicolumn{6}{c}{\dotfill} \\ 0 & 0 & 0 & \dots & \lambda & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda \end{bmatrix} \qquad \text{and} \qquad A = \begin{bmatrix} \lambda & a & 0 & \dots & 0 & 0 \\ 0 & \lambda & a & \dots & 0 & 0 \\ \multicolumn{6}{c}{\dotfill} \\ 0 & 0 & 0 & \dots & \lambda & a \\ 0 & 0 & 0 & \dots & 0 & \lambda \end{bmatrix}$$

The first matrix, called a Jordan Block, has λ's on the diagonal, 1's on the *superdiagonal* (consisting of the entries above the diagonal entries), and 0's elsewhere. The second matrix A has λ's on the diagonal, a's on the superdiagonal, and 0's elsewhere. [Thus, A is a generalization of $J(\lambda, r)$.] One can show that

$$f(t) = (t-\lambda)^r$$

is both the characteristic and minimal polynomial of both $J(\lambda, r)$ and A.

(b) Consider an arbitrary monic polynomial:

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

Let $C(f)$ be the n-square matrix with 1's on the *subdiagonal* (consisting of the entries below the diagonal entries), the negatives of the coefficients in the last column, and 0's elsewhere as follows:

$$C(f) = \begin{bmatrix} 0 & 0 & \dots & 0 & -a_0 \\ 1 & 0 & \dots & 0 & -a_1 \\ 0 & 1 & \dots & 0 & -a_2 \\ \multicolumn{5}{c}{\dotfill} \\ 0 & 0 & \dots & 1 & -a_{n-1} \end{bmatrix}$$

Then $C(f)$ is called the *companion matrix* of the polynomial $f(t)$. Moreover, the minimal polynomial $m(t)$ and the characteristic polynomial $\Delta(t)$ of the companion matrix $C(f)$ are both equal to the original polynomial $f(t)$.

Minimal Polynomial of a Linear Operator

The *minimal polynomial* $m(t)$ of a linear operator T is defined to be the monic polynomial of lowest degree for which T is a root. However, for any polynomial $f(t)$, we have

$$f(T) = 0 \qquad \text{if and only if} \qquad f(A) = 0$$

where A is any matrix representation of T. Accordingly, T and A have the same minimal polynomials. Thus, the above theorems on the minimal polynomial of a matrix also hold for the minimal polynomial of a linear operator. That is, we have the following theorems.

THEOREM 9.15′: The minimal polynomial $m(t)$ of a linear operator T divides every polynomial that has T as a root. In particular, $m(t)$ divides the characteristic polynomial $\Delta(t)$ of T.

THEOREM 9.16′: The characteristic and minimal polynomials of a linear operator T have the same irreducible factors.

THEOREM 9.17′: A scalar λ is an eigenvalue of a linear operator T if and only if λ is a root of the minimal polynomial $m(t)$ of T.

9.8 Characteristic and Minimal Polynomials of Block Matrices

This section discusses the relationship of the characteristic polynomial and the minimal polynomial to certain (square) block matrices.

Characteristic Polynomial and Block Triangular Matrices

Suppose M is a block triangular matrix, say $M = \begin{bmatrix} A_1 & B \\ 0 & A_2 \end{bmatrix}$, where A_1 and A_2 are square matrices. Then $tI - M$ is also a block triangular matrix, with diagonal blocks $tI - A_1$ and $tI - A_2$. Thus,

$$|tI - M| = \begin{vmatrix} tI - A_1 & -B \\ 0 & tI - A_2 \end{vmatrix} = |tI - A_1||tI - A_2|$$

That is, the characteristic polynomial of M is the product of the characteristic polynomials of the diagonal blocks A_1 and A_2.

By induction, we obtain the following useful result.

THEOREM 9.18: Suppose M is a block triangular matrix with diagonal blocks $A_1, A_2, \ldots, A_r$. Then the characteristic polynomial of M is the product of the characteristic polynomials of the diagonal blocks A_i; that is,

$$\Delta_M(t) = \Delta_{A_1}(t)\Delta_{A_2}(t)\ldots\Delta_{A_r}(t)$$

EXAMPLE 9.13 Consider the matrix $M = \left[\begin{array}{cc|cc} 9 & -1 & 5 & 7 \\ 8 & 3 & 2 & -4 \\ \hline 0 & 0 & 3 & 6 \\ 0 & 0 & -1 & 8 \end{array}\right]$.

Then M is a block triangular matrix with diagonal blocks $A = \begin{bmatrix} 9 & -1 \\ 8 & 3 \end{bmatrix}$ and $B = \begin{bmatrix} 3 & 6 \\ -1 & 8 \end{bmatrix}$. Here

$$\begin{array}{llll} \operatorname{tr}(A) = 9 + 3 = 12, & \det(A) = 27 + 8 = 35, & \text{and so} & \Delta_A(t) = t^2 - 12t + 35 = (t-5)(t-7) \\ \operatorname{tr}(B) = 3 + 8 = 11, & \det(B) = 24 + 6 = 30, & \text{and so} & \Delta_B(t) = t^2 - 11t + 30 = (t-5)(t-6) \end{array}$$

Accordingly, the characteristic polynomial of M is the product

$$\Delta_M(t) = \Delta_A(t)\Delta_B(t) = (t-5)^2(t-6)(t-7)$$

Minimal Polynomial and Block Diagonal Matrices

The following theorem (proved in Problem 9.36) holds.

THEOREM 9.19: Suppose M is a block diagonal matrix with diagonal blocks $A_1, A_2, \ldots, A_r$. Then the minimal polynomial of M is equal to the least common multiple (LCM) of the minimal polynomials of the diagonal blocks A_i.

Remark: We emphasize that this theorem applies to block diagonal matrices, whereas the analogous Theorem 9.18 on characteristic polynomials applies to block triangular matrices.

EXAMPLE 9.14 Find the characteristic polynomal $\Delta(t)$ and the minimal polynomial $m(t)$ of the block diagonal matrix:

$$M = \left[\begin{array}{cc|cc|c} 2 & 5 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ \hline 0 & 0 & 4 & 2 & 0 \\ 0 & 0 & 3 & 5 & 0 \\ \hline 0 & 0 & 0 & 0 & 7 \end{array}\right] = \operatorname{diag}(A_1, A_2, A_3), \text{ where } A_1 = \begin{bmatrix} 2 & 5 \\ 0 & 2 \end{bmatrix}, A_2 = \begin{bmatrix} 4 & 2 \\ 3 & 5 \end{bmatrix}, A_3 = [7]$$

Then $\Delta(t)$ is the product of the characterization polynomials $\Delta_1(t)$, $\Delta_2(t)$, $\Delta_3(t)$ of A_1, A_2, A_3, respectively. One can show that

$$\Delta_1(t) = (t-2)^2, \qquad \Delta_2(t) = (t-2)(t-7), \qquad \Delta_3(t) = t - 7$$

Thus, $\Delta(t) = (t-2)^3(t-7)^2$. [As expected, $\deg \Delta(t) = 5$.]

The minimal polynomials $m_1(t)$, $m_2(t)$, $m_3(t)$ of the diagonal blocks A_1, A_2, A_3, respectively, are equal to the characteristic polynomials; that is,

$$m_1(t) = (t-2)^2, \qquad m_2(t) = (t-2)(t-7), \qquad m_3(t) = t - 7$$

But $m(t)$ is equal to the least common multiple of $m_1(t), m_2(t), m_3(t)$. Thus, $m(t) = (t-2)^2(t-7)$.

SOLVED PROBLEMS

Polynomials of Matrices, Characteristic Polynomials

9.1. Let $A = \begin{bmatrix} 1 & -2 \\ 4 & 5 \end{bmatrix}$. Find $f(A)$, where

(a) $f(t) = t^2 - 3t + 7$, (b) $f(t) = t^2 - 6t + 13$

First find $A^2 = \begin{bmatrix} 1 & -2 \\ 4 & 5 \end{bmatrix}\begin{bmatrix} 1 & -2 \\ 4 & 5 \end{bmatrix} = \begin{bmatrix} -7 & -12 \\ 24 & 17 \end{bmatrix}$. Then

(a) $f(A) = A^2 - 3A + 7I = \begin{bmatrix} -7 & -12 \\ 24 & 17 \end{bmatrix} + \begin{bmatrix} -3 & 6 \\ -12 & -15 \end{bmatrix} + \begin{bmatrix} 7 & 0 \\ 0 & 7 \end{bmatrix} = \begin{bmatrix} -3 & -6 \\ 12 & 9 \end{bmatrix}$

(b) $f(A) = A^2 - 6A + 13I = \begin{bmatrix} -7 & -12 \\ 24 & 17 \end{bmatrix} + \begin{bmatrix} -6 & 12 \\ -24 & -30 \end{bmatrix} + \begin{bmatrix} 13 & 0 \\ 0 & 13 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$

[Thus, A is a root of $f(t)$.]

9.2. Find the characteristic polynomial $\Delta(t)$ of each of the following matrices:

(a) $A = \begin{bmatrix} 2 & 5 \\ 4 & 1 \end{bmatrix}$, (b) $B = \begin{bmatrix} 7 & -3 \\ 5 & -2 \end{bmatrix}$, (c) $C = \begin{bmatrix} 3 & -2 \\ 9 & -3 \end{bmatrix}$

Use the formula $(t) = t^2 - \operatorname{tr}(M)\, t + |M|$ for a 2×2 matrix M:

(a) $\operatorname{tr}(A) = 2 + 1 = 3$, $|A| = 2 - 20 = -18$, so $\Delta(t) = t^2 - 3t - 18$
(b) $\operatorname{tr}(B) = 7 - 2 = 5$, $|B| = -14 + 15 = 1$, so $\Delta(t) = t^2 - 5t + 1$
(c) $\operatorname{tr}(C) = 3 - 3 = 0$, $|C| = -9 + 18 = 9$, so $\Delta(t) = t^2 + 9$

9.3. Find the characteristic polynomial $\Delta(t)$ of each of the following matrices:

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

Use the formula $\Delta(t) = t^3 - \mathrm{tr}(A)t^2 + (A_{11} + A_{22} + A_{33})t - |A|$, where A_{ii} is the cofactor of a_{ii} in the 3×3 matrix $A = [a_{ij}]$.

(a) $\mathrm{tr}(A) = 1 + 0 + 5 = 6$,

$$A_{11} = \begin{vmatrix} 0 & 4 \\ 4 & 5 \end{vmatrix} = -16, \qquad A_{22} = \begin{vmatrix} 1 & 3 \\ 6 & 5 \end{vmatrix} = -13, \qquad A_{33} = \begin{vmatrix} 1 & 2 \\ 3 & 0 \end{vmatrix} = -6$$

$$A_{11} + A_{22} + A_{33} = -35, \qquad \text{and} \qquad |A| = 48 + 36 - 16 - 30 = 38$$

Thus, $$\Delta(t) = t^3 - 6t^2 - 35t - 38$$

(b) $\mathrm{tr}(B) = 1 + 2 - 4 = -1$

$$B_{11} = \begin{vmatrix} 2 & 0 \\ 3 & -4 \end{vmatrix} = -8, \qquad B_{22} = \begin{vmatrix} 1 & -2 \\ 0 & -4 \end{vmatrix} = -4, \qquad B_{33} = \begin{vmatrix} 1 & 6 \\ -3 & 2 \end{vmatrix} = 20$$

$$B_{11} + B_{22} + B_{33} = 8, \qquad \text{and} \qquad |B| = -8 + 18 - 72 = -62$$

Thus, $$\Delta(t) = t^3 + t^2 - 8t + 62$$

9.4. Find the characteristic polynomial $\Delta(t)$ of each of the following matrices:

$$\text{(a)} \quad A = \begin{bmatrix} 2 & 5 & 1 & 1 \\ 1 & 4 & 2 & 2 \\ 0 & 0 & 6 & -5 \\ 0 & 0 & 2 & 3 \end{bmatrix}, \quad \text{(b)} \quad B = \begin{bmatrix} 1 & 1 & 2 & 2 \\ 0 & 3 & 3 & 4 \\ 0 & 0 & 5 & 5 \\ 0 & 0 & 0 & 6 \end{bmatrix}$$

(a) A is block triangular with diagonal blocks

$$A_1 = \begin{bmatrix} 2 & 5 \\ 1 & 4 \end{bmatrix} \qquad \text{and} \qquad A_2 = \begin{bmatrix} 6 & -5 \\ 2 & 3 \end{bmatrix}$$

Thus, $$\Delta(t) = \Delta_{A_1}(t)\Delta_{A_2}(t) = (t^2 - 6t + 3)(t^2 - 9t + 28)$$

(b) Because B is triangular, $\Delta(t) = (t-1)(t-3)(t-5)(t-6)$.

9.5. Find the characteristic polynomial $\Delta(t)$ of each of the following linear operators:

(a) $F: \mathbf{R}^2 \to \mathbf{R}^2$ defined by $F(x, y) = (3x + 5y, \quad 2x - 7y)$.

(b) $\mathbf{D}: V \to V$ defined by $\mathbf{D}(f) = df/dt$, where V is the space of functions with basis $S = \{\sin t, \cos t\}$.

The characteristic polynomial $\Delta(t)$ of a linear operator is equal to the characteristic polynomial of any matrix A that represents the linear operator.

(a) Find the matrix A that represents T relative to the usual basis of $\mathbf{R}^2$. We have

$$A = \begin{bmatrix} 3 & 5 \\ 2 & -7 \end{bmatrix}, \qquad \text{so} \qquad \Delta(t) = t^2 - \mathrm{tr}(A)\, t + |A| = t^2 + 4t - 31$$

(b) Find the matrix A representing the differential operator $\mathbf{D}$ relative to the basis S. We have

$$\begin{aligned} \mathbf{D}(\sin t) &= \cos t \;\;\;= 0(\sin t) + 1(\cos t) \\ \mathbf{D}(\cos t) &= -\sin t = -1(\sin t) + 0(\cos t) \end{aligned} \qquad \text{and so} \qquad A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

Therefore, $$\Delta(t) = t^2 - \mathrm{tr}(A)\, t + |A| = t^2 + 1$$

9.6. Show that a matrix A and its transpose A^T have the same characteristic polynomial.

By the transpose operation, $(tI - A)^T = tI^T - A^T = tI - A^T$. Because a matrix and its transpose have the same determinant,

$$\Delta_A(t) = |tI - A| = |(tI - A)^T| = |tI - A^T| = \Delta_{A^T}(t)$$

9.7. Prove Theorem 9.1: Let f and g be polynomials. For any square matrix A and scalar k,

(i) $(f+g)(A) = f(A) + g(A)$, (iii) $(kf)(A) = kf(A)$,
(ii) $(fg)(A) = f(A)g(A)$, (iv) $f(A)g(A) = g(A)f(A)$.

Suppose $f = a_n t^n + \cdots + a_1 t + a_0$ and $g = b_m t^m + \cdots + b_1 t + b_0$. Then, by definition,

$$f(A) = a_n A^n + \cdots + a_1 A + a_0 I \quad \text{and} \quad g(A) = b_m A^m + \cdots + b_1 A + b_0 I$$

(i) Suppose $m \le n$ and let $b_i = 0$ if $i > m$. Then

$$f + g = (a_n + b_n)t^n + \cdots + (a_1 + b_1)t + (a_0 + b_0)$$

Hence,

$$\begin{aligned}(f+g)(A) &= (a_n + b_n)A^n + \cdots + (a_1 + b_1)A + (a_0 + b_0)I \\ &= a_n A^n + b_n A^n + \cdots + a_1 A + b_1 A + a_0 I + b_0 I = f(A) + g(A)\end{aligned}$$

(ii) By definition, $fg = c_{n+m} t^{n+m} + \cdots + c_1 t + c_0 = \sum_{k=0}^{n+m} c_k t^k$, where

$$c_k = a_0 b_k + a_1 b_{k-1} + \cdots + a_k b_0 = \sum_{i=0}^{k} a_i b_{k-i}$$

Hence, $(fg)(A) = \sum_{k=0}^{n+m} c_k A^k$ and

$$f(A)g(A) = \left(\sum_{i=0}^{n} a_i A^i\right)\left(\sum_{j=0}^{m} b_j A^j\right) = \sum_{i=0}^{n}\sum_{j=0}^{m} a_i b_j A^{i+j} = \sum_{k=0}^{n+m} c_k A^k = (fg)(A)$$

(iii) By definition, $kf = ka_n t^n + \cdots + ka_1 t + ka_0$, and so

$$(kf)(A) = ka_n A^n + \cdots + ka_1 A + ka_0 I = k(a_n A^n + \cdots + a_1 A + a_0 I) = kf(A)$$

(iv) By (ii), $g(A)f(A) = (gf)(A) = (fg)(A) = f(A)g(A)$.

9.8. Prove the Cayley–Hamilton Theorem 9.2: Every matrix A is a root of its characterstic polynomial $\Delta(t)$.

Let A be an arbitrary n-square matrix and let $\Delta(t)$ be its characteristic polynomial, say,

$$\Delta(t) = |tI - A| = t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0$$

Now let $B(t)$ denote the classical adjoint of the matrix $tI - A$. The elements of $B(t)$ are cofactors of the matrix $tI - A$ and hence are polynomials in t of degree not exceeding $n-1$. Thus,

$$B(t) = B_{n-1}t^{n-1} + \cdots + B_1 t + B_0$$

where the B_i are n-square matrices over K which are independent of t. By the fundamental property of the classical adjoint (Theorem 8.9), $(tI - A)B(t) = |tI - A|I$, or

$$(tI - A)(B_{n-1}t^{n-1} + \cdots + B_1 t + B_0) = (t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0)I$$

Removing the parentheses and equating corresponding powers of t yields

$$B_{n-1} = I, \qquad B_{n-2} - AB_{n-1} = a_{n-1}I, \qquad \ldots, \qquad B_0 - AB_1 = a_1 I, \qquad -AB_0 = a_0 I$$

Multiplying the above equations by $A^n,\ A^{n-1},\ \ldots,\ A,\ I$, respectively, yields

$$A^n B_{n-1} = A_n I, \qquad A^{n-1}B_{n-2} - A^n B_{n-1} = a_{n-1}A^{n-1}, \qquad \ldots, \qquad AB_0 - A^2 B_1 = a_1 A, \qquad -AB_0 = a_0 I$$

Adding the above matrix equations yields 0 on the left-hand side and $\Delta(A)$ on the right-hand side; that is,

$$0 = A^n + a_{n-1}A^{n-1} + \cdots + a_1 A + a_0 I$$

Therefore, $\Delta(A) = 0$, which is the Cayley–Hamilton theorem.

Eigenvalues and Eigenvectors of 2×2 Matrices

9.9. Let $A = \begin{bmatrix} 3 & -4 \\ 2 & -6 \end{bmatrix}$.

(a) Find all eigenvalues and corresponding eigenvectors.

(b) Find matrices P and D such that P is nonsingular and $D = P^{-1}AP$ is diagonal.

(a) First find the characteristic polynomial $\Delta(t)$ of A:

$$\Delta(t) = t^2 - \mathrm{tr}(A)\, t + |A| = t^2 + 3t - 10 = (t-2)(t+5)$$

The roots $\lambda = 2$ and $\lambda = -5$ of $\Delta(t)$ are the eigenvalues of A. We find corresponding eigenvectors.

(i) Subtract $\lambda = 2$ down the diagonal of A to obtain the matrix $M = A - 2I$, where the corresponding homogeneous system $MX = 0$ yields the eigenvectors corresponding to $\lambda = 2$. We have

$$M = \begin{bmatrix} 1 & -4 \\ 2 & -8 \end{bmatrix}, \qquad \text{corresponding to} \qquad \begin{aligned} x - 4y &= 0 \\ 2x - 8y &= 0 \end{aligned} \qquad \text{or} \qquad x - 4y = 0$$

The system has only one free variable, and $v_1 = (4, 1)$ is a nonzero solution. Thus, $v_1 = (4, 1)$ is an eigenvector belonging to (and spanning the eigenspace of) $\lambda = 2$.

(ii) Subtract $\lambda = -5$ (or, equivalently, add 5) down the diagonal of A to obtain

$$M = \begin{bmatrix} 8 & -4 \\ 2 & -1 \end{bmatrix}, \qquad \text{corresponding to} \qquad \begin{aligned} 8x - 4y &= 0 \\ 2x - y &= 0 \end{aligned} \qquad \text{or} \qquad 2x - y = 0$$

The system has only one free variable, and $v_2 = (1, 2)$ is a nonzero solution. Thus, $v_2 = (1, 2)$ is an eigenvector belonging to $\lambda = 5$.

(b) Let P be the matrix whose columns are v_1 and v_2. Then

$$P = \begin{bmatrix} 4 & 1 \\ 1 & 2 \end{bmatrix} \qquad \text{and} \qquad D = P^{-1}AP = \begin{bmatrix} 2 & 0 \\ 0 & -5 \end{bmatrix}$$

Note that D is the diagonal matrix whose diagonal entries are the eigenvalues of A corresponding to the eigenvectors appearing in P.

Remark: Here P is the change-of-basis matrix from the usual basis of $\mathbf{R}^2$ to the basis $S = \{v_1, v_2\}$, and D is the matrix that represents (the matrix function) A relative to the new basis S.

9.10. Let $A = \begin{bmatrix} 2 & 2 \\ 1 & 3 \end{bmatrix}$.

(a) Find all eigenvalues and corresponding eigenvectors.

(b) Find a nonsingular matrix P such that $D = P^{-1}AP$ is diagonal, and P^{-1}.

(c) Find A^6 and $f(A)$, where $t^4 - 3t^3 - 6t^2 + 7t + 3$.

(d) Find a ‘‘real cube root’’ of B—that is, a matrix B such that $B^3 = A$ and B has real eigenvalues.

(a) First find the characteristic polynomial $\Delta(t)$ of A:

$$\Delta(t) = t^2 - \mathrm{tr}(A)\, t + |A| = t^2 - 5t + 4 = (t-1)(t-4)$$

The roots $\lambda = 1$ and $\lambda = 4$ of $\Delta(t)$ are the eigenvalues of A. We find corresponding eigenvectors.

(i) Subtract $\lambda = 1$ down the diagonal of A to obtain the matrix $M = A - \lambda I$, where the corresponding homogeneous system $MX = 0$ yields the eigenvectors belonging to $\lambda = 1$. We have

$$M = \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}, \qquad \text{corresponding to} \qquad \begin{aligned} x + 2y &= 0 \\ x + 2y &= 0 \end{aligned} \qquad \text{or} \qquad x + 2y = 0$$

The system has only one independent solution; for example, $x = 2, y = -1$. Thus, $v_1 = (2, -1)$ is an eigenvector belonging to (and spanning the eigenspace of) $\lambda = 1$.

(ii) Subtract $\lambda = 4$ down the diagonal of A to obtain

$$M = \begin{bmatrix} -2 & 2 \\ 1 & -1 \end{bmatrix}, \quad \text{corresponding to} \quad \begin{aligned} -2x + 2y &= 0 \\ x - \quad y &= 0 \end{aligned} \quad \text{or} \quad x - y = 0$$

The system has only one independent solution; for example, $x = 1, y = 1$. Thus, $v_2 = (1, 1)$ is an eigenvector belonging to $\lambda = 4$.

(b) Let P be the matrix whose columns are v_1 and v_2. Then

$$P = \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix} \quad \text{and} \quad D = P^{-1}AP = \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}, \quad \text{where} \quad P^{-1} = \begin{bmatrix} \frac{1}{3} & -\frac{1}{3} \\ \frac{1}{3} & \frac{2}{3} \end{bmatrix}$$

(c) Using the diagonal factorization $A = PDP^{-1}$, and $1^6 = 1$ and $4^6 = 4096$, we get

$$A^6 = PD^6P^{-1} = \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 4096 \end{bmatrix} \begin{bmatrix} \frac{1}{3} & -\frac{1}{3} \\ \frac{1}{3} & \frac{2}{3} \end{bmatrix} = \begin{bmatrix} 1366 & 2230 \\ 1365 & 2731 \end{bmatrix}$$

Also, $f(1) = 2$ and $f(4) = -1$. Hence,

$$f(A) = Pf(D)P^{-1} = \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} \frac{1}{3} & -\frac{1}{3} \\ \frac{1}{3} & \frac{2}{3} \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ -1 & 0 \end{bmatrix}$$

(d) Here $\begin{bmatrix} 1 & 0 \\ 0 & \sqrt[3]{4} \end{bmatrix}$ is the real cube root of D. Hence the real cube root of A is

$$B = P\sqrt[3]{D}P^{-1} = \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \sqrt[3]{4} \end{bmatrix} \begin{bmatrix} \frac{1}{3} & -\frac{1}{3} \\ \frac{1}{3} & \frac{2}{3} \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 2 + \sqrt[3]{4} & -2 + 2\sqrt[3]{4} \\ -1 + \sqrt[3]{4} & 1 + 2\sqrt[3]{4} \end{bmatrix}$$

9.11. Each of the following real matrices defines a linear transformation on $\mathbf{R}^2$:

(a) $A = \begin{bmatrix} 5 & 6 \\ 3 & -2 \end{bmatrix}$, (b) $B = \begin{bmatrix} 1 & -1 \\ 2 & -1 \end{bmatrix}$, (c) $C = \begin{bmatrix} 5 & -1 \\ 1 & 3 \end{bmatrix}$

Find, for each matrix, all eigenvalues and a maximum set S of linearly independent eigenvectors. Which of these linear operators are diagonalizable—that is, which can be represented by a diagonal matrix?

(a) First find $\Delta(t) = t^2 - 3t - 28 = (t - 7)(t + 4)$. The roots $\lambda = 7$ and $\lambda = -4$ are the eigenvalues of A. We find corresponding eigenvectors.

(i) Subtract $\lambda = 7$ down the diagonal of A to obtain

$$M = \begin{bmatrix} -2 & 6 \\ 3 & -9 \end{bmatrix}, \quad \text{corresponding to} \quad \begin{aligned} -2x + 6y &= 0 \\ 3x - 9y &= 0 \end{aligned} \quad \text{or} \quad x - 3y = 0$$

Here $v_1 = (3, 1)$ is a nonzero solution.

(ii) Subtract $\lambda = -4$ (or add 4) down the diagonal of A to obtain

$$M = \begin{bmatrix} 9 & 6 \\ 3 & 2 \end{bmatrix}, \quad \text{corresponding to} \quad \begin{aligned} 9x + 6y &= 0 \\ 3x + 2y &= 0 \end{aligned} \quad \text{or} \quad 3x + 2y = 0$$

Here $v_2 = (2, -3)$ is a nonzero solution.

Then $S = \{v_1, v_2\} = \{(3, 1), \ (2, -3)\}$ is a maximal set of linearly independent eigenvectors. Because S is a basis of $\mathbf{R}^2$, A is diagonalizable. Using the basis S, A is represented by the diagonal matrix $D = \mathrm{diag}(7, -4)$.

(b) First find the characteristic polynomial $\Delta(t) = t^2 + 1$. There are no real roots. Thus B, a real matrix representing a linear transformation on $\mathbf{R}^2$, has no eigenvalues and no eigenvectors. Hence, in particular, B is not diagonalizable.

(c) First find $\Delta(t) = t^2 - 8t + 16 = (t-4)^2$. Thus, $\lambda = 4$ is the only eigenvalue of C. Subtract $\lambda = 4$ down the diagonal of C to obtain

$$M = \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}, \qquad \text{corresponding to} \qquad x - y = 0$$

The homogeneous system has only one independent solution; for example, $x = 1,\ y = 1$. Thus, $v = (1, 1)$ is an eigenvector of C. Furthermore, as there are no other eigenvalues, the singleton set $S = \{v\} = \{(1, 1)\}$ is a maximal set of linearly independent eigenvectors of C. Furthermore, because S is not a basis of $\mathbf{R}^2$, C is not diagonalizable.

9.12. Suppose the matrix B in Problem 9.11 represents a linear operator on complex space $\mathbf{C}^2$. Show that, in this case, B is diagonalizable by finding a basis S of $\mathbf{C}^2$ consisting of eigenvectors of B.

The characteristic polynomial of B is still $\Delta(t) = t^2 + 1$. As a polynomial over $\mathbf{C}$, $\Delta(t)$ *does* factor; specifically, $\Delta(t) = (t - i)(t + i)$. Thus, $\lambda = i$ and $\lambda = -i$ are the eigenvalues of B.

(i) Subtract $\lambda = i$ down the diagonal of B to obtain the homogeneous system

$$\begin{aligned} (1-i)x - \qquad\quad y &= 0 \\ 2x + (-1-i)y &= 0 \end{aligned} \qquad \text{or} \qquad (1-i)x - y = 0$$

The system has only one independent solution; for example, $x = 1, y = 1 - i$. Thus, $v_1 = (1,\ 1 - i)$ is an eigenvector that spans the eigenspace of $\lambda = i$.

(ii) Subtract $\lambda = -i$ (or add i) down the diagonal of B to obtain the homogeneous system

$$\begin{aligned} (1+i)x - \qquad\quad y &= 0 \\ 2x + (-1+i)y &= 0 \end{aligned} \qquad \text{or} \qquad (1+i)x - y = 0$$

The system has only one independent solution; for example, $x = 1, y = 1 + i$. Thus, $v_2 = (1,\ 1 + i)$ is an eigenvector that spans the eigenspace of $\lambda = -i$.

As a complex matrix, B is diagonalizable. Specifically, $S = \{v_1, v_2\} = \{(1, 1 - i),\ (1, 1 + i)\}$ is a basis of $\mathbf{C}^2$ consisting of eigenvectors of B. Using this basis S, B is represented by the diagonal matrix $D = \mathrm{diag}(i, -i)$.

9.13. Let L be the linear transformation on $\mathbf{R}^2$ that reflects each point P across the line $y = kx$, where $k > 0$. (See Fig. 9-1.)

(a) Show that $v_1 = (k, 1)$ and $v_2 = (1, -k)$ are eigenvectors of L.

(b) Show that L is diagonalizable, and find a diagonal representation D.

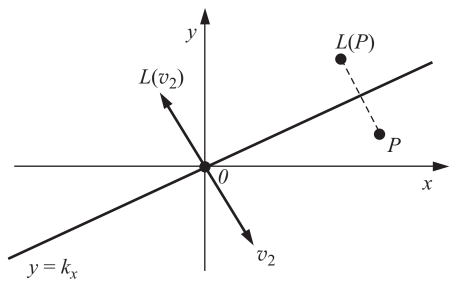

Figure 9-1

(a) The vector $v_1 = (k, 1)$ lies on the line $y = kx$, and hence is left fixed by L; that is, $L(v_1) = v_1$. Thus, v_1 is an eigenvector of L belonging to the eigenvalue $\lambda_1 = 1$.

The vector $v_2 = (1, -k)$ is perpendicular to the line $y = kx$, and hence, L reflects v_2 into its negative; that is, $L(v_2) = -v_2$. Thus, v_2 is an eigenvector of L belonging to the eigenvalue $\lambda_2 = -1$.

(b) Here $S = \{v_1, v_2\}$ is a basis of $\mathbf{R}^2$ consisting of eigenvectors of L. Thus, L is diagonalizable, with the diagonal representation $D = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ (relative to the basis S).

Eigenvalues and Eigenvectors

9.14. Let $A = \begin{bmatrix} 4 & 1 & -1 \\ 2 & 5 & -2 \\ 1 & 1 & 2 \end{bmatrix}$. (a) Find all eigenvalues of A.

(b) Find a maximum set S of linearly independent eigenvectors of A.

(c) Is A diagonalizable? If yes, find P such that $D = P^{-1}AP$ is diagonal.

(a) First find the characteristic polynomial $\Delta(t)$ of A. We have

$$\mathrm{tr}(A) = 4 + 5 + 2 = 11 \qquad \text{and} \qquad |A| = 40 - 2 - 2 + 5 + 8 - 4 = 45$$

Also, find each cofactor A_{ii} of a_{ii} in A:

$$A_{11} = \begin{vmatrix} 5 & -2 \\ 1 & 2 \end{vmatrix} = 12, \qquad A_{22} = \begin{vmatrix} 4 & -1 \\ 1 & 2 \end{vmatrix} = 9, \qquad A_{33} = \begin{vmatrix} 4 & 1 \\ 2 & 5 \end{vmatrix} = 18$$

Hence, $\quad \Delta(t) = t^3 - \mathrm{tr}(A)\,t^2 + (A_{11} + A_{22} + A_{33})t - |A| = t^3 - 11t^2 + 39t - 45$

Assuming Δt has a rational root, it must be among ±1, ±3, ±5, ±9, ±15, ±45. Testing, by synthetic division, we get

$$\begin{array}{r|l} 3 & 1 - 11 + 39 - 45 \\ & \qquad\ \ 3 - 24 + 45 \\ \hline & 1 - 8 + 15 + \ \ 0 \end{array}$$

Thus, $t = 3$ is a root of $\Delta(t)$. Also, $t - 3$ is a factor and $t^2 - 8t + 15$ is a factor. Hence,

$$\Delta(t) = (t-3)(t^2 - 8t + 15) = (t-3)(t-5)(t-3) = (t-3)^2(t-5)$$

Accordingly, $\lambda = 3$ and $\lambda = 5$ are eigenvalues of A.

(b) Find linearly independent eigenvectors for each eigenvalue of A.

(i) Subtract $\lambda = 3$ down the diagonal of A to obtain the matrix

$$M = \begin{bmatrix} 1 & 1 & -1 \\ 2 & 2 & -2 \\ 1 & 1 & -1 \end{bmatrix}, \qquad \text{corresponding to} \qquad x + y - z = 0$$

Here $u = (1, -1, 0)$ and $v = (1, 0, 1)$ are linearly independent solutions.

(ii) Subtract $\lambda = 5$ down the diagonal of A to obtain the matrix

$$M = \begin{bmatrix} -1 & 1 & -1 \\ 2 & 0 & -2 \\ 1 & 1 & -3 \end{bmatrix}, \qquad \text{corresponding to} \qquad \begin{aligned} -x + y - z &= 0 \\ 2x \phantom{{}+y} - 2z &= 0 \\ x + y - 3z &= 0 \end{aligned} \qquad \text{or} \qquad \begin{aligned} x \phantom{{}-2} - z &= 0 \\ y - 2z &= 0 \end{aligned}$$

Only z is a free variable. Here $w = (1, 2, 1)$ is a solution.

Thus, $S = \{u, v, w\} = \{(1, -1, 0),\ (1, 0, 1),\ (1, 2, 1)\}$ is a maximal set of linearly independent eigenvectors of A.

Remark: The vectors u and v were chosen so that they were independent solutions of the system $x + y - z = 0$. On the other hand, w is automatically independent of u and v because w belongs to a different eigenvalue of A. Thus, the three vectors are linearly independent.

(c) A is diagonalizable, because it has three linearly independent eigenvectors. Let P be the matrix with columns u, v, w. Then

$$P = \begin{bmatrix} 1 & 1 & 1 \\ -1 & 0 & 2 \\ 0 & 1 & 1 \end{bmatrix} \qquad \text{and} \qquad D = P^{-1}AP = \begin{bmatrix} 3 & & \\ & 3 & \\ & & 5 \end{bmatrix}$$

9.15. Repeat Problem 9.14 for the matrix $B = \begin{bmatrix} 3 & -1 & 1 \\ 7 & -5 & 1 \\ 6 & -6 & 2 \end{bmatrix}$.

(a) First find the characteristic polynomial $\Delta(t)$ of B. We have

$$\operatorname{tr}(B) = 0, \qquad |B| = -16, \qquad B_{11} = -4, \qquad B_{22} = 0, \qquad B_{33} = -8, \qquad \text{so} \qquad \sum_i B_{ii} = -12$$

Therefore, $\Delta(t) = t^3 - 12t + 16 = (t-2)^2(t+4)$. Thus, $\lambda_1 = 2$ and $\lambda_2 = -4$ are the eigenvalues of B.

(b) Find a basis for the eigenspace of each eigenvalue of B.

(i) Subtract $\lambda_1 = 2$ down the diagonal of B to obtain

$$M = \begin{bmatrix} 1 & -1 & 1 \\ 7 & -7 & 1 \\ 6 & -6 & 0 \end{bmatrix}, \qquad \text{corresponding to} \qquad \begin{aligned} x - y + z &= 0 \\ 7x - 7y + z &= 0 \\ 6x - 6y \phantom{{}+z} &= 0 \end{aligned} \qquad \text{or} \qquad \begin{aligned} x - y + z &= 0 \\ z &= 0 \end{aligned}$$

The system has only one independent solution; for example, $x = 1$, $y = 1$, $z = 0$. Thus, $u = (1, 1, 0)$ forms a basis for the eigenspace of $\lambda_1 = 2$.

(ii) Subtract $\lambda_2 = -4$ (or add 4) down the diagonal of B to obtain

$$M = \begin{bmatrix} 7 & -1 & 1 \\ 7 & -1 & 1 \\ 6 & -6 & 6 \end{bmatrix}, \qquad \text{corresponding to} \qquad \begin{aligned} 7x - y + z &= 0 \\ 7x - y + z &= 0 \\ 6x - 6y + 6z &= 0 \end{aligned} \qquad \text{or} \qquad \begin{aligned} x - y + z &= 0 \\ 6y - 6z &= 0 \end{aligned}$$

The system has only one independent solution; for example, $x = 0$, $y = 1$, $z = 1$. Thus, $v = (0, 1, 1)$ forms a basis for the eigenspace of $\lambda_2 = -4$.

Thus $S = \{u, v\}$ is a maximal set of linearly independent eigenvectors of B.

(c) Because B has at most two linearly independent eigenvectors, B is not similar to a diagonal matrix; that is, B is not diagonalizable.

9.16. Find the algebraic and geometric multiplicities of the eigenvalue $\lambda_1 = 2$ of the matrix B in Problem 9.15.

The algebraic multiplicity of $\lambda_1 = 2$ is 2, because $t - 2$ appears with exponent 2 in $\Delta(t)$. However, the geometric multiplicity of $\lambda_1 = 2$ is 1, because $\dim E_{\lambda_1} = 1$ (where E_{λ_1} is the eigenspace of λ_1).

9.17. Let $T\colon \mathbf{R}^3 \to \mathbf{R}^3$ be defined by $T(x, y, z) = (2x + y - 2z, \quad 2x + 3y - 4z, \quad x + y - z)$. Find all eigenvalues of T, and find a basis of each eigenspace. Is T diagonalizable? If so, find the basis S of $\mathbf{R}^3$ that diagonalizes T, and find its diagonal representation D.

First find the matrix A that represents T relative to the usual basis of $\mathbf{R}^3$ by writing down the coefficients of x, y, z as rows, and then find the characteristic polynomial of A (and T). We have

$$A = [T] = \begin{bmatrix} 2 & 1 & -2 \\ 2 & 3 & -4 \\ 1 & 1 & -1 \end{bmatrix} \qquad \text{and} \qquad \begin{gathered} \operatorname{tr}(A) = 4, \quad |A| = 2 \\ A_{11} = 1, \quad A_{22} = 0, \quad A_{33} = 4 \\ \sum_i A_{ii} = 5 \end{gathered}$$

Therefore, $\Delta(t) = t^3 - 4t^2 + 5t - 2 = (t-1)^2(t-2)$, and so $\lambda = 1$ and $\lambda = 2$ are the eigenvalues of A (and T). We next find linearly independent eigenvectors for each eigenvalue of A.

(i) Subtract $\lambda = 1$ down the diagonal of A to obtain the matrix

$$M = \begin{bmatrix} 1 & 1 & -2 \\ 2 & 2 & -4 \\ 1 & 1 & -2 \end{bmatrix}, \qquad \text{corresponding to} \qquad x + y - 2z = 0$$

Here y and z are free variables, and so there are two linearly independent eigenvectors belonging to $\lambda = 1$. For example, $u = (1, -1, 0)$ and $v = (2, 0, 1)$ are two such eigenvectors.

(ii) Subtract $\lambda = 2$ down the diagonal of A to obtain

$$M = \begin{bmatrix} 0 & 1 & -2 \\ 2 & 1 & -4 \\ 1 & 1 & -3 \end{bmatrix}, \qquad \text{corresponding to} \qquad \begin{aligned} y - 2z &= 0 \\ 2x + y - 4z &= 0 \\ x + y - 3z &= 0 \end{aligned} \qquad \text{or} \qquad \begin{aligned} x + y - 3z &= 0 \\ y - 2z &= 0 \end{aligned}$$

Only z is a free variable. Here $w = (1, 2, 1)$ is a solution.

Thus, T is diagonalizable, because it has three independent eigenvectors. Specifically, choosing

$$S = \{u, v, w\} = \{(1, -1, 0), \quad (2, 0, 1), \quad (1, 2, 1)\}$$

as a basis, T is represented by the diagonal matrix $D = \text{diag}(1, 1, 2)$.

9.18. Prove the following for a linear operator (matrix) T:

(a) The scalar 0 is an eigenvalue of T if and only if T is singular.

(b) If λ is an eigenvalue of T, where T is invertible, then λ^{-1} is an eigenvalue of T^{-1}.

(a) We have that 0 is an eigenvalue of T if and only if there is a vector $v \neq 0$ such that $T(v) = 0v$—that is, if and only if T is singular.

(b) Because T is invertible, it is nonsingular; hence, by (a), $\lambda \neq 0$. By definition of an eigenvalue, there exists $v \neq 0$ such that $T(v) = \lambda v$. Applying T^{-1} to both sides, we obtain

$$v = T^{-1}(\lambda v) = \lambda T^{-1}(v), \qquad \text{and so} \qquad T^{-1}(v) = \lambda^{-1} v$$

Therefore, λ^{-1} is an eigenvalue of T^{-1}.

9.19. Let λ be an eigenvalue of a linear operator $T: V \to V$, and let E_λ consists of all the eigenvectors belonging to λ (called the *eigenspace* of λ). Prove that E_λ is a subspace of V. That is, prove

(a) If $u \in E_\lambda$, then $ku \in E_\lambda$ for any scalar k. (b) If $u, v, \in E_\lambda$, then $u + v \in E_\lambda$.

(a) Because $u \in E_\lambda$, we have $T(u) = \lambda u$. Then $T(ku) = kT(u) = k(\lambda u) = \lambda(ku)$, and so $ku \in E_\lambda$.
(We view the zero vector $0 \in V$ as an ''eigenvector'' of λ in order for E_λ to be a subspace of V.)

(b) As $u, v \in E_\lambda$, we have $T(u) = \lambda u$ and $T(v) = \lambda v$. Then
$T(u + v) = T(u) + T(v) = \lambda u + \lambda v = \lambda(u + v)$, and so $u + v \in E_\lambda$

9.20. Prove Theorem 9.6: The following are equivalent: (i) The scalar λ is an eigenvalue of A.

(ii) The matrix $\lambda I - A$ is singular.

(iii) The scalar λ is a root of the characteristic polynomial $\Delta(t)$ of A.

The scalar λ is an eigenvalue of A if and only if there exists a nonzero vector v such that

$$Av = \lambda v \qquad \text{or} \qquad (\lambda I)v - Av = 0 \qquad \text{or} \qquad (\lambda I - A)v = 0$$

or $\lambda I - A$ is singular. In such a case, λ is a root of $\Delta(t) = |tI - A|$. Also, v is in the eigenspace E_λ of λ if and only if the above relations hold. Hence, v is a solution of $(\lambda I - A)X = 0$.

9.21. Prove Theorem 9.8′: Suppose $v_1, v_2, \ldots, v_n$ are nonzero eigenvectors of T belonging to distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$. Then $v_1, v_2, \ldots, v_n$ are linearly independent.

Suppose the theorem is not true. Let $v_1, v_2, \ldots, v_s$ be a minimal set of vectors for which the theorem is not true. We have $s > 1$, because $v_1 \neq 0$. Also, by the minimality condition, $v_2, \ldots, v_s$ are linearly independent. Thus, v_1 is a linear combination of $v_2, \ldots, v_s$, say,

$$v_1 = a_2 v_2 + a_3 v_3 + \cdots + a_s v_s \tag{1}$$

(where some $a_k \neq 0$). Applying T to (1) and using the linearity of T yields

$$T(v_1) = T(a_2 v_2 + a_3 v_3 + \cdots + a_s v_s) = a_2 T(v_2) + a_3 T(v_3) + \cdots + a_s T(v_s) \tag{2}$$

Because v_j is an eigenvector of T belonging to λ_j, we have $T(v_j) = \lambda_j v_j$. Substituting in (2) yields

$$\lambda_1 v_1 = a_2 \lambda_2 v_2 + a_3 \lambda_3 v_3 + \cdots + a_s \lambda_s v_s \tag{3}$$

Multiplying (1) by λ_1 yields

$$\lambda_1 v_1 = a_2 \lambda_1 v_2 + a_3 \lambda_1 v_3 + \cdots + a_s \lambda_1 v_s \tag{4}$$

Setting the right-hand sides of (3) and (4) equal to each other, or subtracting (3) from (4) yields

$$a_2(\lambda_1 - \lambda_2) v_2 + a_3(\lambda_1 - \lambda_3) v_3 + \cdots + a_s(\lambda_1 - \lambda_s) v_s = 0 \tag{5}$$

Because $v_2, v_3, \ldots, v_s$ are linearly independent, the coefficients in (5) must all be zero. That is,

$$a_2(\lambda_1 - \lambda_2) = 0, \qquad a_3(\lambda_1 - \lambda_3) = 0, \qquad \ldots, \qquad a_s(\lambda_1 - \lambda_s) = 0$$

However, the λ_i are distinct. Hence $\lambda_1 - \lambda_j \neq 0$ for $j > 1$. Hence, $a_2 = 0$, $a_3 = 0, \ldots, a_s = 0$. This contradicts the fact that some $a_k \neq 0$. The theorem is proved.

9.22. Prove Theorem 9.9. Suppose $\Delta(t) = (t - a_1)(t - a_2) \ldots (t - a_n)$ is the characteristic polynomial of an n-square matrix A, and suppose the n roots a_i are distinct. Then A is similar to the diagonal matrix $D = \mathrm{diag}(a_1, a_2, \ldots, a_n)$.

Let $v_1, v_2, \ldots, v_n$ be (nonzero) eigenvectors corresponding to the eigenvalues a_i. Then the n eigenvectors v_i are linearly independent (Theorem 9.8), and hence form a basis of K^n. Accordingly, A is diagonalizable (i.e., A is similar to a diagonal matrix D), and the diagonal elements of D are the eigenvalues a_i.

9.23. Prove Theorem 9.10′: The geometric multiplicity of an eigenvalue λ of T does not exceed its algebraic multiplicity.

Suppose the geometric multiplicity of λ is r. Then its eigenspace E_λ contains r linearly independent eigenvectors $v_1, \ldots, v_r$. Extend the set $\{v_i\}$ to a basis of V, say, $\{v_i, \ldots, v_r, w_1, \ldots, w_s\}$. We have

$$T(v_1) = \lambda v_1, \qquad T(v_2) = \lambda v_2, \qquad \ldots, \qquad T(v_r) = \lambda v_r,$$

$$\begin{aligned} T(w_1) &= a_{11} v_1 + \cdots + a_{1r} v_r + b_{11} w_1 + \cdots + b_{1s} w_s \\ T(w_2) &= a_{21} v_1 + \cdots + a_{2r} v_r + b_{21} w_1 + \cdots + b_{2s} w_s \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ T(w_s) &= a_{s1} v_1 + \cdots + a_{sr} v_r + b_{s1} w_1 + \cdots + b_{ss} w_s \end{aligned}$$

Then $M = \begin{bmatrix} \lambda I_r & A \\ 0 & B \end{bmatrix}$ is the matrix of T in the above basis, where $A = [a_{ij}]^T$ and $B = [b_{ij}]^T$.

Because M is block diagonal, the characteristic polynomial $(t - \lambda)^r$ of the block λI_r must divide the characteristic polynomial of M and hence of T. Thus, the algebraic multiplicity of λ for T is at least r, as required.

Diagonalizing Real Symmetric Matrices and Quadratic Forms

9.24. Let $A = \begin{bmatrix} 7 & 3 \\ 3 & -1 \end{bmatrix}$. Find an orthogonal matrix P such that $D = P^{-1}AP$ is diagonal.

First find the characteristic polynomial $\Delta(t)$ of A. We have

$$\Delta(t) = t^2 - \mathrm{tr}(A)\, t + |A| = t^2 - 6t - 16 = (t-8)(t+2)$$

Thus, the eigenvalues of A are $\lambda = 8$ and $\lambda = -2$. We next find corresponding eigenvectors.

Subtract $\lambda = 8$ down the diagonal of A to obtain the matrix

$$M = \begin{bmatrix} -1 & 3 \\ 3 & -9 \end{bmatrix}, \qquad \text{corresponding to} \qquad \begin{aligned} -x + 3y &= 0 \\ 3x - 9y &= 0 \end{aligned} \qquad \text{or} \qquad x - 3y = 0$$

A nonzero solution is $u_1 = (3, 1)$.

Subtract $\lambda = -2$ (or add 2) down the diagonal of A to obtain the matrix

$$M = \begin{bmatrix} 9 & 3 \\ 3 & 1 \end{bmatrix}, \qquad \text{corresponding to} \qquad \begin{aligned} 9x + 3y &= 0 \\ 3x + y &= 0 \end{aligned} \qquad \text{or} \qquad 3x + y = 0$$

A nonzero solution is $u_2 = (1, -3)$.

As expected, because A is symmetric, the eigenvectors u_1 and u_2 are orthogonal. Normalize u_1 and u_2 to obtain, respectively, the unit vectors

$$\hat{u}_1 = (3/\sqrt{10}, 1/\sqrt{10}) \qquad \text{and} \qquad \hat{u}_2 = (1/\sqrt{10}, -3/\sqrt{10}).$$

Finally, let P be the matrix whose columns are the unit vectors $\hat{u}_1$ and $\hat{u}_2$, respectively. Then

$$P = \begin{bmatrix} 3/\sqrt{10} & 1/\sqrt{10} \\ 1/\sqrt{10} & -3/\sqrt{10} \end{bmatrix} \qquad \text{and} \qquad D = P^{-1}AP = \begin{bmatrix} 8 & 0 \\ 0 & -2 \end{bmatrix}$$

As expected, the diagonal entries in D are the eigenvalues of A.

9.25. Let $B = \begin{bmatrix} 11 & -8 & 4 \\ -8 & -1 & -2 \\ 4 & -2 & -4 \end{bmatrix}$. (a) Find all eigenvalues of B.

(b) Find a maximal set S of nonzero orthogonal eigenvectors of B.

(c) Find an orthogonal matrix P such that $D = P^{-1}BP$ is diagonal.

(a) First find the characteristic polynomial of B. We have

$$\mathrm{tr}(B) = 6, \qquad |B| = 400, \qquad B_{11} = 0, \qquad B_{22} = -60, \qquad B_{33} = -75, \qquad \text{so} \qquad \sum_i B_{ii} = -135$$

Hence, $\Delta(t) = t^3 - 6t^2 - 135t - 400$. If $\Delta(t)$ has an integer root it must divide 400. Testing $t = -5$, by synthetic division, yields

$$\begin{array}{r|l} -5 & 1 - 6 - 135 - 400 \\ & - 5 + 55 + 400 \\ \hline & 1 - 11 - 80 + 0 \end{array}$$

Thus, $t + 5$ is a factor of $\Delta(t)$, and $t^2 - 11t - 80$ is a factor. Thus,

$$\Delta(t) = (t+5)(t^2 - 11t - 80) = (t+5)^2(t-16)$$

The eigenvalues of B are $\lambda = -5$ (multiplicity 2), and $\lambda = 16$ (multiplicity 1).

(b) Find an orthogonal basis for each eigenspace. Subtract $\lambda = -5$ (or, add 5) down the diagonal of B to obtain the homogeneous system

$$16x - 8y + 4z = 0, \qquad -8x + 4y - 2z = 0, \qquad 4x - 2y + z = 0$$

That is, $4x - 2y + z = 0$. The system has two independent solutions. One solution is $v_1 = (0, 1, 2)$. We seek a second solution $v_2 = (a, b, c)$, which is orthogonal to v_1, such that

$$4a - 2b + c = 0, \qquad \text{and also} \qquad b - 2c = 0$$

One such solution is $v_2 = (-5, -8, 4)$.

Subtract $\lambda = 16$ down the diagonal of B to obtain the homogeneous system

$$-5x - 8y + 4z = 0, \qquad -8x - 17y - 2z = 0, \qquad 4x - 2y - 20z = 0$$

This system yields a nonzero solution $v_3 = (4, -2, 1)$. (As expected from Theorem 9.13, the eigenvector v_3 is orthogonal to v_1 and v_2.)

Then v_1, v_2, v_3 form a maximal set of nonzero orthogonal eigenvectors of B.

(c) Normalize v_1, v_2, v_3 to obtain the orthonormal basis:

$$\hat{v}_1 = v_1/\sqrt{5}, \qquad \hat{v}_2 = v_2/\sqrt{105}, \qquad \hat{v}_3 = v_3/\sqrt{21}$$

Then P is the matrix whose columns are $\hat{v}_1, \hat{v}_2, \hat{v}_3$. Thus,

$$P = \begin{bmatrix} 0 & -5/\sqrt{105} & 4/\sqrt{21} \\ 1/\sqrt{5} & -8/\sqrt{105} & -2/\sqrt{21} \\ 2/\sqrt{5} & 4/\sqrt{105} & 1/\sqrt{21} \end{bmatrix} \quad \text{and} \quad D = P^{-1}BP = \begin{bmatrix} -5 & & \\ & -5 & \\ & & 16 \end{bmatrix}$$

9.26. Let $q(x, y) = x^2 + 6xy - 7y^2$. Find an orthogonal substitution that diagonalizes q.

Find the symmetric matrix A that represents q and its characteristic polynomial $\Delta(t)$. We have

$$A = \begin{bmatrix} 1 & 3 \\ 3 & -7 \end{bmatrix} \quad \text{and} \quad \Delta(t) = t^2 + 6t - 16 = (t - 2)(t + 8)$$

The eigenvalues of A are $\lambda = 2$ and $\lambda = -8$. Thus, using s and t as new variables, a diagonal form of q is

$$q(s, t) = 2s^2 - 8t^2$$

The corresponding orthogonal substitution is obtained by finding an orthogonal set of eigenvectors of A.

(i) Subtract $\lambda = 2$ down the diagonal of A to obtain the matrix

$$M = \begin{bmatrix} -1 & 3 \\ 3 & -9 \end{bmatrix}, \quad \text{corresponding to} \quad \begin{array}{r} -x + 3y = 0 \\ 3x - 9y = 0 \end{array} \quad \text{or} \quad -x + 3y = 0$$

A nonzero solution is $u_1 = (3, 1)$.

(ii) Subtract $\lambda = -8$ (or add 8) down the diagonal of A to obtain the matrix

$$M = \begin{bmatrix} 9 & 3 \\ 3 & 1 \end{bmatrix}, \quad \text{corresponding to} \quad \begin{array}{r} 9x + 3y = 0 \\ 3x + \ \ y = 0 \end{array} \quad \text{or} \quad 3x + y = 0$$

A nonzero solution is $u_2 = (-1, 3)$.

As expected, because A is symmetric, the eigenvectors u_1 and u_2 are orthogonal.

Now normalize u_1 and u_2 to obtain, respectively, the unit vectors

$$\hat{u}_1 = (3/\sqrt{10},\ 1/\sqrt{10}) \quad \text{and} \quad \hat{u}_2 = (-1/\sqrt{10},\ 3/\sqrt{10}).$$

Finally, let P be the matrix whose columns are the unit vectors $\hat{u}_1$ and $\hat{u}_2$, respectively, and then $[x, y]^T = P[s, t]^T$ is the required orthogonal change of coordinates. That is,

$$P = \begin{bmatrix} 3/\sqrt{10} & -1/\sqrt{10} \\ 1/\sqrt{10} & 3/\sqrt{10} \end{bmatrix} \quad \text{and} \quad x = \frac{3s - t}{\sqrt{10}}, \qquad y = \frac{s + 3t}{\sqrt{10}}$$

One can also express s and t in terms of x and y by using $P^{-1} = P^T$. That is,

$$s = \frac{3x + y}{\sqrt{10}}, \qquad t = \frac{-x + 3t}{\sqrt{10}}$$

Minimal Polynomial

9.27. Let $A = \begin{bmatrix} 4 & -2 & 2 \\ 6 & -3 & 4 \\ 3 & -2 & 3 \end{bmatrix}$ and $B = \begin{bmatrix} 3 & -2 & 2 \\ 4 & -4 & 6 \\ 2 & -3 & 5 \end{bmatrix}$. The characteristic polynomial of both matrices is $\Delta(t) = (t-2)(t-1)^2$. Find the minimal polynomial $m(t)$ of each matrix.

The minimal polynomial $m(t)$ must divide $\Delta(t)$. Also, each factor of $\Delta(t)$ (i.e., $t-2$ and $t-1$) must also be a factor of $m(t)$. Thus, $m(t)$ must be exactly one of the following:

$$f(t) = (t-2)(t-1) \qquad \text{or} \qquad g(t) = (t-2)(t-1)^2$$

(a) By the Cayley–Hamilton theorem, $g(A) = \Delta(A) = 0$, so we need only test $f(t)$. We have

$$f(A) = (A - 2I)(A - I) = \begin{bmatrix} 2 & -2 & 2 \\ 6 & -5 & 4 \\ 3 & -2 & 1 \end{bmatrix} \begin{bmatrix} 3 & -2 & 2 \\ 6 & -4 & 4 \\ 3 & -2 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Thus, $m(t) = f(t) = (t-2)(t-1) = t^2 - 3t + 2$ is the minimal polynomial of A.

(b) Again $g(B) = \Delta(B) = 0$, so we need only test $f(t)$. We get

$$f(B) = (B - 2I)(B - I) = \begin{bmatrix} 1 & -2 & 2 \\ 4 & -6 & 6 \\ 2 & -3 & 3 \end{bmatrix} \begin{bmatrix} 2 & -2 & 2 \\ 4 & -5 & 6 \\ 2 & -3 & 4 \end{bmatrix} = \begin{bmatrix} -2 & 2 & -2 \\ -4 & 4 & -4 \\ -2 & 2 & -2 \end{bmatrix} \neq 0$$

Thus, $m(t) \neq f(t)$. Accordingly, $m(t) = g(t) = (t-2)(t-1)^2$ is the minimal polynomial of B. [We emphasize that we do not need to compute $g(B)$; we know $g(B) = 0$ from the Cayley–Hamilton theorem.]

9.28. Find the minimal polynomial $m(t)$ of each of the following matrices:

(a) $A = \begin{bmatrix} 5 & 1 \\ 3 & 7 \end{bmatrix}$, (b) $B = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & 3 \\ 0 & 0 & 3 \end{bmatrix}$, (c) $C = \begin{bmatrix} 4 & -1 \\ 1 & 2 \end{bmatrix}$

(a) The characteristic polynomial of A is $\Delta(t) = t^2 - 12t + 32 = (t-4)(t-8)$. Because $\Delta(t)$ has distinct factors, the minimal polynomial $m(t) = \Delta(t) = t^2 - 12t + 32$.

(b) Because B is triangular, its eigenvalues are the diagonal elements $1, 2, 3$; and so its characteristic polynomial is $\Delta(t) = (t-1)(t-2)(t-3)$. Because $\Delta(t)$ has distinct factors, $m(t) = \Delta(t)$.

(c) The characteristic polynomial of C is $\Delta(t) = t^2 - 6t + 9 = (t-3)^2$. Hence the minimal polynomial of C is $f(t) = t - 3$ or $g(t) = (t-3)^2$. However, $f(C) \neq 0$; that is, $C - 3I \neq 0$. Hence,

$$m(t) = g(t) = \Delta(t) = (t-3)^2.$$

9.29. Suppose $S = \{u_1, u_2, \ldots, u_n\}$ is a basis of V, and suppose F and G are linear operators on V such that $[F]$ has 0's on and below the diagonal, and $[G]$ has $a \neq 0$ on the superdiagonal and 0's elsewhere. That is,

$$[F] = \begin{bmatrix} 0 & a_{21} & a_{31} & \ldots & a_{n1} \\ 0 & 0 & a_{32} & \ldots & a_{n2} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & a_{n,n-1} \\ 0 & 0 & 0 & \ldots & 0 \end{bmatrix}, \qquad [G] = \begin{bmatrix} 0 & a & 0 & \ldots & 0 \\ 0 & 0 & a & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & a \\ 0 & 0 & 0 & \ldots & 0 \end{bmatrix}$$

Show that (a) $F^n = 0$, (b) $G^{n-1} \neq 0$, but $G^n = 0$. (These conditions also hold for $[F]$ and $[G]$.)

(a) We have $F(u_1) = 0$ and, for $r > 1$, $F(u_r)$ is a linear combination of vectors preceding u_r in S. That is,

$$F(u_r) = a_{r1}u_1 + a_{r2}u_2 + \cdots + a_{r,r-1}u_{r-1}$$

Hence, $F^2(u_r) = F(F(u_r))$ is a linear combination of vectors preceding u_{r-1}, and so on. Hence, $F^r(u_r) = 0$ for each r. Thus, for each r, $F^n(u_r) = F^{n-r}(0) = 0$, and so $F^n = 0$, as claimed.

(b) We have $G(u_1) = 0$ and, for each $k > 1$, $G(u_k) = au_{k-1}$. Hence, $G^r(u_k) = a^r u_{k-r}$ for $r < k$. Because $a \neq 0$, $a^{n-1} \neq 0$. Therefore, $G^{n-1}(u_n) = a^{n-1}u_1 \neq 0$, and so $G^{n-1} \neq 0$. On the other hand, by (a), $G^n = 0$.

9.30. Let B be the matrix in Example 9.12(a) that has 1's on the diagonal, a's on the superdiagonal, where $a \neq 0$, and 0's elsewhere. Show that $f(t) = (t - \lambda)^n$ is both the characteristic polynomial $\Delta(t)$ and the minimum polynomial $m(t)$ of A.

Because A is triangular with λ's on the diagonal, $\Delta(t) = f(t) = (t - \lambda)^n$ is its characteristic polynomial. Thus, $m(t)$ is a power of $t - \lambda$. By Problem 9.29, $(A - \lambda I)^{r-1} \neq 0$. Hence, $m(t) = \Delta(t) = (t - \lambda)^n$.

9.31. Find the characteristic polynomial $\Delta(t)$ and minimal polynomial $m(t)$ of each matrix:

$$\text{(a)}\ M = \begin{bmatrix} 4 & 1 & 0 & 0 & 0 \\ 0 & 4 & 1 & 0 & 0 \\ 0 & 0 & 4 & 0 & 0 \\ 0 & 0 & 0 & 4 & 1 \\ 0 & 0 & 0 & 0 & 4 \end{bmatrix}, \quad \text{(b)}\quad M' = \begin{bmatrix} 2 & 7 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & -2 & 4 \end{bmatrix}$$

(a) M is block diagonal with diagonal blocks

$$A = \begin{bmatrix} 4 & 1 & 0 \\ 0 & 4 & 1 \\ 0 & 0 & 4 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 4 & 1 \\ 0 & 4 \end{bmatrix}$$

The characteristic and minimal polynomial of A is $f(t) = (t - 4)^3$ and the characteristic and minimal polynomial of B is $g(t) = (t - 4)^2$. Then

$$\Delta(t) = f(t)g(t) = (t - 4)^5 \quad \text{but} \quad m(t) = \text{LCM}[f(t), g(t)] = (t - 4)^3$$

(where LCM means least common multiple). We emphasize that the exponent in $m(t)$ is the size of the largest block.

(b) Here M' is block diagonal with diagonal blocks $A' = \begin{bmatrix} 2 & 7 \\ 0 & 2 \end{bmatrix}$ and $B' = \begin{bmatrix} 1 & 1 \\ -2 & 4 \end{bmatrix}$ The characteristic and minimal polynomial of A' is $f(t) = (t - 2)^2$. The characteristic polynomial of B' is $g(t) = t^2 - 5t + 6 = (t - 2)(t - 3)$, which has distinct factors. Hence, $g(t)$ is also the minimal polynomial of B. Accordingly,

$$\Delta(t) = f(t)g(t) = (t - 2)^3(t - 3) \quad \text{but} \quad m(t) = \text{LCM}[f(t), g(t)] = (t - 2)^2(t - 3)$$

9.32. Find a matrix A whose minimal polynomial is $f(t) = t^3 - 8t^2 + 5t + 7$.

Simply let $A = \begin{bmatrix} 0 & 0 & -7 \\ 1 & 0 & -5 \\ 0 & 1 & 8 \end{bmatrix}$, the companion matrix of $f(t)$ [defined in Example 9.12(b)].

9.33. Prove Theorem 9.15: The minimal polynomial $m(t)$ of a matrix (linear operator) A divides every polynomial that has A as a zero. In particular (by the Cayley–Hamilton theorem), $m(t)$ divides the characteristic polynomial $\Delta(t)$ of A.

Suppose $f(t)$ is a polynomial for which $f(A) = 0$. By the division algorithm, there exist polynomials $q(t)$ and $r(t)$ for which $f(t) = m(t)q(t) + r(t)$ and $r(t) = 0$ or $\deg r(t) < \deg m(t)$. Substituting $t = A$ in this equation, and using that $f(A) = 0$ and $m(A) = 0$, we obtain $r(A) = 0$. If $r(t) \neq 0$, then $r(t)$ is a polynomial of degree less than $m(t)$ that has A as a zero. This contradicts the definition of the minimal polynomial. Thus, $r(t) = 0$, and so $f(t) = m(t)q(t)$; that is, $m(t)$ divides $f(t)$.

9.34. Let $m(t)$ be the minimal polynomial of an n-square matrix A. Prove that the characteristic polynomial $\Delta(t)$ of A divides $[m(t)]^n$.

Suppose $m(t) = t^r + c_1 t^{r-1} + \cdots + c_{r-1}t + c_r$. Define matrices B_j as follows:

$$\begin{array}{lll}
B_0 = I & \text{so} & I = B_0 \\
B_1 = A + c_1 I & \text{so} & c_1 I = B_1 - A = B_1 - AB_0 \\
B_2 = A^2 + c_1 A + c_2 I & \text{so} & c_2 I = B_2 - A(A + c_1 I) = B_2 - AB_1 \\
\cdots\cdots & & \cdots\cdots \\
B_{r-1} = A^{r-1} + c_1 A^{r-2} + \cdots + c_{r-1} I & \text{so} & c_{r-1} I = B_{r-1} - AB_{r-2}
\end{array}$$

Then

$$-AB_{r-1} = c_r I - (A^r + c_1 A^{r-1} + \cdots + c_{r-1}A + c_r I) = c_r I - m(A) = c_r I$$

Set
$$B(t) = t^{r-1}B_0 + t^{r-2}B_1 + \cdots + tB_{r-2} + B_{r-1}$$

Then

$$\begin{aligned}
(tI - A)B(t) &= (t^r B_0 + t^{r-1}B_1 + \cdots + tB_{r-1}) - (t^{r-1}AB_0 + t^{r-2}AB_1 + \cdots + AB_{r-1}) \\
&= t^r B_0 + t^{r-1}(B_1 - AB_0) + t^{r-2}(B_2 - AB_1) + \cdots + t(B_{r-1} - AB_{r-2}) - AB_{r-1} \\
&= t^r I + c_1 t^{r-1} I + c_2 t^{r-2} I + \cdots + c_{r-1} tI + c_r I = m(t)I
\end{aligned}$$

Taking the determinant of both sides gives $|tI - A||B(t)| = |m(t)I| = [m(t)]^n$. Because $|B(t)|$ is a polynomial, $|tI - A|$ divides $[m(t)]^n$; that is, the characteristic polynomial of A divides $[m(t)]^n$.

9.35. Prove Theorem 9.16: The characteristic polynomial $\Delta(t)$ and the minimal polynomial $m(t)$ of A have the same irreducible factors.

Suppose $f(t)$ is an irreducible polynomial. If $f(t)$ divides $m(t)$, then $f(t)$ also divides $\Delta(t)$ [because $m(t)$ divides $\Delta(t)$]. On the other hand, if $f(t)$ divides $\Delta(t)$, then by Problem 9.34, $f(t)$ also divides $[m(t)]^n$. But $f(t)$ is irreducible; hence, $f(t)$ also divides $m(t)$. Thus, $m(t)$ and $\Delta(t)$ have the same irreducible factors.

9.36. Prove Theorem 9.19: The minimal polynomial $m(t)$ of a block diagonal matrix M with diagonal blocks A_i is equal to the least common multiple (LCM) of the minimal polynomials of the diagonal blocks A_i.

We prove the theorem for the case $r = 2$. The general theorem follows easily by induction. Suppose $M = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$, where A and B are square matrices. We need to show that the minimal polynomial $m(t)$ of M is the LCM of the minimal polynomials $g(t)$ and $h(t)$ of A and B, respectively.

Because $m(t)$ is the minimal polynomial of $M, m(M) = \begin{bmatrix} m(A) & 0 \\ 0 & m(B) \end{bmatrix} = 0$, and $m(A) = 0$ and $m(B) = 0$. Because $g(t)$ is the minimal polynomial of A, $g(t)$ divides $m(t)$. Similarly, $h(t)$ divides $m(t)$. Thus $m(t)$ is a multiple of $g(t)$ and $h(t)$.

Now let $f(t)$ be another multiple of $g(t)$ and $h(t)$. Then $f(M) = \begin{bmatrix} f(A) & 0 \\ 0 & f(B) \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = 0$. But $m(t)$ is the minimal polynomial of M; hence, $m(t)$ divides $f(t)$. Thus, $m(t)$ is the LCM of $g(t)$ and $h(t)$.

9.37. Suppose $m(t) = t^r + a_{r-1}t^{r-1} + \cdots + a_1 t + a_0$ is the minimal polynomial of an n-square matrix A. Prove the following:

(a) A is nonsingular if and only if the constant term $a_0 \neq 0$.

(b) If A is nonsingular, then A^{-1} is a polynomial in A of degree $r - 1 < n$.

(a) The following are equivalent: (i) A is nonsingular, (ii) 0 is not a root of $m(t)$, (iii) $a_0 \neq 0$. Thus, the statement is true.

(b) Because A is nonsingular, $a_0 \neq 0$ by (a). We have

$$m(A) = A^r + a_{r-1}A^{r-1} + \cdots + a_1 A + a_0 I = 0$$

Thus,
$$-\frac{1}{a_0}(A^{r-1} + a_{r-1}A^{r-2} + \cdots + a_1 I)A = I$$

Accordingly,
$$A^{-1} = -\frac{1}{a_0}(A^{r-1} + a_{r-1}A^{r-2} + \cdots + a_1 I)$$

SUPPLEMENTARY PROBLEMS

Polynomials of Matrices

9.38. Let $A = \begin{bmatrix} 2 & -3 \\ 5 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}$. Find $f(A)$, $g(A)$, $f(B)$, $g(B)$, where $f(t) = 2t^2 - 5t + 6$ and $g(t) = t^3 - 2t^2 + t + 3$.

9.39. Let $A = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$. Find A^2, A^3, A^n, where $n > 3$, and A^{-1}.

9.40. Let $B = \begin{bmatrix} 8 & 12 & 0 \\ 0 & 8 & 12 \\ 0 & 0 & 8 \end{bmatrix}$. Find a real matrix A such that $B = A^3$.

9.41. For each matrix, find a polynomial having the following matrix as a root:

(a) $A = \begin{bmatrix} 2 & 5 \\ 1 & -3 \end{bmatrix}$, (b) $B = \begin{bmatrix} 2 & -3 \\ 7 & -4 \end{bmatrix}$, (c) $C = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 2 & 3 \\ 2 & 1 & 4 \end{bmatrix}$

9.42. Let A be any square matrix and let $f(t)$ be any polynomial. Prove (a) $(P^{-1}AP)^n = P^{-1}A^nP$.
(b) $f(P^{-1}AP) = P^{-1}f(A)P$. (c) $f(A^T) = [f(A)]^T$. (d) If A is symmetric, then $f(A)$ is symmetric.

9.43. Let $M = \text{diag}[A_1, \ldots, A_r]$ be a block diagonal matrix, and let $f(t)$ be any polynomial. Show that $f(M)$ is block diagonal and $f(M) = \text{diag}[f(A_1), \ldots, f(A_r)]$.

9.44. Let M be a block triangular matrix with diagonal blocks $A_1, \ldots, A_r$, and let $f(t)$ be any polynomial. Show that $f(M)$ is also a block triangular matrix, with diagonal blocks $f(A_1), \ldots, f(A_r)$.

Eigenvalues and Eigenvectors

9.45. For each of the following matrices, find all eigenvalues and corresponding linearly independent eigenvectors:

(a) $A = \begin{bmatrix} 2 & -3 \\ 2 & -5 \end{bmatrix}$, (b) $B = \begin{bmatrix} 2 & 4 \\ -1 & 6 \end{bmatrix}$, (c) $C = \begin{bmatrix} 1 & -4 \\ 3 & -7 \end{bmatrix}$

When possible, find the nonsingular matrix P that diagonalizes the matrix.

9.46. Let $A = \begin{bmatrix} 2 & -1 \\ -2 & 3 \end{bmatrix}$.

(a) Find eigenvalues and corresponding eigenvectors.
(b) Find a nonsingular matrix P such that $D = P^{-1}AP$ is diagonal.
(c) Find A^8 and $f(A)$ where $f(t) = t^4 - 5t^3 + 7t^2 - 2t + 5$.
(d) Find a matrix B such that $B^2 = A$.

9.47. Repeat Problem 9.46 for $A = \begin{bmatrix} 5 & 6 \\ -2 & -2 \end{bmatrix}$.

9.48. For each of the following matrices, find all eigenvalues and a maximum set S of linearly independent eigenvectors:

(a) $A = \begin{bmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{bmatrix}$, (b) $B = \begin{bmatrix} 3 & -1 & 1 \\ 7 & -5 & 1 \\ 6 & -6 & 2 \end{bmatrix}$, (c) $C = \begin{bmatrix} 1 & 2 & 2 \\ 1 & 2 & -1 \\ -1 & 1 & 4 \end{bmatrix}$

Which matrices can be diagonalized, and why?

9.49. For each of the following linear operators $T: \mathbf{R}^2 \to \mathbf{R}^2$, find all eigenvalues and a basis for each eigenspace:

(a) $T(x, y) = (3x + 3y,\ x + 5y)$, (b) $T(x, y) = (3x - 13y,\ x - 3y)$.

9.50. Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ be a real matrix. Find necessary and sufficient conditions on a, b, c, d so that A is diagonalizable—that is, so that A has two (real) linearly independent eigenvectors.

9.51. Show that matrices A and A^T have the same eigenvalues. Give an example of a 2×2 matrix A where A and A^T have different eigenvectors.

9.52. Suppose v is an eigenvector of linear operators F and G. Show that v is also an eigenvector of the linear operator $kF + k'G$, where k and k' are scalars.

9.53. Suppose v is an eigenvector of a linear operator T belonging to the eigenvalue λ. Prove

(a) For $n > 0$, v is an eigenvector of T^n belonging to λ^n.

(b) $f(\lambda)$ is an eigenvalue of $f(T)$ for any polynomial $f(t)$.

9.54. Suppose $\lambda \neq 0$ is an eigenvalue of the composition $F \circ G$ of linear operators F and G. Show that λ is also an eigenvalue of the composition $G \circ F$. [*Hint:* Show that $G(v)$ is an eigenvector of $G \circ F$.]

9.55. Let $E: V \to V$ be a projection mapping; that is, $E^2 = E$. Show that E is diagonalizable and, in fact, can be represented by the diagonal matrix $M = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$, where r is the rank of E.

Diagonalizing Real Symmetric Matrices and Quadratic Forms

9.56. For each of the following symmetric matrices A, find an orthogonal matrix P and a diagonal matrix D such that $D = P^{-1}AP$:

(a) $A = \begin{bmatrix} 5 & 4 \\ 4 & -1 \end{bmatrix}$, (b) $A = \begin{bmatrix} 4 & -1 \\ -1 & 4 \end{bmatrix}$, (c) $A = \begin{bmatrix} 7 & 3 \\ 3 & -1 \end{bmatrix}$

9.57. For each of the following symmetric matrices B, find its eigenvalues, a maximal orthogonal set S of eigenvectors, and an orthogonal matrix P such that $D = P^{-1}BP$ is diagonal:

(a) $B = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$, (b) $B = \begin{bmatrix} 2 & 2 & 4 \\ 2 & 5 & 8 \\ 4 & 8 & 17 \end{bmatrix}$

9.58. Using variables s and t, find an orthogonal substitution that diagonalizes each of the following quadratic forms:

(a) $q(x, y) = 4x^2 + 8xy - 11y^2$, (b) $q(x, y) = 2x^2 - 6xy + 10y^2$

9.59. For each of the following quadratic forms $q(x, y, z)$, find an orthogonal substitution expressing x, y, z in terms of variables r, s, t, and find $q(r, s, t)$:

(a) $q(x, y, z) = 5x^2 + 3y^2 + 12xz$, (b) $q(x, y, z) = 3x^2 - 4xy + 6y^2 + 2xz - 4yz + 3z^2$

9.60. Find a real 2×2 symmetric matrix A with eigenvalues:

(a) $\lambda = 1$ and $\lambda = 4$ and eigenvector $u = (1, 1)$ belonging to $\lambda = 1$;

(b) $\lambda = 2$ and $\lambda = 3$ and eigenvector $u = (1, 2)$ belonging to $\lambda = 2$.

In each case, find a matrix B for which $B^2 = A$.

Characteristic and Minimal Polynomials

9.61. Find the characteristic and minimal polynomials of each of the following matrices:

$$\text{(a)}\quad A = \begin{bmatrix} 3 & 1 & -1 \\ 2 & 4 & -2 \\ -1 & -1 & 3 \end{bmatrix}, \text{(b)}\quad B = \begin{bmatrix} 3 & 2 & -1 \\ 3 & 8 & -3 \\ 3 & 6 & -1 \end{bmatrix}$$

9.62. Find the characteristic and minimal polynomials of each of the following matrices:

$$\text{(a)}\quad A = \begin{bmatrix} 2 & 5 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 4 & 2 & 0 \\ 0 & 0 & 3 & 5 & 0 \\ 0 & 0 & 0 & 0 & 7 \end{bmatrix}, \quad \text{(b)}\quad B = \begin{bmatrix} 4 & -1 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 & 0 \\ 0 & 0 & 3 & 1 & 0 \\ 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 3 \end{bmatrix}, \quad \text{(c)}\quad C = \begin{bmatrix} 3 & 2 & 0 & 0 & 0 \\ 1 & 4 & 0 & 0 & 0 \\ 0 & 0 & 3 & 1 & 0 \\ 0 & 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 0 & 4 \end{bmatrix}$$

9.63. Let $A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 2 \\ 0 & 0 & 1 \end{bmatrix}$. Show that A and B have different characteristic polynomials (and so are not similar) but have the same minimal polynomial. Thus, nonsimilar matrices may have the same minimal polynomial.

9.64. Let A be an n-square matrix for which $A^k = 0$ for some $k > n$. Show that $A^n = 0$.

9.65. Show that a matrix A and its transpose A^T have the same minimal polynomial.

9.66. Suppose $f(t)$ is an irreducible monic polynomial for which $f(A) = 0$ for a matrix A. Show that $f(t)$ is the minimal polynomial of A.

9.67. Show that A is a scalar matrix kI if and only if the minimal polynomial of A is $m(t) = t - k$.

9.68. Find a matrix A whose minimal polynomial is (a) $t^3 - 5t^2 + 6t + 8$, (b) $t^4 - 5t^3 - 2t + 7t + 4$.

9.69. Let $f(t)$ and $g(t)$ be monic polynomials (leading coefficient one) of minimal degree for which A is a root. Show $f(t) = g(t)$. [Thus, the minimal polynomial of A is unique.]

ANSWERS TO SUPPLEMENTARY PROBLEMS

Notation: $M = [R_1; \quad R_2; \quad \ldots]$ denotes a matrix M with rows $R_1, R_2, \ldots$.

9.38. $f(A) = [-26, -3; \quad 5, -27]$, $\quad g(A) = [-40, 39; \quad -65, -27]$,
$f(B) = [3, 6; \quad 0, 9]$, $\quad g(B) = [3, 12; \quad 0, 15]$

9.39. $A^2 = [1, 4; \quad 0, 1]$, $\quad A^3 = [1, 6; \quad 0, 1]$, $\quad A^n = [1, 2n; \quad 0, 1]$, $\quad A^{-1} = [1, -2; \quad 0, 1]$

9.40. Let $A = [2, a, b; \quad 0, 2, c; \quad 0, 0, 2]$. Set $B = A^3$ and then $a = 1$, $b = -\frac{1}{2}$, $c = 1$

9.41. Find $\Delta(t)$: (a) $t^2 + t - 11$, (b) $t^2 + 2t + 13$, (c) $t^3 - 7t^2 + 6t - 1$

9.45. (a) $\lambda = 1, u = (3, 1)$; $\lambda = -4, v = (1, 2)$, (b) $\lambda = 4, u = (2, 1)$,
(c) $\lambda = -1, u = (2, 1)$; $\lambda = -5, v = (2, 3)$. Only A and C can be diagonalized; use $P = [u, v]$.

9.46. (a) $\lambda = 1, u = (1, 1)$; $\lambda = 4, v = (1, -2)$,
(b) $P = [u, v]$,
(c) $f(A) = [3, 1;\quad 2, 1]$, $A^8 = [21\,846, -21\,845;\quad -43\,690, 43\,691]$,
(d) $B = \left[\frac{4}{3}, -\frac{1}{3};\quad -\frac{2}{3}, \frac{5}{3}\right]$

9.47. (a) $\lambda = 1, u = (3, -2)$; $\lambda = 2, v = (2, -1)$, (b) $P = [u, v]$,
(c) $f(A) = [2, -6;\quad 2, 9]$, $A^8 = [1021, 1530;\quad -510, -764]$,
(d) $B = [-3 + 4\sqrt{2},\quad -6 + 6\sqrt{2};\quad 2 - 2\sqrt{2},\quad 4 - 3\sqrt{2}]$

9.48. (a) $\lambda = -2, u = (1, 1, 0), v = (1, 0, -1); \lambda = 4, w = (1, 1, 2)$,
(b) $\lambda = 2, u = (1, 1, 0)$; $\lambda = -4, v = (0, 1, 1)$,
(c) $\lambda = 3, u = (1, 1, 0), v = (1, 0, 1)$; $\lambda = 1, w = (2, -1, 1)$. Only A and C can be diagonalized; use $P = [u, v, w]$.

9.49. (a) $\lambda = 2, u = (3, -1)$; $\lambda = 6, v = (1, 1)$, (b) No real eigenvalues

9.50. We need $[-\mathrm{tr}(A)]^2 - 4[\det(A)] \geq 0$ or $(a - d)^2 + 4bc \geq 0$.

9.51. $A = [1, 1;\quad 0, 1]$

9.56. (a) $P = [2, -1;\quad 1, 2]/\sqrt{5}$, $D = [7, 0;\quad 0, 3]$,
(b) $P = [1, 1;\quad 1, -1]/\sqrt{2}$, $D = [3, 0;\quad 0, 5]$,
(c) $P = [3, -1;\quad 1, 3]/\sqrt{10}$, $D = [8, 0;\quad 0, 2]$

9.57. (a) $\lambda = -1$, $u = (1, -1, 0)$, $v = (1, 1, -2)$; $\lambda = 2$, $w = (1, 1, 1)$,
(b) $\lambda = 1$, $u = (2, 1, -1)$, $v = (2, -3, 1)$; $\lambda = 22$, $w = (1, 2, 4)$;
Normalize u, v, w, obtaining $\hat{u}, \hat{v}, \hat{w}$, and set $P = [\hat{u}, \hat{v}, \hat{w}]$. (*Remark:* u and v are not unique.)

9.58. (a) $x = (4s + t)/\sqrt{17}$, $y = (-s + 4t)/\sqrt{17}$, $q(s, t) = 5s^2 - 12t^2$,
(b) $x = (3s - t)/\sqrt{10}$, $y = (s + 3t)/\sqrt{10}$, $q(s, t) = s^2 + 11t^2$

9.59. (a) $x = (3s + 2t)/\sqrt{13}$, $y = r$, $z = (2s - 3t)/\sqrt{13}$, $q(r, s, t) = 3r^2 + 9s^2 - 4t^2$,
(b) $x = 5Ks + Lt$, $y = Jr + 2Ks - 2Lt$, $z = 2Jr - Ks - Lt$, where $J = 1/\sqrt{5}$, $K = 1/\sqrt{30}$, $L = 1/\sqrt{6}$; $q(r, s, t) = 2r^2 + 2s^2 + 8t^2$

9.60. (a) $A = \frac{1}{2}[5, -3;\ -3, 5]$, $B = \frac{1}{2}[3, -1;\ -1, 3]$,
(b) $A = \frac{1}{5}[14, -2;\ -2, 11]$, $B = \frac{1}{5}[\sqrt{2} + 4\sqrt{3}, 2\sqrt{2} - 2\sqrt{3};\ 2\sqrt{2} - 2\sqrt{3}, 4\sqrt{2} + \sqrt{3}]$

9.61. (a) $\Delta(t) = m(t) = (t - 2)^2(t - 6)$, (b) $\Delta(t) = (t - 2)^2(t - 6)$, $m(t) = (t - 2)(t - 6)$

9.62. (a) $\Delta(t) = (t - 2)^3(t - 7)^2$, $m(t) = (t - 2)^2(t - 7)$,
(b) $\Delta(t) = (t - 3)^5$, $m(t) = (t - 3)^3$,
(c) $\Delta(t) = (t - 2)^2(t - 4)^2(t - 5)$, $m(t) = (t - 2)(t - 4)(t - 5)$

9.68. Let A be the companion matrix [Example 9.12(b)] with last column: (a) $[-8, -6, 5]^T$, (b) $[-4, -7, 2, 5]^T$

9.69. *Hint*: A is a root of $h(t) = f(t) - g(t)$, where $h(t) \equiv 0$ or the degree of $h(t)$ is less than the degree of $f(t)$.