Chapter 8

Eigenvalues and Eigenvectors, Diagonalization

8.1 INTRODUCTION

Consider an n-square matrix A over a field K. Recall (Section 4.13) that A induces a function $f: K^n \to K^n$ defined by

$$f(X) = AX$$

where X is any point (column vector) in K^n. (We then view A as the matrix which represents the function f relative to the usual basis E for K^n.)

Suppose a new basis is chosen for K^n, say

$$S = \{u_1, u_2, \ldots, u_n\}$$

(Geometrically, S determines a new coordinate system for K^n.) Let P be the matrix whose columns are the vectors $u_1, u_2, \ldots, u_n$. Then (Section 5.11) P is the change-of-basis matrix from the usual basis E to S. Also, by Theorem 5.27,

$$X' = P^{-1}X$$

gives the coordinates of X in the new basis S. Furthermore, the matrix

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

represents the function f in the new system S; that is, $f(X') = BX'$.

The following two questions are addressed in this chapter:

(1) Given a matrix A, can we find a nonsingular matrix P (which represents a new coordinate system S), so that

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

is a diagonal matrix? If the answer is yes, then we say that A is *diagonalizable.*

(2) Given a real matrix A, can we find an orthogonal matrix P (which represents a new orthonormal system S) so that

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

is a diagonal matrix? If the answer is yes, then we say that A is *orthogonally diagonalizable.*

Recall that matrices A and B are said to be similar (orthogonally similar) if there exists a nonsingular (orthogonal) matrix P such that $B = P^{-1}AP$. What is in question, then, is whether or not a given matrix A is similar (orthogonally similar) to a diagonal matrix.

The answers are closely related to the roots of certain polynomials associated with A. The particular underlying field K also plays an important part in this theory since the existence of roots of the polynomials depends on K. In this connection, see the Appendix (page 446).

8.2 POLYNOMIALS IN MATRICES

Consider a polynomial $f(t)$ over a field K; say

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

Recall that if A is a square matrix over K, 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 or zero of the polynomial $f(t)$ if $f(A) = 0$.

Example 8.1. Let $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$, and let $f(t) = 2t^2 - 3t + 7$, $g(t) = t^2 - 5t - 2$. Then

$$f(A) = 2\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}^2 - 3\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} + 7\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 18 & 14 \\ 21 & 39 \end{pmatrix}$$

and

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

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

The following theorem, proved in Problem 8.26, applies.

Theorem 8.1: Let f and g be polynomials over K, and let A be an n-square matrix over K. Then

(i) $(f + g)(A) = f(A) + g(A)$

(ii) $(fg)(A) = f(A)g(A)$

(iii) $(kf)(A) = kf(A)$ for all $k \in K$

(iv) $f(A)g(A) = g(A)f(A)$

By (iv), any two polynomials in the matrix A commute.

8.3 CHARACTERISTIC POLYNOMIAL, CAYLEY–HAMILTON THEOREM

Consider an n-square matrix A over a field K:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

The matrix $tI_n - A$, where I_n is the n-square identity matrix and t is an indeterminate, is called the *characteristic matrix* of A:

$$tI_n - A = \begin{pmatrix} t - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & t - a_{22} & \cdots & -a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ -a_{n1} & -a_{n2} & \cdots & t - a_{nn} \end{pmatrix}$$

Its determinant

$$\Delta_A(t) = \det(tI_n - A)$$

which is a polynomial in t, is called the *characteristic polynomial* of A. We also call

$$\Delta_A(t) = \det(tI_n - A) = 0$$

the *characteristic equation* of A.

Now each term in the determinant contains one and only one entry from each row and from each column; hence the above characteristic polynomial is of the form

$$\Delta_A(t) = (t - a_{11})(t - a_{22}) \cdots (t - a_{nn}) + \text{terms with at most } n - 2 \text{ factors of the form } t - a_{ii}$$

Accordingly,

$$\Delta_A(t) = t^n - (a_{11} + a_{22} + \cdots + a_{nn})t^{n-1} + \text{terms of lower degree}$$

Recall that the trace of A is the sum of its diagonal elements. Thus the characteristic polynomial $\Delta_A(t) = \det(tI_n - A)$ of A is a monic polynomial of degree n, and the coefficient of t^{n-1} is the negative of the trace of A. (A polynomial is *monic* if its leading coefficient is 1.)

Furthermore, if we set $t = 0$ in $\Delta_A(t)$, we obtain

$$\Delta_A(0) = |-A| = (-1)^n |A|$$

But $\Delta_A(0)$ is the constant term of the polynomial $\Delta_A(t)$. Thus the constant term of the characteristic polynomial of the matrix A is $(-1)^n|A|$ where n is the order of A.

We now state one of the most important theorems in linear algebra (proved in Problem 8.27):

Cayley–Hamilton Theorem 8.2: Every matrix is a zero of its characteristic polynomial.

Example 8.2. Let $B = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}$. Its characteristic polynomial is

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

As expected from the Cayley–Hamilton Theorem, B is a zero of $\Delta(t)$:

$$\Delta(B) = B^2 - 3B - 4I = \begin{pmatrix} 7 & 6 \\ 9 & 10 \end{pmatrix} + \begin{pmatrix} -3 & -6 \\ -9 & -6 \end{pmatrix} + \begin{pmatrix} -4 & 0 \\ 0 & -4 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

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 $tI = P^{-1}tIP$,

$$\begin{aligned} |tI - B| &= |tI - P^{-1}AP| = |P^{-1}tIP - P^{-1}AP| \\ &= |P^{-1}(tI - A)P| = |P^{-1}||tI - A||P| \end{aligned}$$

Since determinants are scalars and commute, and since $|P^{-1}||P| = 1$, we finally obtain

$$|tI - B| = |tI - A|$$

Thus we have proved

Theorem 8.3: Similar matrices have the same characteristic polynomial.

Characteristic Polynomials of Degree Two and Three

Let A be a matrix of order two or three. Then there is an easy formula for its characteristic polynomial $\Delta(t)$. Specifically:

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

$$\Delta(t) = t^2 - (a_{11} + a_{22})t + \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = t^2 - (\operatorname{tr} A)t + \det(A)$$

(Here tr A denotes the trace of A, that is, the sum of the diagonal elements of A.)

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

$$\Delta(t) = t^3 - (a_{11} + a_{22} + a_{33})t^2 + \left(\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \right) t - \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

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

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

Consider again a 3-square matrix $A = (a_{ij})$. As noted above,

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

are the coefficients of its characteristic polynomial with alternating signs. On the other hand, each S_k is the sum of all the principal minors of A of order k. The next theorem, whose proof lies beyond the scope of this Outline, tells us that this result is true in general.

Theorem 8.4: 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 and Block Triangular Matrices

Suppose M is a block triangular matrix, say $M = \begin{pmatrix} A_1 & B \\ 0 & A_2 \end{pmatrix}$ where A_1 and A_2 are square matrices. Then the characteristic matrix of M,

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

is also a block triangular matrix with diagonal blocks $tI - A_1$ and $tI - A_2$. Thus, by Theorem 7.12,

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

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 8.5: 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,

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

Example 8.3. 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 $A = \begin{pmatrix} 9 & -1 \\ 8 & 3 \end{pmatrix}$ and $B = \begin{pmatrix} 3 & 6 \\ -1 & 8 \end{pmatrix}$. Here

$$\operatorname{tr} A = 9 + 3 = 12 \qquad \det(A) = 27 + 8 = 35 \qquad \text{and so} \qquad \Delta_A(t) = t^2 - 12t + 35 = (t-5)(t-7)$$

$$\operatorname{tr} B = 3 + 8 = 11 \qquad \det(B) = 24 + 6 = 30 \qquad \text{and so} \qquad \Delta_B(t) = t^2 - 11t + 30 = (t-5)(t-6)$$

Accordingly, the characteristic polynomial of M is the product

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

8.4 EIGENVALUES AND EIGENVECTORS

Let A be an n-square matrix over a field K. A scalar $\lambda \in K$ is called an *eigenvalue* of A if there exists a nonzero (column) vector $v \in K^n$ for which

$$Av = \lambda v$$

Every vector satisfying this relation is then called an *eigenvector* of A belonging to the eigenvalue λ. Note that each scalar multiple kv is such an eigenvector since

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

The set E_λ of all eigenvectors belonging to λ is a subspace of K^n (Problem 8.16), 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.

Example 8.4. Let $A = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}$ and let $v_1 = (2, 3)^T$ and $v_2 = (1, -1)^T$. Then

$$Av_1 = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}\begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 8 \\ 12 \end{pmatrix} = 4\begin{pmatrix} 2 \\ 3 \end{pmatrix} = 4v_1$$

and

$$Av_2 = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}\begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix} = (-1)v_2$$

Thus v_1 and v_2 are eigenvectors of A belonging, respectively, to the eigenvalues $\lambda_1 = 4$ and $\lambda_2 = -1$ of A.

The following theorem, proved in Problem 8.28, is the main tool for computing eigenvalues and eigenvectors (Section 8.5).

Theorem 8.6: Let A be an n-square matrix over a field K. Then the following are equivalent:

(i) A scalar $\lambda \in K$ is an eigenvalue of A.

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

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

The eigenspace E_λ of λ is the solution space of the homogeneous system $MX = (\lambda I - A)X = 0$. Sometimes it is more convenient to solve the homogeneous system $(A - \lambda I)X = 0$; both systems, of course, yield the same solution space.

Some matrices may have no eigenvalues and hence no eigenvectors. However, using the Fundamental Theorem of Algebra (every polynomial over **C** has a root) and Theorem 8.6, we obtain the following result.

Theorem 8.7: Let A be an n-square matrix over the complex field **C**. Then A has at least one eigenvalue.

Now suppose λ is an eigenvalue of a matrix A. The *algebraic multiplicity* of λ is defined to be the multiplicity of λ as a root of the characteristic polynomial of A. The *geometric multiplicity* of λ is defined to be the dimension of its eigenspace.

The following theorem, proved in Problem 10.27, applies.

Theorem 8.8: Let λ be an eigenvalue of a matrix A. Then the geometric multiplicity of λ does not exceed its algebraic multiplicity.

Diagonalizable Matrices

A matrix A is said to be *diagonalizable* (under similarity) if there exists a nonsingular matrix P such that $D = P^{-1}AP$ is a diagonal matrix, i.e., if A is similar to a diagonal matrix D. The following theorem, proved in Problem 8.29, characterizes such matrices.

Theorem 8.9: 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 $D = \operatorname{diag}(k_1, k_2, \ldots, k_n)$. Then

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

and, more generally, for any polynomial $f(t)$,

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

Furthermore, if the diagonal entries of D are nonnegative, then the following matrix B is a "square root" of A:

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

that is, $B^2 = A$.

Example 8.5. Consider the matrix $A = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}$. By Example 8.4, A has two linearly independent eigenvectors $\begin{pmatrix} 2 \\ 3 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$. Set $P = \begin{pmatrix} 2 & 1 \\ 3 & -1 \end{pmatrix}$, and so $P^{-1} = \begin{pmatrix} \frac{1}{5} & \frac{1}{5} \\ \frac{3}{5} & -\frac{2}{5} \end{pmatrix}$. Then A is similar to the diagonal matrix

$$B = P^{-1}AP = \begin{pmatrix} \frac{1}{5} & \frac{1}{5} \\ \frac{3}{5} & -\frac{2}{5} \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}\begin{pmatrix} 2 & 1 \\ 3 & -1 \end{pmatrix} = \begin{pmatrix} 4 & 0 \\ 0 & -1 \end{pmatrix}$$

As expected, the diagonal elements 4 and -1 of the diagonal matrix B are the eigenvalues corresponding to the given eigenvectors. In particular, A has the factorization

$$A = PDP^{-1} = \begin{pmatrix} 2 & 1 \\ 3 & -1 \end{pmatrix}\begin{pmatrix} 4 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} \frac{1}{5} & \frac{1}{5} \\ \frac{3}{5} & -\frac{2}{5} \end{pmatrix}$$

Accordingly,

$$A^4 = PD^4P^{-1} = \begin{pmatrix} 2 & 1 \\ 3 & -1 \end{pmatrix}\begin{pmatrix} 256 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} \frac{1}{5} & \frac{1}{5} \\ \frac{3}{5} & -\frac{2}{5} \end{pmatrix} = \begin{pmatrix} 103 & 102 \\ 153 & 154 \end{pmatrix}$$

Furthermore, if $f(t) = t^3 - 7t^2 + 9t - 2$, then

$$f(A) = Pf(D)P^{-1} = \begin{pmatrix} 2 & 1 \\ 3 & -1 \end{pmatrix}\begin{pmatrix} -14 & 0 \\ 0 & -19 \end{pmatrix}\begin{pmatrix} \frac{1}{5} & \frac{1}{5} \\ \frac{3}{5} & -\frac{2}{5} \end{pmatrix} = \begin{pmatrix} -17 & 2 \\ 3 & -16 \end{pmatrix}$$

Remark: Throughout this chapter, we use the fact that the inverse of the matrix

$$P = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \text{is the matrix} \quad P^{-1} = \begin{pmatrix} d/|P| & -b/|P| \\ -c/|P| & a/|P| \end{pmatrix}$$

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 $|P|$.

The following two theorems, proved in Problems 8.30 and 8.31, respectively, will be subsequently used.

Theorem 8.10: Let $v_1, \ldots, v_n$ be nonzero eigenvectors of a matrix A belonging to distinct eigenvalues $\lambda_1, \ldots, \lambda_n$. Then $v_1, \ldots, v_n$ are linearly independent.

Theorem 8.11: 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 a diagonal matrix whose diagonal elements are the a_i.

8.5 COMPUTING EIGENVALUES AND EIGENVECTORS, DIAGONALIZING MATRICES

This section computes the eigenvalues and eigenvectors for a given square matrix A and determines whether or not a nonsingular matrix P exists such that $P^{-1}AP$ is diagonal. Specifically, the following algorithm will be applied to the matrix A.

Diagonalization Algorithm 8.5:

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 $M = A - \lambda I$ by subtracting λ down the diagonal of A, or form $M' = \lambda I - A$ by substituting $t = \lambda$ in $tI - A$.

(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 $m = n$, let P be the matrix whose columns are the eigenvectors $v_1, v_2, \ldots, v_n$. Then

$$D = P^{-1}AP = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \cdots & \\ & & & \lambda_n \end{pmatrix}$$

where λ_i is the eigenvalue corresponding to the eigenvector v_i.

Example 8.6. The Diagonalization Algorithm is applied to $A = \begin{pmatrix} 4 & 2 \\ 3 & -1 \end{pmatrix}$.

1. The characteristic polynomial $\Delta(t)$ of A is the determinant

$$\Delta(t) = |tI - A| = \begin{vmatrix} t-4 & -1 \\ -3 & t+1 \end{vmatrix} = t^2 - 3t - 10 = (t-5)(t+2)$$

Alternatively, $\operatorname{tr} A = 4 - 1 = 3$ and $|A| = -4 - 6 = -10$; so $\Delta(t) = t^2 - 3t - 10$.

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 $M = \begin{pmatrix} -1 & 2 \\ 3 & -6 \end{pmatrix}$. The eigenvectors belonging to $\lambda_1 = 5$ form the solution of the homogeneous system $MX = 0$, that is,

$$\begin{pmatrix} -1 & 2 \\ 3 & -6 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \text{or} \quad \begin{cases} -x + 2y = 0 \\ 3x - 6y = 0 \end{cases} \quad \text{or} \quad -x + 2y = 0$$

The system has only one independent solution; for example, $x = 2$, $y = 1$. Thus $v_1 = (2, 1)$ is an eigenvector which 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 $M = \begin{pmatrix} 6 & 2 \\ 3 & 1 \end{pmatrix}$ which yields the homogeneous system

$$\begin{cases} 6x + 2y = 0 \\ 3x + \ \ y = 0 \end{cases} \quad \text{or} \quad 3x + y = 0$$

The system has only one independent solution; for example, $x = -1$, $y = 3$. Thus $v_2 = (-1, 3)$ is an eigenvector which spans the eigenspace of $\lambda_2 = -2$.

4. Let P be the matrix whose columns are the above eigenvectors: $P = \begin{pmatrix} 2 & -1 \\ 1 & 3 \end{pmatrix}$. Then $P^{-1} = \begin{pmatrix} \frac{3}{7} & \frac{1}{7} \\ -\frac{1}{7} & \frac{2}{7} \end{pmatrix}$ and $D = P^{-1}AP$ is the diagonal matrix whose diagonal entries are the respective eigenvalues:

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

Accordingly, A has the "diagonal factorization"

$$A = PDP^{-1} = \begin{pmatrix} 2 & -1 \\ 1 & 3 \end{pmatrix}\begin{pmatrix} 5 & 0 \\ 0 & -2 \end{pmatrix}\begin{pmatrix} \frac{3}{7} & \frac{1}{7} \\ -\frac{1}{7} & \frac{2}{7} \end{pmatrix}$$

If $f(t) = t^4 - 4t^3 - 3t^2 + 5$, then we can calculate $f(5) = 55$, $f(-2) = 41$; thus

$$f(A) = Pf(D)P^{-1} = \begin{pmatrix} 2 & -1 \\ 1 & 3 \end{pmatrix}\begin{pmatrix} 55 & 0 \\ 0 & 41 \end{pmatrix}\begin{pmatrix} \frac{3}{7} & \frac{1}{7} \\ -\frac{1}{7} & \frac{2}{7} \end{pmatrix} = \begin{pmatrix} 53 & 4 \\ 6 & 43 \end{pmatrix}$$

Example 8.7. Consider the matrix $B = \begin{pmatrix} 5 & 1 \\ -4 & 1 \end{pmatrix}$. Here $\operatorname{tr} B = 5 + 1 = 6$ and $|B| = 5 + 4 = 9$. Hence $\Delta(t) = t^2 - 6t + 9 = (t - 3)^2$ is the characteristic polynomial of B. Accordingly, $\lambda = 3$ is the only eigenvalue of B.

Subtract $\lambda = 3$ down the diagonal of B to obtain the matrix $M = \begin{pmatrix} 2 & 1 \\ -4 & -2 \end{pmatrix}$ which corresponds to the homogeneous system

$$\begin{cases} 2x + y = 0 \\ -4x - 2y = 0 \end{cases} \qquad \text{or} \qquad 2x + y = 0$$

The system has only one independent solution; for example, $x = 1$, $y = -2$. Thus $v = (1, -2)$ is the only independent eigenvector of the matrix B. Accordingly, B is not diagonalizable since there does not exist a basis consisting of eigenvectors of B.

Example 8.8. Consider the matrix $A = \begin{pmatrix} 2 & -5 \\ 1 & -2 \end{pmatrix}$. Here $\operatorname{tr} A = 2 - 2 = 0$ and $|A| = -4 + 5 = 1$. 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 **R**. Then $\Delta(t)$ has no (real) roots. Thus A has no eigenvalues and no eigenvectors, and so A is not diagonizable.

(*b*) A is a matrix over the complex field **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 **C** for which

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

Therefore, A is diagonalizable (over **C**).

8.6 DIAGONALIZING REAL SYMMETRIC MATRICES

There are many real matrices A which are not diagonalizable. In fact, some such matrices may not have any (real) eigenvalues. However, if A is a real *symmetric* matrix, then these problems do not exist. Namely:

Theorem 8.12: Let A be a real symmetric matrix. Then each root λ of its characteristic polynomial is real.

Theorem 8.13: Let A be a real symmetric matrix. Suppose u and v are nonzero eigenvectors of A belonging to distinct eigenvalues λ_1 and λ_2. Then u and v are orthogonal, i.e., $\langle u, v \rangle = 0$.

The above two theorems gives us the following fundamental result:

Theorem 8.14: Let A be a real symmetric matrix. Then there exists an orthogonal matrix P such that $D = P^{-1}AP$ is diagonal.

We can choose the columns of the above matrix P to be normalized orthogonal eigenvectors of A; then the diagonal entries of D are the corresponding eigenvalues.

Example 8.9. Let $A = \begin{pmatrix} 2 & -2 \\ -2 & 5 \end{pmatrix}$. We find an orthogonal matrix P such that $P^{-1}AP$ is diagonal. Here $\operatorname{tr} A = 2 + 5 = 7$ and $|A| = 10 - 4 = 6$. Hence $\Delta(t) = t^2 - 7t + 6 = (t - 6)(t - 1)$ is the characteristic polynomial of A. The eigenvalues of A are 6 and 1. Subtract $\lambda = 6$ down the diagonal of A to obtain the corresponding homogeneous system of linear equations

$$-4x - 2y = 0 \qquad -2x - y = 0$$

A nonzero solution is $v_1 = (1, -2)$. Next subtract $\lambda = 1$ down the diagonal of A to find the corresponding homogeneous system

$$+x - 2y = 0 \qquad -2x + 4y = 0$$

A nonzero solution is $(2, 1)$. As expected from Theorem 8.13, v_1 and v_2 are orthogonal. Normalize v_1 and v_2 to obtain the orthonormal vectors

$$u_1 = (1/\sqrt{5}, -2/\sqrt{5}) \qquad u_2 = (2/\sqrt{5}, 1/\sqrt{5})$$

Finally let P be the matrix whose columns are u_1 and u_2, respectively. Then

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

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

Application to Quadratic Forms

Recall (Section 4.12) that a real quadratic form $q(x_1, x_2, \ldots, x_n)$ can be expressed in the matrix form

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

where $X = (x_1, \ldots, x_n)^T$ and A is a real symmetric matrix, and recall that under a change of variables $X = PY$, where $Y = (y_1, \ldots, y_n)$ and P is a nonsingular matrix, the quadratic form has the form

$$q(Y) = Y^T BY$$

where $B = P^T AP$. (Thus B is congruent to A.)

Now if P is an orthogonal matrix, then $P^T = P^{-1}$. In such a case, $B = P^T AP = P^{-1}AP$ and so B is orthogonally similar to A. Accordingly, the above method for diagonalizing a real symmetric matrix A can be used to diagonalize a quadratic form q under an orthogonal change of coordinates, as follows.

Orthogonal Diagonalization Algorithm 8.6:

The input is a quadratic form $q(X)$.

Step 1. Find the symmetric matrix A which represents q and find its characteristic polynomial $\Delta(t)$.

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.

Then $X = PY$ is the required orthogonal change of coordinates, and the diagonal entries of $P^T AP$ will be the eigenvalues $\lambda_1, \ldots, \lambda_n$ which correspond to the columns of P.

8.7 MINIMUM POLYNOMIAL

Let A be an n-square matrix over a field K and let $J(A)$ denote the collection of all polynomials $f(t)$ for which $f(A) = 0$. [Note $J(A)$ is not empty since the characteristic polynomial $\Delta_A(t)$ of A belongs to $J(A)$.] Let $m(t)$ be the monic polynomial of minimal degree in $J(A)$. Then $m(t)$ is called the *minimum polynomial* of A. [Such a polynomial $m(t)$ exists and is unique (Problem 8.25).]

Theorem 8.15: The minimum polynomial $m(t)$ of A divides every polynomial which has A as a zero. In particular, $m(t)$ divides the characteristic polynomial $\Delta(t)$ of A.

(The proof is given in Problem 8.32.) There is an even stronger relationship between $m(t)$ and $\Delta(t)$.

Theorem 8.16: The characteristic and minimum polynomials of a matrix A have the same irreducible factors.

This theorem, proved in Problem 8.33(*b*), does not say that $m(t) = \Delta(t)$; only that any irreducible factor of one must divide the other. In particular, since a linear factor is irreducible, $m(t)$ and $\Delta(t)$ have the same linear factors; hence they have the same roots. Thus we have:

Theorem 8.17: A scalar λ is an eigenvalue of the matrix A if and only if λ is a root of the minimum polynomial of A.

Example 8.10. Find the minimum polynomial $m(t)$ of $A = \begin{pmatrix} 2 & 2 & -5 \\ 3 & 7 & -15 \\ 1 & 2 & -4 \end{pmatrix}$.

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

$$\Delta(t) = |tI - A| = \begin{vmatrix} t-2 & -2 & 5 \\ -3 & t-7 & 15 \\ -1 & -2 & t+4 \end{vmatrix} = t^3 - 5t^2 + 7t - 3 = (t-1)^2(t-3)$$

Alternatively, $\Delta(t) = t^3 - (\text{tr } A)t^2 + (A_{11} + A_{22} + A_{33})t - |A| = t^3 - 5t^2 + 7t - 3 = (t-1)^2(t-3)$ (where A_{ii} is the cofactor of a_{ii} in A).

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

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

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

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

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

Example 8.11. Consider the following n-square matrix where $a \neq 0$:

$$M = \begin{pmatrix} \lambda & a & 0 & \ldots & 0 & 0 \\ 0 & \lambda & a & \ldots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \ldots & \lambda & a \\ 0 & 0 & 0 & \ldots & 0 & \lambda \end{pmatrix}$$

Note that M has λ's on the diagonal, a's on the superdiagonal, and 0s elsewhere. This matrix, especially when $a = 1$, is important in linear algebra. One can show that

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

is both the characteristic and minimum polynomial of M.

Example 8.12. Consider an arbitrary monic polynomial $f(t) = t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0$. Let A be the n-square matrix with 1s on the subdiagonal, the negatives of the coefficients in the last column and 0s elsewhere as follows:

$$A = \begin{pmatrix} 0 & 0 & \ldots & 0 & -a_0 \\ 1 & 0 & \ldots & 0 & -a_1 \\ 0 & 1 & \ldots & 0 & -a_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \ldots & 1 & -a_{n-1} \end{pmatrix}$$

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

Minimum Polynomial and Block Diagonal Matrices

The following theorem, proved in Problem 8.34, applies.

Theorem 8.18: Suppose M is a block diagonal matrix with diagonal blocks $A_1, A_2, \ldots, A_r$. Then the minimum polynomial of M is equal to the least common multiple (LCM) of the minimum polynomials of the diagonal blocks A_i.

Remark: We emphasize that this theorem applies to block diagonal matrices, whereas the analogous Theorem 8.5 on characteristic polynomials applies to block triangular matrices.

Example 8.13. Find the characteristic polynomial $\Delta(t)$ and the minimum polynomial $m(t)$ of the matrix

$$A = \begin{pmatrix} 2 & 5 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 4 & 2 & 0 \\ 0 & 0 & 3 & 5 & 0 \\ 0 & 0 & 0 & 0 & 7 \end{pmatrix}$$

Note A is a block diagonal matrix with diagonal blocks

$$A_1 = \begin{pmatrix} 2 & 5 \\ 0 & 2 \end{pmatrix} \qquad A_2 = \begin{pmatrix} 4 & 2 \\ 3 & 5 \end{pmatrix} \qquad A_3 = (7)$$

Then $\Delta(t)$ is the product of the characteristic polynomials $\Delta_1(t)$, $\Delta_2(t)$, and $\Delta_3(t)$ of A_1, A_2, and A_3, respectively. Since A_1 and A_3 are triangular, $\Delta_1(t) = (t-2)^2$ and $\Delta_3(t) = (t-7)$. Also,

$$\Delta_2(t) = t^2 - (\text{tr } A_2)t + |A_2| = t^2 - 9t + 14 = (t-2)(t-7)$$

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

The minimum polynomials $m_1(t)$, $m_2(t)$, and $m_3(t)$ of the diagonal blocks A_1, A_2, and 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 IN MATRICES, CHARACTERISTIC POLYNOMIAL

8.1. Let $A = \begin{pmatrix} 1 & -2 \\ 4 & 5 \end{pmatrix}$. Find $f(A)$ where: (*a*) $f(t) = t^2 - 3t + 7$, and (*b*) $f(t) = t^2 - 6t + 13$.

(*a*) $$f(A) = A^2 - 3A + 7I = \begin{pmatrix} 1 & -2 \\ 4 & 5 \end{pmatrix}^2 - 3\begin{pmatrix} 1 & -2 \\ 4 & 5 \end{pmatrix} + 7\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$
$$= \begin{pmatrix} -7 & -12 \\ 24 & 17 \end{pmatrix} + \begin{pmatrix} -3 & 6 \\ -12 & -15 \end{pmatrix} + \begin{pmatrix} 7 & 0 \\ 0 & 7 \end{pmatrix} = \begin{pmatrix} -3 & -6 \\ 12 & 9 \end{pmatrix}$$

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

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

8.2. Find the characteristic polynomial $\Delta(t)$ of the matrix $A = \begin{pmatrix} 2 & -3 \\ 5 & 1 \end{pmatrix}$.

Form the characteristic matrix $tI - A$:

$$tI - A = \begin{pmatrix} t & 0 \\ 0 & t \end{pmatrix} + \begin{pmatrix} -2 & 3 \\ -5 & -1 \end{pmatrix} = \begin{pmatrix} t-2 & 3 \\ -5 & t-1 \end{pmatrix}$$

The characteristic polynomial $\Delta(t)$ of A is its determinant:

$$\Delta(t) = |tI - A| = \begin{vmatrix} t-2 & 3 \\ -5 & t-1 \end{vmatrix} = (t-2)(t-1) + 15 = t^2 - 3t + 17$$

Alternatively, $\operatorname{tr} A = 2 + 1 = 3$ and $|A| = 2 + 15 = 17$; hence $\Delta(t) = t^2 - 3t + 17$.

8.3. Find the characteristic polynomial $\Delta(t)$ of the matrix $A = \begin{pmatrix} 1 & 6 & -2 \\ -3 & 2 & 0 \\ 0 & 3 & -4 \end{pmatrix}$.

$$\Delta(t) = |tI - A| = \begin{vmatrix} t-1 & -6 & 2 \\ 3 & t-2 & 0 \\ 0 & -3 & t+4 \end{vmatrix} = (t-1)(t-2)(t+4) - 18 + 18(t+4) = t^3 + t^2 - 8t + 62$$

Alternatively, $\operatorname{tr} A = 1 + 2 - 4 = -1$, $A_{11} = \begin{vmatrix} 2 & 0 \\ 3 & -4 \end{vmatrix} = -8$, $A_{22} = \begin{vmatrix} 1 & -2 \\ 0 & -4 \end{vmatrix} = -4$, $A_{33} = \begin{vmatrix} 1 & 6 \\ -3 & 2 \end{vmatrix} = 2 + 18 = 20$, $A_{11} + A_{22} + A_{33} = -8 - 4 + 20 = 8$, and $|A| = -8 + 18 - 72 = -62$. Thus

$$\Delta(t) = t^3 - (\operatorname{tr} A)t^2 + (A_{11} + A_{22} + A_{33})t - |A| = t^3 + t^2 - 8t + 62$$

8.4. Find the characteristic polynomials of the following matrices:

$$(a)\quad R = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 2 & 8 & -6 \\ 0 & 0 & 3 & -5 \\ 0 & 0 & 0 & 4 \end{pmatrix}, \qquad (b)\quad S = \begin{pmatrix} 2 & 5 & 7 & -9 \\ 1 & 4 & -6 & 4 \\ 0 & 0 & 6 & -5 \\ 0 & 0 & 2 & 3 \end{pmatrix}$$

(*a*) Since R is triangular, $\Delta(t) = (t-1)(t-2)(t-3)(t-4)$.

(*b*) Note S is block triangular with diagonal blocks $A_1 = \begin{pmatrix} 2 & 5 \\ 1 & 4 \end{pmatrix}$ and $A_2 = \begin{pmatrix} 6 & -5 \\ 2 & 3 \end{pmatrix}$. Thus

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

EIGENVALUES AND EIGENVECTORS

8.5. Let $A = \begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix}$. Find: (*a*) all eigenvalues of A and the corresponding eigenspaces, (*b*) an invertible matrix P such that $D = P^{-1}AP$ is diagonal, and (*c*) A^5 and $f(A)$ where $f(t) = t^4 - 3t^3 - 7t^2 + 6t - 15$.

(*a*) Form the characteristic matrix $tI - A$ of A:

$$tI - A = \begin{pmatrix} t & 0 \\ 0 & t \end{pmatrix} - \begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} t-1 & -4 \\ -2 & t-3 \end{pmatrix} \tag{1}$$

The characteristic polynomial $\Delta(t)$ of A is its determinant:

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

Alternatively, $\operatorname{tr} A = 1 + 3 = 4$ and $|A| = 3 - 8 = -5$, so $\Delta(t) = t^2 - 4t - 5$. The roots $\lambda_1 = 5$ and $\lambda_2 = -1$ of the characteristic polynomial $\Delta(t)$ are the eigenvalues of A.

We obtain the eigenvectors of A belonging to the eigenvalue $\lambda_1 = 5$. Substitute $t = 5$ in the characteristic matrix (*I*) to obtain the matrix $M = \begin{pmatrix} 4 & -4 \\ -2 & 2 \end{pmatrix}$. The eigenvectors belonging to $\lambda_1 = 5$ form the solution of the homogeneous system $MX = 0$, that is,

$$\begin{pmatrix} 4 & -4 \\ -2 & 2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \text{or} \quad \begin{cases} 4x - 4y = 0 \\ -2x + 2y = 0 \end{cases} \quad \text{or} \quad x - y = 0$$

The system has only one independent solution; for example, $x = 1$, $y = 1$. Thus $v_1 = (1, 1)$ is an eigenvector which spans the eigenspace of $\lambda_1 = 5$.

We obtain the eigenvectors of A belonging to the eigenvalue $\lambda_2 = -1$. Substitute $t = -1$ into $tI - A$ to obtain $M = \begin{pmatrix} -2 & -4 \\ -2 & -4 \end{pmatrix}$ which yields the homogeneous system

$$\begin{cases} -2x - 4y = 0 \\ -2x - 4y = 0 \end{cases} \quad \text{or} \quad x + 2y = 0$$

The system has only one independent solution; for example, $x = 2$, $y = -1$. Thus $v_2 = (2, -1)$ is an eigenvector which spans the eigenspace of $\lambda_2 = -1$.

(*b*) Let P be the matrix whose columns are the above eigenvectors: $P = \begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix}$. Then $D = P^{-1}AP$ is the diagonal matrix whose diagonal entries are the respective eigenvalues:

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

[*Remark*: Here P is the change-of-basic matrix from the usual basis E of $\mathbf{R}^2$ to the basis $S = \{v_1, v_2\}$. Hence D is the matrix representation of the function determined by A in this new basis.]

(*c*) Use the diagonal factorization of A,

$$A = PDP^{-1} = \begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 5 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & -\frac{1}{3} \end{pmatrix}$$

and $5^5 = 3125$ and $(-1)^5 = -1$ to obtain:

$$A^5 = PD^5P^{-1} = \begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 3125 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & -\frac{1}{3} \end{pmatrix} = \begin{pmatrix} 1041 & 2084 \\ 1042 & 2083 \end{pmatrix}$$

Also, since $f(5) = 90$ and $f(-1) = -24$,

$$f(A) = Pf(D)P^{-1} = \begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 90 & 0 \\ 0 & -24 \end{pmatrix}\begin{pmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & -\frac{1}{3} \end{pmatrix} = \begin{pmatrix} 14 & 76 \\ 38 & 52 \end{pmatrix}$$

8.6. Find all eigenvalues and a maximal set S of linearly independent eigenvectors for the following matrices:

$$(a)\quad A = \begin{pmatrix} 5 & 6 \\ 3 & -2 \end{pmatrix} \qquad (b)\quad C = \begin{pmatrix} 5 & -1 \\ 1 & 3 \end{pmatrix}$$

Which of the matrices can be diagonalized? If so, find the required nonsingular matrix P.

(*a*) Find the characteristic polynomial $\Delta(t) = t^2 - 3t - 28 = (t-7)(t+4)$. Thus the eigenvalues of A are $\lambda_1 = 7$ and $\lambda_2 = -4$.

(i) Subtract $\lambda_1 = 7$ down the diagonal of A to obtain $M = \begin{pmatrix} -2 & 6 \\ 3 & -9 \end{pmatrix}$ which corresponds to the system

$$\begin{cases} -2x + 6y = 0 \\ 3x - 9y = 0 \end{cases} \quad \text{or} \quad x - 3y = 0$$

Here $v_1 = (3, 1)$ is a nonzero solution (spanning the solution space) and so v_1 is the eigenvector of $\lambda_1 = 7$.

(ii) Subtract $\lambda_2 = -4$ (or add 4) down the diagonal of A to obtain $M = \begin{pmatrix} 9 & 6 \\ 3 & 2 \end{pmatrix}$ which corresponds to the system $3x + 2y = 0$. Here $v_2 = (2, -3)$ is a solution and hence an eigenvector of $\lambda_2 = -4$.

Then $S = \{v_1 = (3, 1), v_2 = (2, -3)\}$ is a maximal set of linearly independent eigenvectors of A. Since S is a basis for $\mathbf{R}^2$, A is diagonalizable. Let P be the matrix whose columns are v_1 and v_2. Then

$$P = \begin{pmatrix} 3 & 2 \\ 1 & -3 \end{pmatrix} \quad \text{and} \quad P^{-1}AP = \begin{pmatrix} 7 & 0 \\ 0 & -4 \end{pmatrix}$$

(b) Find $\Delta(t) = t^2 - 8t + 16 = (t - 4)^2$. Thus $\lambda = 4$ is the only eigenvalue. Subtract $\lambda = 4$ down the diagonal of C to obtain $M = \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}$ which corresponds to the homogeneous system $x + y = 0$. Here $v = (1, 1)$ is a nonzero solution of the system and hence v is an eigenvector of C belonging to $\lambda = 4$. Since there are no other eigenvalues, the singleton set $S = \{v = (1, 1)\}$ is a maximal set of linearly independent eigenvectors. Furthermore, C is not diagonalizable since the number of linearly independent eigenvectors is not equal to the dimension of the vector space $\mathbf{R}^2$. In particular, no such nonsingular matrix P exists.

8.7. Let $A = \begin{pmatrix} 2 & 2 \\ 1 & 3 \end{pmatrix}$. Find: (a) all eigenvalues of A and the corresponding eigenvectors; (b) an invertible matrix P such that $D = P^{-1}AP$ is diagonal; (c) A^6; and (d) a "positive square root" of A, i.e., a matrix B, having nonnegative eigenvalues, such that $B^2 = A$.

(a) Here $\Delta(t) = t^2 - \text{tr}\, A + |A| = t^2 - 5t + 4 = (t - 1)(t - 4)$. Hence $\lambda_1 = 1$ and $\lambda_2 = 4$ are eigenvalues of A. We find corresponding eigenvectors:

(i) Subtract $\lambda_1 = 1$ down the diagonal of A to obtain $M = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}$ which corresponds to the homogeneous system $x + 2y = 0$. Here $v_1 = (2, -1)$ is a nonzero solution of the system and so an eigenvector of A belonging to $\lambda_1 = 1$.

(ii) Subtract $\lambda_2 = 4$ down the diagonal of A to obtain $M = \begin{pmatrix} -2 & 2 \\ 1 & -1 \end{pmatrix}$ which corresponds to the homogeneous system $x - y = 0$. Here $v_2 = (1, 1)$ is a nonzero solution and so an eigenvector of A belonging to $\lambda_2 = 4$.

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

$$P = \begin{pmatrix} 2 & 1 \\ -1 & 1 \end{pmatrix} \quad \text{and} \quad D = P^{-1}AP = \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix}$$

(c) Use the diagonal factorization of A,

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

to obtain

$$A^6 = PD^6P^{-1} = \begin{pmatrix} 2 & 1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 4096 \end{pmatrix}\begin{pmatrix} \frac{1}{3} & -\frac{1}{3} \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix} = \begin{pmatrix} 1366 & 2730 \\ 1365 & 2731 \end{pmatrix}$$

(*d*) Here $\begin{pmatrix} \pm 1 & 0 \\ 0 & \pm 2 \end{pmatrix}$ are square roots of D. Hence

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

is the positive square root of A.

8.8. Suppose $A = \begin{pmatrix} 4 & 1 & -1 \\ 2 & 5 & -2 \\ 1 & 1 & 2 \end{pmatrix}$. Find: (*a*) the characteristic polynomial $\Delta(t)$ of A, (*b*) the eigenvalues of A, and (*c*) a maximal set of linearly independent eigenvectors of A. (*d*) Is A diagonalizable? If yes, find P such that $P^{-1}AP$ is diagonal.

(*a*) We have

$$\Delta(t) = |tI - A| = \begin{vmatrix} t-4 & -1 & 1 \\ -2 & t-5 & 2 \\ -1 & -1 & t-2 \end{vmatrix} = t^3 - 11t^2 + 39t - 45$$

Alternatively, $\Delta(t) = t^3 - (\text{tr } A)t^2 + (A_{11} + A_{22} + A_{33})t - |A| = t^3 - 11t^2 + 39t - 45$. (Here A_{ii} is the cofactor of a_{ii} in the matrix A.)

(*b*) Assuming $\Delta(t)$ has a rational root, it must be among $\pm 1, \pm 3, \pm 5, \pm 9, \pm 15, \pm 45$. Testing by synthetic division, we get

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

Thus $t = 3$ is a root of $\Delta(t)$ and $t - 3$ is a factor, giving

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

Accordingly, $\lambda_1 = 3$ and $\lambda_2 = 5$ are the eigenvalues of A.

(*c*) Find independent eigenvectors for each eigenvalue of A.

(i) Subtract $\lambda_1 = 3$ down the diagonal of A to obtain the matrix $M = \begin{pmatrix} 1 & 1 & -1 \\ 2 & 2 & -2 \\ 1 & 1 & -1 \end{pmatrix}$ which corresponds to the homogeneous system $x + y - z = 0$. Here $u = (1, -1, 0)$ and $v = (1, 0, 1)$ are two independent solutions.

(ii) Subtract $\lambda_2 = 5$ down the diagonal of A to obtain $M = \begin{pmatrix} -1 & 1 & -1 \\ 2 & 0 & -2 \\ 1 & 1 & -3 \end{pmatrix}$ which corresponds to the homogeneous system

$$\begin{cases} -x + y - z = 0 \\ 2x - 2z = 0 \\ x + y - 3z = 0 \end{cases} \quad \text{or} \quad \begin{cases} x - z = 0 \\ y - 2z = 0 \end{cases}$$

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

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

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

(*d*) A is diagonalizable since it has three linearly independent eigenvectors. Let P be the matrix with column u, v, w. Then

$$P = \begin{pmatrix} 1 & 1 & 1 \\ -1 & 0 & 2 \\ 0 & 1 & 1 \end{pmatrix} \quad \text{and} \quad P^{-1}AP = \begin{pmatrix} 3 & & \\ & 3 & \\ & & 5 \end{pmatrix}$$

8.9. Suppose $B = \begin{pmatrix} 3 & -1 & 1 \\ 7 & -5 & 1 \\ 6 & -6 & 2 \end{pmatrix}$ Find: (*a*) the characteristic polynomial $\Delta(t)$ and eigenvalues of B; and (*b*) a maximal set S of linearly independent eigenvectors of B. (*c*) Is B diagonalizable? If yes, find P such tht $P^{-1}BP$ is diagonal.

(*a*) We have:

$tr(B) = 3 - 5 + 2 = 0$, $B_{11} = -10 + 6 = -4$, $B_{22} = 6 - 6 = 0$, $B_{33} = -15 + 7 = -8$,
$|B| = -30 - 6 - 42 + 30 + 18 + 14 = -16$

Therefore, $\Delta(t) = t^3 - 12t + 16 = (t-2)^2(t+4)$ and so $\lambda = 2$ and $\lambda = 4$ are the eigvenvalues of B.

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

(i) Subtract $\lambda = 2$ down the diagonal of B to obtain the homogeneous system

$$\begin{pmatrix} 1 & -1 & 1 \\ 7 & -7 & 1 \\ 6 & -6 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad \text{or} \quad \begin{cases} x - y + z = 0 \\ 7x - 7y + z = 0 \\ 6x - 6y = 0 \end{cases} \quad \text{or} \quad \begin{cases} x - y + z = 0 \\ x - y = 0 \end{cases}$$

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

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

$$\begin{pmatrix} 7 & -1 & 1 \\ 7 & -1 & 1 \\ 6 & -6 & 6 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad \text{or} \quad \begin{cases} 7x - y + z = 0 \\ 7x - y + z = 0 \\ 6x - 6y + 6z = 0 \end{cases} \quad \text{or} \quad \begin{cases} 7x - y + z = 0 \\ x = 0 \end{cases}$$

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

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

(*c*) Since B has at most two independent eigenvectors, B is not similar to a diagonal matrix, i.e., B is not diagonalizable.

8.10. Find the algebraic and geometric multiplicities of the eigenvalue $\lambda = 2$ for matrix B in Problem 8.9.

The algebraic multiplicity of $\lambda = 2$ is two since $t - 2$ appears with exponent 2 in $\Delta(t)$. However, the geometric multiplicity of $\lambda = 2$ is one since $\dim E_\lambda = 1$.

8.11. Let $A = \begin{pmatrix} 1 & -1 \\ 2 & -1 \end{pmatrix}$. Find all eigenvalues and corresponding eigenvectors of A assuming A is a real matrix. Is A diagonalizable? If yes, find P such that $P^{-1}AP$ is diagonal.

The characteristic polynomial of A is $\Delta(t) = t^2 + 1$ which has no root in $\mathbf{R}$. Thus A, viewed as a real matrix, has no eigenvalues and no eigenvectors, and hence A is not diagonalizable over $\mathbf{R}$.

8.12. Repeat Problem 8.11 assuming now that A is a matrix over the complex field $\mathbf{C}$.

The characteristic polynomial of A is still $\Delta(t) = t^2 + 1$. (It does not depend on the field K.) Over $\mathbf{C}$, $\Delta(t)$ does factor; specifically, $\Delta(t) = t^2 + 1 = (t - i)(t + i)$. Thus $\lambda_1 = i$ and $\lambda_2 = -i$ are eigenvalues of A.

(i) Substitute $t = i$ in $tI - A$ to obtain the homogeneous system

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

The system has only one independent solution, e.g., $x = 1$, $y = 1 - i$. Thus $v_1 = (1, 1 - i)$ is an eigenvector which spans the eigenspace of $\lambda_1 = i$.

(ii) Substitute $t = -i$ into $tI - A$ to obtain the homogeneous system

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

The system has only one independent solution, e.g., $x = 1$, $y = 1 + i$. Thus $v_2 = (1, 1 + i)$ is an eigenvector of A which spans the eigenspace of $\lambda_2 = -i$.

As a complex matrix, A is diagonalizable. Let P be the matrix whose columns are v_1 and v_2. Then

$$P = \begin{pmatrix} 1 & 1 \\ 1-i & 1+i \end{pmatrix} \quad \text{and} \quad P^{-1}AP = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$$

8.13. Let $B = \begin{pmatrix} 2 & 4 \\ 3 & 1 \end{pmatrix}$. Find: (*a*) all eigenvalues of B and the corresponding eigenvectors; (*b*) an invertible matrix P such that $D = P^{-1}BP$ is diagonal; and (*c*) B^6.

(*a*) Here $\Delta(t) = t^2 - \operatorname{tr} B + |B| = t^2 - 3t - 10 = (t - 5)(t + 2)$. Thus $\lambda_1 = 5$ and $\lambda_2 = -2$ are the eigenvalues of B.

(i) Subtract $\lambda_1 = 5$ down the diagonal of B to obtain $M = \begin{pmatrix} -3 & 4 \\ 3 & -4 \end{pmatrix}$ which corresponds to the homogeneous system $3x - 4y = 0$. Here $v_1 = (4, 3)$ is a nonzero solution.

(ii) Subtract $\lambda_2 = -2$ (or add 2) down the diagonal of B to obtain $M = \begin{pmatrix} 4 & 4 \\ 3 & 3 \end{pmatrix}$ which corresponds to the system $x + y = 0$ which has a nonzero solution $v_2 = (1, -1)$.

(Since B has two independent eigenvectors, B is diagonalizable.)

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

$$P = \begin{pmatrix} 4 & 1 \\ 3 & -1 \end{pmatrix} \quad \text{and} \quad D = P^{-1}BP = \begin{pmatrix} 5 & 0 \\ 0 & -2 \end{pmatrix}$$

(*c*) Use the diagonal factorization of B,

$$B = PDP^{-1} = \begin{pmatrix} 4 & 1 \\ 3 & -1 \end{pmatrix}\begin{pmatrix} 5 & 0 \\ 0 & -2 \end{pmatrix}\begin{pmatrix} \frac{1}{7} & \frac{1}{7} \\ \frac{3}{7} & -\frac{4}{7} \end{pmatrix}$$

to obtain ($5^6 = 15\,625$, $(-2)^6 = 64$):

$$B^6 = PD^6P^{-1} = \begin{pmatrix} 4 & 1 \\ 3 & -1 \end{pmatrix}\begin{pmatrix} 15\,625 & 0 \\ 0 & 64 \end{pmatrix}\begin{pmatrix} \frac{1}{7} & \frac{1}{7} \\ \frac{3}{7} & -\frac{4}{7} \end{pmatrix} = \begin{pmatrix} 8956 & 8892 \\ 6669 & 6733 \end{pmatrix}$$

8.14. Determine whether or not A is diagonalizable where $A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 2 & 3 \\ 0 & 0 & 3 \end{pmatrix}$.

Since A is triangular, the eigenvalues of A are the diagonal elements 1, 2, and 3. Since they are distinct, A has three independent eigenvectors and thus A is similar to a diagonal matrix (Theorem 8.11). (We emphasize that here we do not need to compute eigenvectors to tell that A is diagonalizable. We will have to compute eigenvectors if we want to find P such that $P^{-1}AP$ is diagonal.)

8.15. Suppose A and B are n-square matrices.

(a) Show that 0 is an eigenvalue of A if and only if A is singular.

(b) Show that AB and BA have the same eigenvalues.

(c) Suppose A is nonsingular (invertible) and λ is an eigenvalue of A. Show that λ^{-1} is an eigenvalue of A^{-1}.

(d) Show that A and its transpose A^T have the same characteristic polynomial.

(a) We have that 0 is an eigenvalue of A if and only if there exists a nonzero vector v such that $A(v) = 0v = 0$; i.e., if and only if A is singular.

(b) By part (a) and the fact that the product of nonsingular matrices is nonsingular, the following statements are equivalent: (i) 0 is an eigenvalue of AB, (ii) AB is singular, (iii) A or B is singular, (iv) BA is singular, (v) 0 is an eigenvalue of BA.

Now suppose λ is a nonzero eigenvalue of AB. Then there exists a nonzero vector v such that $ABv = \lambda v$. Set $w = Bv$. Since $\lambda \neq 0$ and $v \neq 0$,

$$Aw = ABv = \lambda v \neq 0 \qquad \text{and so} \qquad w \neq 0$$

But w is an eigenvector of BA belonging to the eigenvalue λ since

$$BAw = BABv = B\lambda v = \lambda Bv = \lambda w$$

Hence λ is an eigenvalue of BA. Similarly, any nonzero eigenvalue of BA is also an eigenvalue of AB.

Thus AB and BA have the same eigenvalues.

(c) By part (a) $\lambda \neq 0$. By definition of an eigenvalue, there exists a nonzero vector v for which $A(v) = \lambda v$. Applying A^{-1} to both sides, we obtain $v = A^{-1}(\lambda v) = \lambda A^{-1}(v)$. Hence $A^{-1}(v) = \lambda^{-1}v$; that is, λ^{-1} is an eigenvalue of A^{-1}.

(d) Since a matrix and its transpose have the same determinant, $|tI - A| = |(tI - A)^T| = |tI - A^T|$. Thus A and A^T have the same characteristic polynomial.

8.16. Let λ be an eigenvalue of an n-square matrix A over K. Let E_λ be the eigenspace of λ, i.e., the set of all eigenvectors of A belonging to λ. Show that E_λ is a subspace of K^n, that is, show that: (a) if $v \in E_\lambda$, then $kv \in E_\lambda$ for any scalar $k \in K$; and (b) if $u, v \in E_\lambda$, then $u + v \in E_\lambda$.

(a) Since $v \in E_\lambda$, we have $A(v) = \lambda v$. Then

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

Thus $kv \in E_\lambda$. [We must allow the zero vector of K^n to serve as the "eigenvector" corresponding to $k = 0$, to make E_λ a *subspace*.]

(b) Since $u, v \in E_\lambda$, we have $A(u) = \lambda v$ and $A(v) = \lambda v$. Then

$$A(u + v) = A(u) + A(v) = \lambda u + \lambda v = \lambda(u + v)$$

Thus $u + v \in E_\lambda$.

DIAGONALIZING REAL SYMMETRIC MATRICES AND REAL QUADRATIC FORMS

8.17. Let $A = \begin{pmatrix} 3 & 2 \\ 2 & 3 \end{pmatrix}$. Find a (real) orthogonal matrix P for which P^TAP is diagonal.

The characteristic polynomial $\Delta(t)$ of A is

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

and thus the eigenvalues of A are 5 and 1.

Subtract $\lambda = 5$ down the diagonal of A to obtain the corresponding homogeneous system of linear equations

$$-2x + 2y = 0 \qquad 2x - 2y = 0$$

A nonzero solution is $v_1 = (1, 1)$. Normalize v_1 to find the unit solution $u_1 = (1/\sqrt{2}, 1/\sqrt{2})$.

Next subtract $\lambda = 1$ down the diagonal of A to obtain the corresponding homogeneous system of linear equations

$$2x + 2y = 0 \qquad 2x + 2y = 0$$

A nonzero solution is $v_2 = (1, -1)$. Normalize v_2 to find the unit solution $u_2 = (1/\sqrt{2}, -1/\sqrt{2})$.

Finally let P be the matrix whose columns are u_1 and u_2, respectively; then

$$P = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix} \quad \text{and} \quad P^TAP = \begin{pmatrix} 5 & 0 \\ 0 & 1 \end{pmatrix}$$

As expected, the diagonal entries of P^TAP are the eigenvalues of A.

8.18. Suppose $C = \begin{pmatrix} 11 & -8 & 4 \\ -8 & -1 & -2 \\ 4 & -2 & -4 \end{pmatrix}$. Find: (*a*) the characteristic polynomial $\Delta(t)$ of C; (*b*) the eigenvalues of C or, in other words, the roots of $\Delta(t)$; (*c*) a maximal set S of nonzero orthogonal eigenvectors of C; and (*d*) an orthogonal matrix P such that $P^{-1}CP$ is diagonal.

(*a*) We have

$$\Delta(t) = t^3 - (\operatorname{tr} C)t^2 + (C_{11} + C_{22} + C_{33})t - |C| = t^3 - 6t^2 - 135t - 400$$

[Here C_{ii} is the cofactor of c_{ii} in $C = (c_{ij})$.]

(*b*) If $\Delta(t)$ has a rational root, it must divide 400. Testing $t = -5$, we get

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

Thus $t + 5$ is a factor of $\Delta(t)$ and

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

Accordingly, the eigenvalues of C are $\lambda = -5$ (with multiplicity two) and $\lambda = 16$ (with multiplicity one).

(*c*) Find an orthogonal basis for each eigenspace.

Subtract $\lambda = -5$ down the diagonal of C 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; i.e., such that

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

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

Subtract $\lambda = 16$ down the diagonal of C 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 8.13, 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 C.

(*d*) Normalize v_1, v_2, v_3 to obtain the orthonormal basis

$$u_1 = (0, 1/\sqrt{5}, 2/\sqrt{5}) \qquad u_2 = (-5/\sqrt{105}, -8/\sqrt{105}, 4/\sqrt{105}) \qquad u_3 = (4/\sqrt{21}, -2/\sqrt{21}, 1/\sqrt{21})$$

Then P is the matrix whose columns are u_1, u_2, u_3. Thus

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

8.19. Let $q(x, y) = 3x^2 - 6xy + 11y^2$. Find an orthogonal change of coordinates which diagonalizes q.

Find the symmetric matrix A representing q and its characteristic polynomial $\Delta(t)$:

$$A = \begin{pmatrix} 3 & -3 \\ -3 & 11 \end{pmatrix} \quad \text{and} \quad \Delta(t) = \begin{vmatrix} t-3 & 3 \\ 3 & t-11 \end{vmatrix} = t^2 - 14t + 24 = (t-2)(t-12)$$

The eigenvalues are 2 and 12; hence a diagonal form of q is

$$q(x', y') = 2x'^2 + 12y'^2$$

The corresponding change of coordinates is obtained by finding a corresponding set of eigenvectors of A.

Subtract $\lambda = 2$ down the diagonal of A to obtain the homogeneous system

$$x - 3y = 0, \qquad -3x + 9y = 0$$

A nonzero solution is $v_1 = (3,1)$. Next subtract $\lambda = 12$ down the diagonal of A to obtain the homogeneous system

$$-9x - 3y = 0, \qquad -3x - y = 0$$

A nonzero solution is $v_2 = (-1, 3)$. Normalize v_1 and v_2 to obtain the orthonormal basis

$$u_1 = (3/\sqrt{10}, 1/\sqrt{10}) \qquad u_2 = (-1/\sqrt{10}, 3/\sqrt{10})$$

The change-of-basis matrix P and the required change of coordinates follow:

$$P = \begin{pmatrix} 3/\sqrt{10} & -1/\sqrt{10} \\ 1/\sqrt{10} & 3/\sqrt{10} \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} x \\ y \end{pmatrix} = P\begin{pmatrix} x' \\ y' \end{pmatrix} \quad \text{or} \quad \begin{cases} x = (3x' - y')/\sqrt{10} \\ y = (x' + 3y')/\sqrt{10} \end{cases}$$

One can also express x' and y' in terms of x and y by using $P^{-1} = P^T$, that is,

$$x' = (3x + y)/\sqrt{10} \qquad y' = (-x + 3y)/\sqrt{10}$$

8.20. Consider the quadratic form $q(x, y, z) = 3x^2 + 2xy + 3y^2 + 2xz + 2yz + 3z^2$. Find:

(*a*) The symmetric matrix A which represents q and its characteristic polynomial $\Delta(t)$,

(*b*) The eigenvalues of A or, in other words, the roots of $\Delta(t)$,

(*c*) A maximal set S of nonzero orthogonal eigenvectors of A,

(*d*) An orthogonal change of coordinates which diagonalizes q.

(*a*) Recall $A = (a_{ij})$ is the symmetric matrix where a_{ii} is the coefficient of x_i^2 and $a_{ij} = a_{ji}$ is one-half the coefficient of $x_i x_j$. Thus

$$A = \begin{pmatrix} 3 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 3 \end{pmatrix} \quad \text{and} \quad \Delta(t) = \begin{vmatrix} t-3 & -1 & -1 \\ -1 & t-3 & -1 \\ -1 & -1 & t-3 \end{vmatrix} = t^3 - 9t^2 + 24t - 20$$

(*b*) If $\Delta(t)$ has a rational root, it must divide the constant 20, or, in other words, it must be among $\pm 1, \pm 2, \pm 4, \pm 5, \pm 10, \pm 20$. Testing $t = 2$, we get

$$\begin{array}{r|l} 2 & 1 - 9 + 24 - 20 \\ & \quad\;\; 2 - 14 + 20 \\ \hline & 1 - 7 + 10 + 0 \end{array}$$

Thus $t - 2$ is a factor of $\Delta(t)$, and we find

$$\Delta(t) = (t - 2)(t^2 - 7t + 10) = (t - 2)^2(t - 5)$$

Hence the eigenvalues of A are 2 (with multiplicity two) and 5 (with multiplicity one).

(*c*) Find an orthogonal basis for each eigenspace.

Subtract $\lambda = 2$ down the diagonal of A to obtain the corresponding homogeneous system

$$x + y + z = 0 \qquad x + y + z = 0 \qquad x + y + z = 0$$

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

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

For example, $v_2 = (2, -1, -1)$. Thus $v_1 = (0, 1, -1)$, $v_2 = (2, -1, -1)$ form an orthogonal basis for the eigenspace of $\lambda = 2$.

Subtract $\lambda = 5$ down the diagonal of A to obtain the corresponding homogeneous system

$$-2x + y + z = 0 \qquad x - 2y + z = 0 \qquad x + y - 2z = 0$$

This system yields a nonzero solution $v_3 = (1, 1, 1)$. (As expected from Theorem 8.13, 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 A.

(*d*) Normalize v_1, v_2, v_3 to obtain the orthonormal basis

$$u_1 = (0, 1/\sqrt{2}, -1/\sqrt{2}) \qquad u_2 = (2/\sqrt{6}, -1/\sqrt{6}, -1/\sqrt{6}) \qquad u_3 = (1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3})$$

Let P be the matrix whose columns are u_1, u_2, u_3. Then

$$P = \begin{pmatrix} 0 & 2/\sqrt{6} & 1/\sqrt{3} \\ 1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ -1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \end{pmatrix} \qquad \text{and} \qquad P^TAP = \begin{pmatrix} 2 & & \\ & 2 & \\ & & 5 \end{pmatrix}$$

Thus the required orthogonal change of coordinates is

$$x = \frac{2y'}{\sqrt{6}} + \frac{z'}{\sqrt{3}}$$

$$y = \frac{x'}{\sqrt{2}} - \frac{y'}{\sqrt{6}} + \frac{z'}{\sqrt{3}}$$

$$z = -\frac{x'}{\sqrt{2}} - \frac{y'}{\sqrt{6}} + \frac{z'}{\sqrt{3}}$$

Under this change of coordinates, q is transformed into the diagonal form

$$q(x', y', z') = 2x'^2 + 2y'^2 + 5z'^2$$

MINIMUM POLYNOMIAL

8.21. Find the minimum polynomial $m(t)$ of the matrix $A = \begin{pmatrix} 4 & -2 & 2 \\ 6 & -3 & 4 \\ 3 & -2 & 3 \end{pmatrix}$.

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

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

Alternatively, $\Delta(t) = t^3 - (\text{tr } A)t^2 + (A_{11} + A_{22} + A_{33})t - |A| = t^3 - 4t^2 + 5t - 2 = (t-2)(t-1)^2$. (Here A_{ii} is the cofactor of a_{ii} in A.)

The minimum polynomial $m(t)$ must divide $\Delta(t)$. Also, each irreducible factor of $\Delta(t)$, that is, $t-2$ and $t-1$, must also be a factor of $m(t)$. Thus $m(t)$ is exactly only of the following:

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

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

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

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

8.22. Find the minimum polynomial $m(t)$ of the matrix, where $a \neq 0$. $B = \begin{pmatrix} \lambda & a & 0 \\ 0 & \lambda & a \\ 0 & 0 & \lambda \end{pmatrix}$.

The characteristic polynomial of B is $\Delta(t) = (t-\lambda)^3$. [Note $m(t)$ is exactly one of $t-\lambda$, $(t-\lambda)^2$, or $(t-\lambda)^3$.] We find $(B - \lambda I)^2 \neq 0$; thus $m(t) = \Delta(t) = (t-\lambda)^3$.

(*Remark:* This matrix is a special case of Example 8.11 and Problem 8.61.)

8.23. Find the minimum polynomial $m(t)$ of the following matrix: $M' = \begin{pmatrix} 4 & 1 & 0 & 0 & 0 \\ 0 & 4 & 1 & 0 & 0 \\ 0 & 0 & 4 & 0 & 0 \\ 0 & 0 & 0 & 4 & 1 \\ 0 & 0 & 0 & 0 & 4 \end{pmatrix}$.

Here M' is block diagonal with diagonal blocks

$$A' = \begin{pmatrix} 4 & 1 & 0 \\ 0 & 4 & 1 \\ 0 & 0 & 4 \end{pmatrix} \qquad \text{and} \qquad B' = \begin{pmatrix} 4 & 1 \\ 0 & 4 \end{pmatrix}$$

The characteristic and minimum polynomial of A' is $f(t) = (t-4)^3$, and the characteristic and minimum polynomial of B' is $g(t) = (t-4)^2$. Thus $\Delta(t) = f(t)g(t) = (t-4)^5$ is the characteristic polynomial of M', but $m(t) = \text{LCM}\,[f(t), g(t)] = (t-4)^3$ (which is the size of the largest block) is the minimum polynomial of M'.

8.24. Find a matrix A whose minimum polynomial is:

(*a*) $f(t) = t^3 - 8t^2 + 5t + 7$, (*b*) $f(t) = t^4 - 3t^3 - 4t^2 + 5t + 6$

Let A be the companion matrix (see Example 8.12) of $f(t)$. Then

$$(a)\quad A = \begin{pmatrix} 0 & 0 & -7 \\ 1 & 0 & -5 \\ 0 & 1 & 8 \end{pmatrix}, \qquad (b)\quad A = \begin{pmatrix} 0 & 0 & 0 & -6 \\ 1 & 0 & 0 & -5 \\ 0 & 1 & 0 & 4 \\ 0 & 0 & 1 & 3 \end{pmatrix}$$

(*Remark:* The polynomial $f(t)$ is also the characteristic polynomial of A.)

8.25. Show that the minimum polynomial of a matrix A exists and is unique.

By the Cayley-Hamilton Theorem, A is a zero of some nonzero polynomial (see also Problem 8.37). Let n be the lowest degree for which a polynomial $f(t)$ exists such that $f(A) = 0$. Dividing $f(t)$ by its leading coefficient, we obtain a monic polynomial $m(t)$ of degree n which has A as a zero. Suppose $m'(t)$ is another monic polynomial of degree n for which $m'(A) = 0$. Then the difference $m(t) - m'(t)$ is a nonzero polynomial of degree less than n which has A as a zero. This contradicts the original assumption on n; hence $m(t)$ is the unique minimum polynomial.

PROOFS OF THEOREMS

8.26. Prove Theorem 8.1. (i) $(f+g)(A) = f(A) + g(A)$, (ii) $(fg)(A) = f(A)g(A)$, (iii) $(kf)(A) = kf(A)$.

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,

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

(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), $g(A)f(A) = (gf)(A) = (fg)(A) = f(A)g(A)$.

8.27. Prove the Cayley-Hamilton Theorem 8.2. Every matrix is a root of its characteristic polynomial.

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 7.9), $(tI - A)B(t) = |tI - A|I$, 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 parentheses and equating the coefficients of corresponding powers of t,

$$\begin{aligned} B_{n-1} &= I \\ B_{n-2} - AB_{n-1} &= a_{n-1}I \\ B_{n-3} - AB_{n-2} &= a_{n-2}I \\ &\cdots\cdots \\ B_0 - AB_1 &= a_1 I \\ -AB_0 &= a_0 I \end{aligned}$$

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

$$\begin{aligned} A^n B_{n-1} &= A^n \\ A^{n-1}B_{n-2} - A^n B_{n-1} &= a_{n-1}A^{n-1} \\ A^{n-2}B_{n-3} - A^{n-1}B_{n-2} &= a_{n-2}A^{n-2} \\ &\cdots\cdots \\ AB_0 - A^2 B_1 &= a_1 A \\ -AB_0 &= a_0 I \end{aligned}$$

Adding the above matrix equations,

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

or $\Delta(A) = 0$, which is the Cayley-Hamilton Theorem.

8.28. Prove Theorem 8.6.

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 $M = \lambda I - A$ 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 $(\lambda I - A)X = 0$.

8.29. Prove Theorem 8.9.

Suppose A has n linearly independent eigenvectors $v_1, v_2, \ldots, v_n$ with corresponding eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$. Let P be the matrix whose columns are $v_1, \ldots, v_n$. Then P is nonsingular. Also, the columns of AP are $Av_1, \ldots, Av_n$. But $Av_k = \lambda v_k$. Hence the columns of AP are $\lambda_1 v_1, \ldots, \lambda_n v_n$. On the other hand, let $D = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$, that is, the diagonal matrix with diagonal entries λ_k. Then PD is also a matrix with columns $\lambda_k v_k$. Accordingly,

$$AP = PD \qquad \text{and hence} \qquad D = P^{-1}AP$$

as required.

Conversely, suppose there exists a nonsingular matrix P for which

$$P^{-1}AP = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n) = D \qquad \text{and so} \qquad AP = PD$$

Let $v_1, v_2, \ldots, v_n$ be the column vectors of P. Then the columns of AP are Av_k and the columns of PD are $\lambda_k v_k$. Accordingly, since $AP = PD$, we have

$$Av_1 = \lambda_1 v_1, \; Av_2 = \lambda_2 v_2, \; \ldots, \; Av_n = \lambda_n v_n$$

Furthermore, since P is nonsingular, $v_1, v_2, \ldots, v_n$ are nonzero and hence, they are eigenvectors of A belonging to the eigenvalues that are the diagonal elements of D. Moreover, they are linearly independent. Thus the theorem is proved.

8.30. Prove Theorem 8.10.

The proof is by induction on n. If $n = 1$, then v_1 is linearly independent since $v_1 \neq 0$. Assume $n > 1$. Suppose

$$a_1 v_1 + a_2 v_2 + \cdots + a_n v_n = 0 \tag{1}$$

where the a_i are scalars. Multiply (*1*) by A and obtain

$$a_1 Av_1 + a_2 Av_2 + \cdots + a_n Av_n = A0 = 0$$

By hypothesis, $Av_i = \lambda_i v_i$. Thus on substitution we obtain

$$a_1 \lambda_1 v_1 + a_2 \lambda_2 v_2 + \cdots + a_n \lambda_n v_n = 0 \tag{2}$$

On the other hand, multiplying (*1*) by λ_n, we get

$$a_1 \lambda_n v_1 + a_2 \lambda_n v_2 + \cdots + a_n \lambda_n v_n = 0 \tag{3}$$

Subtracting (*3*) from (*2*) yields

$$a_1(\lambda_1 - \lambda_n)v_1 + a_2(\lambda_2 - \lambda_n)v_2 + \cdots + a_{n-1}(\lambda_{n-1} - \lambda_n)v_{n-1} = 0$$

By induction, $v_1, v_2, \ldots, v_{n-1}$ are linearly independent; hence each of the above coefficients is 0. Since the λ_i are distinct, $\lambda_i - \lambda_n \neq 0$ for $i \neq n$. Hence $a_1 = \cdots = a_{n-1} = 0$. Substituting this into (*1*), we get $a_n v_n = 0$, and hence $a_n = 0$. Thus the v_i are linearly independent.

8.31. Prove Theorem 8.11.

By Theorem 8.6, the a_i are eigenvalues of A. Let v_i be corresponding eigenvectors. By Theorem 8.10, the v_i are linearly independent and hence form a basis of K^n. Thus A is diagonalizable by Theorem 8.9.

8.32. Prove Theorem 8.15. The minimum polynomial $m(t)$ of A divides $f(t)$ whenever $f(A) = 0$.

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 $m(t)$ which has A as a zero; this contradicts the definition of the minimum polynomial. Thus $r(t) = 0$ and so $f(t) = m(t)q(t)$, i.e., $m(t)$ divides $f(t)$.

8.33. Let $m(t)$ be the minimum polynomial of an n-square matrix A.

(*a*) Show that the characteristic polynomial of A divides $(m(t))^n$.

(*b*) Prove Theorem 8.16. $m(t)$ and $\Delta(t)$ have the same irreducible factors.

(*a*) Suppose $m(t) = t^r + c_1 t^{r-1} + \cdots + c_{r-1}t + c_r$. Consider the following matrices:

$$\begin{aligned} B_0 &= I \\ B_1 &= A + c_1 I \\ B_2 &= A^2 + c_1 A + c_2 I \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ B_{r-1} &= A^{r-1} + c_1 A^{r-2} + \cdots + c_{r-1} I \end{aligned}$$

Then

$$\begin{aligned} B_0 &= I \\ B_1 - AB_0 &= c_1 I \\ B_2 - AB_1 &= c_2 I \\ &\cdots\cdots\cdots \\ B_{r-1} - AB_{r-2} &= c_{r-1} I \end{aligned}$$

Also,

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

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

Then

$$\begin{aligned}(tI - A) \cdot B(t) &= (t^rB_0 + t^{r-1}B_1 + \cdots + tB_{r-1}) - (t^{r-1}AB_0 + t^{r-2}AB_1 + \cdots + AB_{r-1}) \\ &= t^rB_0 + t^{r-1}(B_1 - AB_0) + t^{r-2}(B_2 - AB_1) + \cdots + t(B_{r-1} - AB_{r-2}) - AB_{r-1} \\ &= t^rI + c_1t^{r-1}I + c_2t^{r-2}I + \cdots + c_{r-1}tI + c_rI \\ &= m(t)I\end{aligned}$$

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

(*b*) Suppose $f(t)$ is an irreducible polynomial. If $f(t)$ divides $m(t)$ then, since $m(t)$ divides $\Delta(t)$, $f(t)$ divides $\Delta(t)$. On the other hand, if $f(t)$ divides $\Delta(t)$ then, by part (*a*), $f(t)$ 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.

8.34. Prove Theorem 8.18.

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

Since $m(t)$ is the minimum polynomial of M, $m(M) = \begin{pmatrix} m(A) & 0 \\ 0 & m(B) \end{pmatrix} = 0$ and hence $m(A) = 0$ and $m(B) = 0$. Since $g(t)$ is the minimum 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 $f(M) = \begin{pmatrix} f(A) & 0 \\ 0 & f(B) \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = 0$. But $m(t)$ is the minimum polynomial of M; hence $m(t)$ divides $f(t)$. Thus $m(t)$ is the least common multiple of $g(t)$ and $h(t)$.

8.35. Suppose A is a real symmetric matrix viewed as a matrix over **C**.

(*a*) Prove that $\langle Au, v\rangle = \langle u, Av\rangle$ for the inner product in $\mathbf{C}^n$.

(*b*) Prove Theorems 8.12 and 8.13 for the matrix A.

(*a*) We use the fact that the inner product in $\mathbf{C}^n$ is defined by $\langle u, v\rangle = u^T\bar{v}$. Since A is real symmetric, $A = A^T = \bar{A}$. Thus

$$\langle Au, v\rangle = (Au)^T\bar{v} = u^TA^T\bar{v} = u^T\bar{A}\bar{v} = u^T\overline{Av} = \langle u, Av\rangle$$

(*b*) We use the fact that in $\mathbf{C}^n$, $\langle ku, v\rangle = k\langle u, v\rangle$ but $\langle u, kv\rangle = \bar{k}\langle u, v\rangle$.

(1) There exists $v \neq 0$ such that $Av = \lambda v$. Then

$$\lambda\langle v, v\rangle = \langle \lambda v, v\rangle = \langle Av, v\rangle = \langle v, Av\rangle = \langle v, \lambda v\rangle = \bar{\lambda}\langle v, v\rangle$$

But $\langle v, v\rangle \neq 0$ since $v \neq 0$. Thus $\lambda = \bar{\lambda}$ and so λ is real.

(2) Here $Au = \lambda_1 u$ and $Av = \lambda_2 v$ and, by (1), λ_2 is real. Then

$$\lambda_1\langle u, v\rangle = \langle \lambda_1 u, v\rangle = \langle Au, v\rangle = \langle u, Av\rangle = \langle u, \lambda_2 v\rangle = \bar{\lambda}_2\langle u, v\rangle = \lambda_2\langle u, v\rangle$$

Since $\lambda_1 \neq \lambda_2$, we have $\langle u, v\rangle = 0$.

MISCELLANEOUS PROBLEMS

8.36. Suppose A be a 2×2 symmetric matrix with eigenvalues 1 and 9 and suppose $u = (1, 3)^T$ is an eigenvector belonging to the eigenvalue 1. Find: (*a*) an eigenvector v belonging to the eigenvalue 9, (*b*) the matrix A, and (*c*) a square root of A, i.e., a matrix B such that $B^2 = A$.

(*a*) Since A is symmetric, v must be orthogonal to u. Set $v = (-3, 1)^T$.

(*b*) Let P be the matrix whose columns are the eigenvectors u and v. Then, by the diagonal factorization of A, we have

$$A = PDP^{-1} = \begin{pmatrix} 1 & -3 \\ 3 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 9 \end{pmatrix}\begin{pmatrix} \frac{1}{10} & \frac{3}{10} \\ -\frac{3}{10} & \frac{1}{10} \end{pmatrix} = \begin{pmatrix} \frac{41}{5} & -\frac{12}{5} \\ -\frac{12}{5} & \frac{9}{5} \end{pmatrix}$$

(Alternatively, A is the matrix for which $Au = u$ and $Av = 9v$.)

(*c*) Use the diagonal factorization of A to obtain

$$B = P\sqrt{D}\,P^{-1} = \begin{pmatrix} 1 & -3 \\ 3 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}\begin{pmatrix} \frac{1}{10} & \frac{3}{10} \\ -\frac{3}{10} & \frac{1}{10} \end{pmatrix} = \begin{pmatrix} \frac{14}{5} & -\frac{3}{5} \\ -\frac{3}{5} & \frac{6}{5} \end{pmatrix}$$

8.37. Let A be an n-square matrix. Without using the Cayley–Hamilton theorem, show that A is a root of a nonzero polynomial.

Let $N = n^2$. Consider the following $N + 1$ matrices

$$I, A, A^2, \dots, A^N$$

Recall that the vector space V of $n \times n$ matrices has dimension $N = n^2$. Thus the above $N + 1$ matrices are linearly dependent. Thus there exist scalars $a_0, a_1, a_2, \dots, a_N$, not all zero, for which

$$a_N A^N + \cdots + a_1 A + a_0 I = 0$$

Thus A is a root of the polynomial $f(t) = a_N t^N + \cdots + a_1 t + a_0$.

8.38. Suppose A is an n-square matrix. Prove the following:

(*a*) A is nonsingular if and only if the constant term of the minimum polynomial of A is not zero.

(*b*) If A is nonsingular, then A^{-1} is equal to a polynomial in A of degree not exceeding n.

(*a*) Suppose $f(t) = t^r + a_{r-1}t^{r-1} + \cdots + a_1 t + a_0$ is the minimum (characteristic) polynomial of A. Then the following are equivalent: (i) A is nonsingular, (ii) 0 is not a root of $f(t)$, and (iii) the constant term a_0 is not zero. Thus the statement is true.

(*b*) Let $m(t)$ be the minimum polynomial of A. Then $m(t) = t^r + a_{r-1}t^{r-1} + \cdots + a_1 t + a_0$, where $r \le n$. Since A is nonsingular, $a_0 \ne 0$ by part (*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,

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

8.39. Let F be an extension of a field K. Let A be an n-square matrix over K. Note that A may also be viewed as a matrix $\hat{A}$ over F. Clearly $|tI - A| = |tI - \hat{A}|$, that is, A and $\hat{A}$ have the same characteristic polynomial. Show that A and $\hat{A}$ also have the same minimum polynomial.

Let $m(t)$ and $m'(t)$ be the minimum polynomials of A and $\hat{A}$, respectively. Now $m'(t)$ divides every polynomial over F which has A as a zero. Since $m(t)$ has A as a zero and since $m(t)$ may be viewed as a polynomial over F, $m'(t)$ divides $m(t)$. We show now that $m(t)$ divides $m'(t)$.

Since $m'(t)$ is a polynomial over F which is an extension of K, we may write

$$m'(t) = f_1(t)b_1 + f_2(t)b_2 + \cdots + f_n(t)b_n$$

where $f_i(t)$ are polynomials over K, and $b_1, \ldots, b_n$ belong to F and are linearly independent over K. We have

$$m'(A) = f_1(A)b_1 + f_2(A)b_2 + \cdots + f_n(A)b_n = 0 \qquad (1)$$

Let $a_{ij}^{(k)}$ denote the ij-entry of $f_k(A)$. The above matrix equation implies that, for each pair (i, j),

$$a_{ij}^{(1)}b_1 + a_{ij}^{(2)}b_2 + \cdots + a_{ij}^{(n)}b_n = 0$$

Since the b_i are linearly independent over K and since the $a_{ij}^{(k)} \in K$, every $a_{ij}^{(k)} = 0$. Then

$$f_1(A) = 0, \; f_2(A) = 0, \ldots, f_n(A) = 0$$

Since the $f_i(t)$ are polynomials over K which have A as a zero and since $m(t)$ is the minimum polynomial of A as a matrix over K, $m(t)$ divides each of the $f_i(t)$. Accordingly, by (1), $m(t)$ must also divide $m'(t)$. But monic polynomials which divide each other are necessarily equal. That is, $m(t) = m'(t)$, as required.

Supplementary Problems

POLYNOMIALS IN MATRICES

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

8.41. Let $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Find A^2, A^3, A^n.

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

8.43. Show that, for any square matrix A, $(P^{-1}AP)^n = P^{-1}A^nP$ where P is invertible. More generally, show that $f(P^{-1}AP) = P^{-1}f(A)P$ for any polynomial $f(t)$.

8.44. Let $f(t)$ be any polynomial. Show that (a) $f(A^T) = (f(A))^T$, and (b) if A is symmetric, then $f(A)$ is symmetric.

EIGENVALUES AND EIGENVECTORS

8.45. Let $A = \begin{pmatrix} 5 & 6 \\ -2 & -2 \end{pmatrix}$. Find: (a) all eigenvalues and linearly independent eigenvectors; (b) P such that $D = P^{-1}AP$ is diagonal; (c) A^{10} and $f(A)$ where $f(t) = t^4 - 5t^3 + 7t^2 - 2t + 5$; and (d) B such that $B^2 = A$.

8.46. For each of the following matrices, find all eigenvalues and a basis for each eigenspace:

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

When possible, find invertible matrices P_1, P_2, and P_3 such that $P_1^{-1}AP_1$, $P_2^{-1}BP_2$, and $P_3^{-1}CP_3$ are diagonal.

8.47. Consider the matrices $A = \begin{pmatrix} 2 & -1 \\ 1 & 4 \end{pmatrix}$ and $B = \begin{pmatrix} 3 & -1 \\ 13 & -3 \end{pmatrix}$. Find all eigenvalues and linearly independent eigenvectors assuming (*a*) A and B are matrices over the real field **R**, and (*b*) A and B are matrices over the complex field **C**.

8.48. Suppose v is a nonzero eigenvector of matrices A and B. Show that v is also an eigenvector of the matrix $kA + k'B$ where k and k' are any scalars.

8.49. Suppose v is a nonzero eigenvector of a matrix A belonging to the eigenvalue λ. Show that for $n > 0$, v is also an eigenvector of A^n belonging to λ^n.

8.50. Suppose λ is an eigenvalue of a matrix A. Show that $f(\lambda)$ is an eigenvalue of $f(A)$ for any polynomial $f(t)$.

8.51. Show that similar matrices have the same eigenvalues.

8.52. Show that matrices A and A^T have the same eigenvalues. Give an example where A and A^T have different eigenvectors.

CHARACTERISTIC AND MINIMUM POLYNOMIALS

8.53. Find the characteristic and minimum polynomials of each of the following matrices:

$$A = \begin{pmatrix} 2 & 5 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 4 & 2 & 0 \\ 0 & 0 & 3 & 5 & 0 \\ 0 & 0 & 0 & 0 & 7 \end{pmatrix} \qquad B = \begin{pmatrix} 3 & 1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 3 & 1 & 0 \\ 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 3 \end{pmatrix} \qquad C = \begin{pmatrix} \lambda & 0 & 0 & 0 & 0 \\ 0 & \lambda & 0 & 0 & 0 \\ 0 & 0 & \lambda & 0 & 0 \\ 0 & 0 & 0 & \lambda & 0 \\ 0 & 0 & 0 & 0 & \lambda \end{pmatrix}$$

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

8.55. Consider a square block matrix $M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$. Show that $tI - M = \begin{pmatrix} tI - A & -B \\ -C & tI - D \end{pmatrix}$ is the characteristic matrix of M.

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

8.57. Show that a matrix A and its transpose A^T have the same minimum polynomial.

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

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

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

8.61. Consider the following n-square matrices (where $a \neq 0$):

$$N = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & \dots & 0 & 0 \end{pmatrix} \qquad M = \begin{pmatrix} \lambda & a & 0 & \dots & 0 & 0 \\ 0 & \lambda & a & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \lambda & a \\ 0 & 0 & 0 & \dots & 0 & \lambda \end{pmatrix}$$

Here N has 1s on the first diagonal above the main diagonal and 0s elsewhere, and M has λ's on the main diagonal, a's on the first diagonal above the main diagonal and 0s elsewhere.

(*a*) Show that, for $k < n$, N^k has 1s on the kth diagonal above the main diagonal and 0s elsewhere, and show that $N^n = 0$.

(*b*) Show that the characteristic polynomial and minimal polynomial of N is $f(t) = t^n$.

(*c*) Show that the characteristic and minimum polynomial of M is $g(t) = (t - \lambda)^n$. (Hint: Note that $M = \lambda I + aN$.)

DIAGONALIZATION

8.62. Let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ be a matrix over the real field **R**. Find necessary and sufficient conditions on a, b, c, and d so that A is diagonalizable, i.e., has two linearly independent eigenvectors.

8.63. Repeat Problem 8.62 for the case that A is a matrix over the complex field **C**.

8.64. Show that a matrix A is diagonalizable if and only if its minimum polynomial is a product of distinct linear factors.

8.65. Suppose E is a matrix such that $E^2 = E$.

(*a*) Find the minimum polynomial $m(t)$ of E.

(*b*) Show that E is diagonalizable and, moreover, E is similar to the diagonal matrix $A = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$ where r is the rank of E.

DIAGONALIZATION OF REAL SYMMETRIC MATRICES AND QUADRATIC FORMS

8.66. For each of the following symmetric matrices A, find an orthogonal matrix P for which $P^{-1}AP$ is diagonal:

$$(a)\quad A = \begin{pmatrix} 1 & 2 \\ 2 & -2 \end{pmatrix}, \qquad (b)\quad A = \begin{pmatrix} 5 & 4 \\ 4 & -1 \end{pmatrix}, \qquad (c)\quad A = \begin{pmatrix} 7 & 3 \\ 3 & -1 \end{pmatrix}$$

8.67. Find an orthogonal transformation of coordinates which diagonalizes each quadratic form:

$$(a)\quad q(x, y) = 2x^2 - 6xy + 10y^2, \qquad (b)\quad q(x, y) = x^2 + 8xy - 5y^2$$

8.68. Find an orthogonal transformation of coordinates which diagonalizes the following quadratic form $q(x, y, z) = 2xy + 2xz + 2yz$.

8.69. Let A be a 2×2 real symmetric matrix with eigenvalues 2 and 3, and let $u = (1, 2)$ be an eigenvector belonging to 2. Find an eigenvector v belonging to 3 and find A.

Answers to Supplementary Problems

8.40. $f(A) = \begin{pmatrix} -26 & -3 \\ 5 & -27 \end{pmatrix}$, $\quad g(A) = \begin{pmatrix} -40 & 39 \\ -65 & -27 \end{pmatrix}$, $\quad f(B) = \begin{pmatrix} 3 & 6 \\ 0 & 9 \end{pmatrix}$, $\quad g(B) = \begin{pmatrix} 3 & 12 \\ 0 & 15 \end{pmatrix}$

8.41. $A^2 = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$, $\quad A^3 = \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}$, $\quad A^n = \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix}$

8.42. *Hint:* Let $A = \begin{pmatrix} 2 & a & b \\ 0 & 2 & c \\ 0 & 0 & 2 \end{pmatrix}$. Set $B = A^3$ and then obtain conditions on a, b, and c.

8.45. (*a*) $\lambda_1 = 1, u = (3, -2); \lambda_2 = 2, v = (2, -1)$ (*b*) $P = \begin{pmatrix} 3 & 2 \\ -2 & -1 \end{pmatrix}$

(*c*) $A^{10} = \begin{pmatrix} 4093 & 6138 \\ -2046 & -3066 \end{pmatrix}, f(A) = \begin{pmatrix} 2 & -6 \\ 2 & 9 \end{pmatrix}$ (*d*) $B = \begin{pmatrix} -3 + 4\sqrt{2} & -6 + 6\sqrt{2} \\ 2 - 2\sqrt{2} & 4 - 3\sqrt{2} \end{pmatrix}$

8.46. (*a*) $\lambda_1 = 2, u = (1, -1, 0), v = (1, 0, -1); \lambda_2 = 6, w = (1, 2, 1)$

(*b*) $\lambda_1 = 3, u = (1, 1, 0), v = (1, 0, 1); \lambda_2 = 1, w = (2, -1, 1)$

(*c*) $\lambda = 1, u = (1, 0, 0), v = (0, 0, 1)$

Let $P_1 = \begin{pmatrix} 1 & 1 & 1 \\ -1 & 0 & 2 \\ 0 & -1 & 1 \end{pmatrix}$ and $P_2 = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 0 & -1 \\ 0 & 1 & 1 \end{pmatrix}$. P_3 does not exist since C has at most two linearly independent eigenvectors, and so cannot be diagonalized.

8.47. (*a*) For A, $\lambda = 3, u = (1, -1)$; B has no eigenvalues (in $\mathbf{R}$);

(*b*) For A, $\lambda = 3, u = (1, -1)$; for B, $\lambda_1 = 2i, u = (1, 3 - 2i); \lambda_2 = -2i, v = (1, 3 + 2i)$.

8.52. Let $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Then $\lambda = 1$ is the only eigenvalue and $v = (1, 0)$ spans the eigenspace of $\lambda = 1$. On the other hand, for $A^T = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$, $\lambda = 1$ is still the only eigenvalue, but $w = (0, 1)$ spans the eigenspace of $\lambda = 1$.

8.53. (*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*) $\Delta(t) = (t - \lambda)^5; m(t) = t - \lambda$

8.60. (*a*) $A = \begin{pmatrix} 0 & 0 & -8 \\ 1 & 0 & -6 \\ 0 & 1 & 5 \end{pmatrix}$, (*b*) $A = \begin{pmatrix} 0 & 0 & 0 & -4 \\ 1 & 0 & 0 & -7 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 5 \end{pmatrix}$

8.65. (*a*) If $E = I$, $m(t) = (t - 1)$; if $E = 0$, $m(t) = t$; otherwise $m(t) = t(t - 1)$.

(*b*) Hint: Use (*a*)

8.66. (*a*) $P = \begin{pmatrix} 2/\sqrt{5} & -1/\sqrt{5} \\ -1/\sqrt{5} & 2/\sqrt{5} \end{pmatrix}$, (*b*) $P = \begin{pmatrix} 2/\sqrt{5} & -1/\sqrt{5} \\ -1/\sqrt{5} & 2/\sqrt{5} \end{pmatrix}$, (*c*) $P = \begin{pmatrix} 3/\sqrt{10} & -1/\sqrt{10} \\ 1/\sqrt{10} & 3/\sqrt{10} \end{pmatrix}$

8.67. (*a*) $x = (3x' - y')/\sqrt{10}, y = (x' + 3y')/\sqrt{10}$, (*b*) $x = (2x' - y')/\sqrt{5}, y = (x' + 2y')/\sqrt{5}$

8.68. $x = x'/\sqrt{3} + y'/\sqrt{2} + z'/\sqrt{6}, y = x'/\sqrt{3} - y'/\sqrt{2} + z'/\sqrt{6}, z = x'/\sqrt{3} - 2z'/\sqrt{6}$

8.69. $v = (2, -1), A = \begin{pmatrix} \frac{14}{5} & -\frac{2}{5} \\ -\frac{2}{5} & \frac{11}{5} \end{pmatrix}$

Similarity and Linear Operators

Suppose A and B are square matrices for which there exists an invertible matrix P such that $B = P^{-1}AP$; then (Section 4.13) B is said to be *similar to* A or is said to be obtained from A by a *similarity transformation*. By Theorem 10.4 and the above remark, we have the following basic result.

Theorem 10.5: Two matrices A and B represent the same linear operator T if and only if they are similar to each other.

That is, all the matrix representations of the linear operator T form an equivalence class of similar matrices.

Now suppose f is a function on square matrices which assigns the same value to similar matrices; that is, $f(A) = f(B)$ whenever A is similar to B. Then f induces a function, also denoted by f, on linear operators T in the following natural way: $f(T) = f([T]_S)$ where S is any basis. The function is well-defined by Theorem 10.5. Three important examples of such functions are:

(1) determinant, (2) trace, and (3) characteristic polynomial

Thus the determinant, trace, and characteristic polynomial of a linear operator T are well-defined.

Example 10.4. Let F be the linear operator on $\mathbf{R}^2$ defined by $F(x, y) = (2x - 3y, 4x + y)$. By Example 10.34, the matrix representation of T relative to the usual basis for $\mathbf{R}^2$ is

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

Accordingly:

(i) $\det(T) = \det(A) = 2 + 12 = 14$ is the determinant of T.
(ii) $\operatorname{tr} T = \operatorname{tr} A = 2 + 1 = 3$ is the trace of T.
(iii) $\Delta_T(t) = \Delta_A(t) = t^2 - 3t + 14$ is the characteristic polynomial of T.

By Example 10.3, another matrix representation of T is the matrix

$$B = \begin{pmatrix} 44 & 101 \\ -18 & -41 \end{pmatrix}$$

Using this matrix, we obtain:

(i) $\det(T) = \det(A) = -1804 + 1818 = 14$ is the determinant of T.
(ii) $\operatorname{tr} T = \operatorname{tr} A = 44 - 41 = 3$ is the trace of T.
(iii) $\Delta_T(t) = \Delta_B(t) = t^2 - 3t + 14$ is the characteristic polynomial of T.

As expected, both matrices yield the same results.

10.4 DIAGONALIZATION OF LINEAR OPERATORS

A linear operator T on a vector space V is said to be *diagonalizable* if T 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

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

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

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

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 in Chapter 8.

A scalar $\lambda \in K$ is called an *eigenvalue* of T if there exists a nonzero vector $v \in V$ for which

$$T(v) = \lambda v$$

Every vector satisfying this relation is called an *eigenvector* of T *belonging* to the eigenvalue λ. The set E_λ of all such vectors 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 following theorems apply.

Theorem 10.6: T can be represented by a diagonal matrix D (or T is diagonalizable) 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 10.7: Nonzero eigenvectors $u_1, u_2, \ldots, u_r$ of T, belonging, respectively, to distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_r$, are linearly independent. (See Problem 10.26 for the proof.)

Theorem 10.8: T is a root of its characteristic polynomial $\Delta(t)$.

Theorem 10.9: The scalar λ is an eigenvalue of T if and only if λ is a root of the characteristic polynomial $\Delta(t)$ of T.

Theorem 10.10: The geometric multiplicity of an eigenvalue λ of T does not exceed its algebraic multiplicity. (See Problem 10.27 for the proof.)

Theorem 10.11: Suppose A is a matrix representation of T. Then T is diagonalizable if and only if A is diagonalizable.

Remark: Theorem 10.11 reduces the investigation of the diagonalization of a linear operator T to the diagonalization of a matrix A which was discussed in detail in Chapter 8.

Example 10.5

(*a*) Let V be the vector space of real functions for which $S = \{\sin\theta, \cos\theta\}$ is a basis, and let $\mathbf{D}$ be the differential operator on V. Then

$$\mathbf{D}(\sin\theta) = \cos\theta = 0(\sin\theta) + 1(\cos\theta)$$
$$\mathbf{D}(\cos\theta) = -\sin\theta = -1(\sin\theta) + 0(\cos\theta)$$

Hence $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ is the matrix representation of $\mathbf{D}$ in the basis S. Therefore,

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

is the characteristic polynomial of both A and $\mathbf{D}$. Thus A and $\mathbf{D}$ have no (real) eigenvalues and, in particular, $\mathbf{D}$ is not diagonalizable.

(*b*) Consider the functions $e^{a_1 t}, e^{a_2 t}, \ldots, e^{a_r t}$ where $a_1, a_2, \ldots, a_r$ are distinct real numbers. Let $\mathbf{D}$ be the differential operator; hence $\mathbf{D}(e^{a_k t}) = a_k e^{a_k t}$. Accordingly, the functions $e^{a_k t}$ are eigenvectors of $\mathbf{D}$ belonging to distinct eigenvalues. Thus, by Theorem 10.7, the functions are linearly independent.

(*c*) Let $T: \mathbf{R}^2 \to \mathbf{R}^2$ be the linear operator which rotates each vector $v \in \mathbf{R}^2$ by an angle $\theta = 90^\circ$ (as shown in Fig. 10-1). Note that no nonzero vector is a multiple of itself. Hence T has no eigenvalues and so no eigenvectors.

Since the mapping $v \mapsto [v]_{S'}$ is onto K^n, we have $P^{-1}[T]_S PX = [T]_{S'} X$ for every $X \in K^n$. Thus $P^{-1}[T]_S P = [T]_{S'}$, as claimed.

DIAGONALIZATION OF LINEAR OPERATORS, EIGENVALUES AND EIGENVECTORS

10.20. Find the eigenvalues and linearly independent eigenvectors of the following linear operator on $\mathbf{R}^2$, and, if it is diagonalizable, find a diagonal representation D: $F(x, y) = (6x - y, 3x + 2y)$.

First find the matrix A which represents F in the usual basis of $\mathbf{R}^2$ by writing down the coefficients of x and y as rows:

$$A = \begin{pmatrix} 6 & -1 \\ 3 & 2 \end{pmatrix}$$

The characteristic polynomial $\Delta(t)$ of F is then

$$\Delta(t) = t^2 - (\text{tr } A)t + |A| = t^2 - 8t + 15 = (t - 3)(t - 5)$$

Thus $\lambda_1 = 3$ and $\lambda_2 = 5$ are eigenvalues of F. We find the corresponding eigenvectors as follows:

(i) Subtract $\lambda_1 = 3$ down the diagonal of A to obtain the matrix $M = \begin{pmatrix} 3 & -1 \\ 3 & -1 \end{pmatrix}$ which corresponds to the homogeneous system $3x - y = 0$. Here $v_1 = (1, 3)$ is a nonzero solution and hence an eigenvector of F belonging to $\lambda_1 = 3$.

(ii) Subtract $\lambda_2 = 5$ down the diagonal of A to obtain $M = \begin{pmatrix} 1 & -1 \\ 3 & -3 \end{pmatrix}$ which corresponds to the system $x - y = 0$. Here $v_2 = (1, 1)$ is a nonzero solution and hence an eigenvector of F belonging to $\lambda_2 = 5$.

Then $S = \{v_1, v_2\}$ is a basis of $\mathbf{R}^2$ consisting of eigenvectors of F. Thus F is diagonalizable, with the matrix representation $D = \begin{pmatrix} 3 & 0 \\ 0 & 5 \end{pmatrix}$.

10.21. Let L be the linear operator on $\mathbf{R}^2$ which reflects points across the line $y = kx$ (where $k \neq 0$). See Fig. 10-2.

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

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

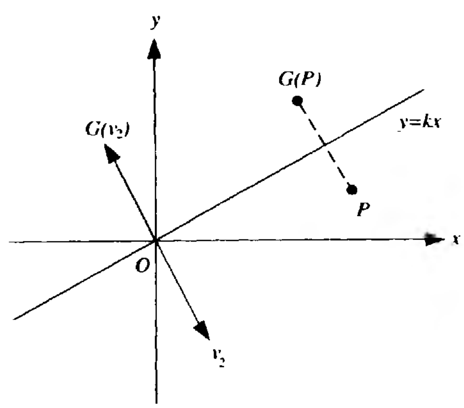

Fig. 10-2

(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 (relative to S) $D = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

10.22. Find all eigenvalues and a basis of each eigenspace of the operator $T: \mathbf{R}^3 \to \mathbf{R}^3$ defined by $T(x, y, z) = (2x + y, y - z, 2y + 4z)$. Is T diagonalizable? If so, find such a representation D.

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

$$A = [T] = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 1 & -1 \\ 0 & 2 & 4 \end{pmatrix} \text{ and so } \Delta(t) = |tI - A| = \begin{vmatrix} t-2 & -1 & 0 \\ 0 & t-1 & 1 \\ 0 & -2 & t-4 \end{vmatrix} = (t-2)^2(t-3)$$

Thus $\lambda = 2$ and $\lambda = 3$ are the eigenvalues of T.

We find a basis of the eigenspace E_2 of $\lambda = 2$. Subtract $\lambda = 2$ down the diagonal of A to obtain the homogeneous system

$$\begin{aligned} y \quad\quad &= 0 \\ -y - z &= 0 \\ 2y + 2z &= 0 \end{aligned} \qquad \text{or} \qquad \begin{aligned} y \quad\quad &= 0 \\ y + z &= 0 \end{aligned}$$

The system has only one independent solution, e.g., $x = 1, y = 0, z = 0$. Thus $u = (1, 0, 0)$ forms a basis of the eigenspace E_2.

We find a basis of the eigenspace E_3 of $\lambda = 3$. Subtract $\lambda = 3$ down the diagonal of A to obtain the homogeneous system

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

The system has only one independent solution, e.g., $x = 1, y = 1, z = -2$. Thus $v = (1, 1, -2)$ forms a basis of the eigenspace E_3.

Observe that T is not diagonalizable, since T has only two linearly independent eigenvectors.

10.23. Show that 0 is an eigenvalue of T if and only if T is singular.

We have that 0 is an eigenvalue of T if and only if there exists a nonzero vector v such that $T(v) = 0v = 0$, i.e., if and only if T is singular.

10.24. Suppose λ is an eigenvalue of an invertible operator T. Show that λ^{-1} is an eigenvalue of T^{-1}.

Since T is invertible, it is also nonsingular; hence, by Problem 10.23 $\lambda \neq 0$.

By definition of an eigenvalue, there exists a nonzero vector v for which $T(v) = \lambda v$. Applying T^{-1} to both sides, we obtain $v = T^{-1}(\lambda v) = \lambda T^{-1}(v)$. Hence $T^{-1}(v) = \lambda^{-1}v$; that is, λ^{-1} is an eigenvalue of T^{-1}.

10.25. Suppose $\dim V = n$. Let $T: V \to V$ be an invertible operator. Show that T^{-1} is equal to a polynomial in T of degree not exceeding n.

Let $m(t)$ be the minimum polynomial of T. Then $m(t) = t^r + a_{r-1}t^{r-1} + \cdots + a_1 t + a_0$, where $r \leq n$. Since T is invertible, $a_0 \neq 0$. We have

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

Hence

$$-\frac{1}{a_0}(T^{r-1} + a_{r-1}T^{r-2} + \cdots + a_1 I)T = I \quad \text{and} \quad T^{-1} = -\frac{1}{a_0}(T^{r-1} + a_{r-1}T^{r-2} + \cdots + a_1 I)$$

10.26. Prove Theorem 10.7.

The proof is by induction on n. If $n = 1$, then u_1 is linearly independent since $u_1 \neq 0$. Assume $n > 1$. Suppose

$$a_1u_1 + a_2u_2 + \cdots + a_nu_n = 0 \tag{1}$$

where the a_i are scalars. Applying T to the above relation, we obtain by linearity

$$a_1T(u_1) + a_2T(u_2) + \cdots + a_nT(u_n) = T(0) = 0$$

But by hypothesis $T(u_i) = \lambda_i u_i$; hence

$$a_1\lambda_1u_1 + a_2\lambda_2u_2 + \cdots + a_n\lambda_nu_n = 0 \tag{2}$$

On the other hand, multiplying (*1*) by λ_n.

$$a_1\lambda_nu_1 + a_2\lambda_nu_2 + \cdots + a_n\lambda_nu_n = 0 \tag{3}$$

Now subtracting (*3*) from (*2*),

$$a_1(\lambda_1 - \lambda_n)u_1 + a_2(\lambda_2 - \lambda_n)u_2 + \cdots + a_{n-1}(\lambda_{n-1} - \lambda_n)u_{n-1} = 0$$

By induction, $u_1, u_2, \ldots, u_{n-1}$ are linearly independent; hence each of the above coefficients is 0. Since the λ_i are distinct, $\lambda_i - \lambda_n \neq 0$ for $i \neq n$. Hence $a_1 = \cdots = a_{n-1} = 0$. Substituting this into (*1*) we get $a_nu_n = 0$, and hence $a_n = 0$. Thus the u_i are linearly independent.

10.27. Prove Theorem 10.10.

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

$$\begin{aligned}
T(v_1) &= \lambda v_1 \\
T(v_2) &= \lambda v_2 \\
&\cdots\cdots\cdots \\
T(v_r) &= \lambda v_r \\
T(w_1) &= a_{11}v_1 + \cdots + a_{1r}v_r + b_{11}w_1 + \cdots + b_{1s}w_s \\
T(w_2) &= a_{21}v_1 + \cdots + a_{2r}v_r + b_{21}w_1 + \cdots + b_{2s}w_s \\
&\cdots\cdots\cdots \\
T(w_s) &= a_{s1}v_1 + \cdots + a_{sr}v_r + b_{s1}w_1 + \cdots + b_{ss}w_s
\end{aligned}$$

The matrix of T in the above basis is

$$M = \left(\begin{array}{cccc|cccc}
\lambda & 0 & \ldots & 0 & a_{11} & a_{21} & \ldots & a_{s1} \\
0 & \lambda & \ldots & 0 & a_{12} & a_{22} & \ldots & a_{s2} \\
\cdots & & & & \cdots & & & \\
0 & 0 & \ldots & \lambda & a_{1r} & a_{2r} & \ldots & a_{sr} \\
\hline
0 & 0 & \ldots & 0 & b_{11} & b_{21} & \ldots & b_{r1} \\
0 & 0 & \ldots & 0 & b_{12} & b_{22} & \ldots & b_{r2} \\
\cdots & & & & \cdots & & & \\
0 & 0 & \ldots & 0 & b_{1s} & b_{2s} & \ldots & b_{ss}
\end{array}\right) = \left(\begin{array}{c|c} \lambda I_r & A \\ \hline 0 & B \end{array}\right)$$

where $A = (a_{ij})^T$ and $B = (b_{ij})^T$.

Since M is a block triangular matrix, the characteristic polynomial of λI_r, which is $(t - \lambda)^r$, must divide the characteristic polynomial of M and hence that of T. Thus the algebraic multiplicity of λ for the operator T is at least r, as required.

10.28. Let $\{v_1, \ldots, v_n\}$ be a basis of V. Let $T: V \to V$ be an operator for which $T(v_1) = 0$, $T(v_2) = a_{21}v_1$, $T(v_3) = a_{31}v_1 + a_{32}v_2, \ldots, T(v_n) = a_{n1}v_1 + \cdots + a_{n,n-1}v_{n-1}$. Show that $T^n = \mathbf{0}$.

It suffices to show that

$$T^j(v_j) = 0 \qquad (*)$$

for $j = 1, \ldots, n$. For then it follows that

$$T^n(v_j) = T^{n-j}(T^j(v_j)) = T^{n-j}(0) = 0, \qquad \text{for } j = 1, \ldots, n$$

and, since $\{v_1, \ldots, v_n\}$ is a basis, $T^n = \mathbf{0}$.

We prove $(*)$ by induction on j. The case $j = 1$ is true by hypothesis. The inductive step follows (for $j = 2, \ldots, n$) from

$$\begin{aligned} T^j(v_j) &= T^{j-1}(T(v_j)) = T^{j-1}(a_{j1}v_1 + \cdots + a_{j,j-1}v_{j-1}) \\ &= a_{j1}T^{j-1}(v_1) + \cdots + a_{j,j-1}T^{j-1}(v_{j-1}) \\ &= a_{j1}0 + \cdots + a_{j,j-1}0 = 0 \end{aligned}$$

Remark: Observe that the matrix representation of T in the above basis is triangular with diagonal elements 0:

$$\begin{pmatrix} 0 & a_{21} & a_{31} & \ldots & a_{n1} \\ 0 & 0 & a_{32} & \ldots & a_{n2} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \ldots & a_{n,n-1} \\ 0 & 0 & 0 & \ldots & 0 \end{pmatrix}$$

MATRIX REPRESENTATIONS OF LINEAR MAPPINGS

10.29. Let $F: \mathbf{R}^3 \to \mathbf{R}^2$ be the linear mapping defined by $F(x, y, z) = (3x + 2y - 4z, x - 5y + 3z)$.

(a) Find the matrix of F in the following bases of $\mathbf{R}^3$ and $\mathbf{R}^2$:

$$S = \{w_1 = (1, 1, 1), w_2 = (1, 1, 0), w_3 = (1, 0, 0)\} \qquad S' = \{u_1 = (1, 3), u_2 = (2, 5)\}$$

(b) Verify that the action of F is preserved by its matrix representation; that is, for any $v \in \mathbf{R}^3$, $[F]_S^{S'}[v]_S = [F(v)]_{S'}$

(a) From Problem 10.2, $(a, b) = (-5a + 2b)u_1 + (3a - b)u_2$. Thus

$$\begin{aligned} F(w_1) &= F(1, 1, 1) = (1, -1) = -7u_1 + 4u_2 \\ F(w_2) &= F(1, 1, 0) = (5, -4) = -33u_1 + 19u_2 \\ F(w_3) &= F(1, 0, 0) = (3, 1) = -13u_1 + 8u_2 \end{aligned}$$

Write the coordinates of $F(w_1)$, $F(w_2)$, $F(w_3)$ as columns to get

$$[F]_S^{S'} = \begin{pmatrix} -7 & -33 & -13 \\ 4 & 19 & 8 \end{pmatrix}$$

(b) If $v = (x, y, z)$ then, by Problem 10.3, $v = zw_1 + (y - z)w_2 + (x - y)w_3$. Also,

$$F(v) = (3x + 2y - 4z, x - 5y + 3z) = (-13x - 20y + 26z)u_1 + (8x + 11y - 15z)u_2$$

Hence $\qquad [v]_S = (z, y - z, x - y)^T \qquad$ and $\qquad [F(v)]_{S'} = \begin{pmatrix} -13x - 20y + 26z \\ 8x + 11y - 15z \end{pmatrix}$

Thus $\qquad [F]_S^{S'}[v]_S = \begin{pmatrix} -7 & -33 & -13 \\ 4 & 19 & 8 \end{pmatrix} \begin{pmatrix} z \\ y - z \\ x - y \end{pmatrix} = \begin{pmatrix} -13x - 20y + 26z \\ 8x + 11y - 15z \end{pmatrix} = [F(v)]_{S'}$