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

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

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

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

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

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

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

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

SOLVED PROBLEMS

Polynomials of Matrices, Characteristic Polynomials

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

(a) A is block triangular with diagonal blocks

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Hence,

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Removing the parentheses and equating corresponding powers of t yields

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

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

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

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

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

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

Eigenvalues and Eigenvectors of 2 × 2 Matrices

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

(a) Find all eigenvalues and corresponding eigenvectors.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

(a) Find all eigenvalues and corresponding eigenvectors.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

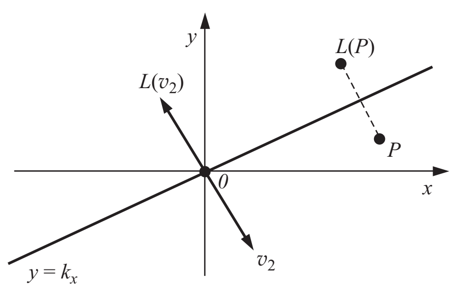

Figure 9-1

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

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

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

Eigenvalues and Eigenvectors

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

9.21. Prove Theorem 9.8$'$: Suppose $v_1, v_2, \ldots, v_n$ are nonzero eigenvectors of T belonging to distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$. Then $v_1, v_2, \ldots, v_n$ are linearly independent.

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

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

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

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

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

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

Multiplying (1) by λ_1 yields

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Diagonalizing Real Symmetric Matrices and Quadratic Forms

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Minimal Polynomial

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

(a) M is block diagonal with diagonal blocks

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

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

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

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

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

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

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

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

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

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