Chapter 4

Square Matrices, Elementary Matrices

4.1 INTRODUCTION

Matrices with the same number of rows as columns are called *square matrices*. These matrices play a major role in linear algebra and will be used throughout the text. This chapter introduces us to these matrices and certain of their elementary properties.

This chapter also introduces us to the elementary matrices which are closely related to the elementary row operations in Chapter 1. We use these matrices to justify two algorithms—one which finds the inverse of a matrix, and the other which diagonalizes a quadratic form.

The scalars in this chapter are real numbers unless otherwise stated or implied. However, we will discuss the special case of complex matrices and some of their properties.

4.2 SQUARE MATRICES

A *square matrix* is a matrix with the same number of rows as columns. An $n \times n$ square matrix is said to be of *order n* and is called an *n-square matrix*.

Recall that not every two matrices can be added or multiplied. However, if we only consider square matrices of some given order n, then this inconvenience disappears. Specifically, the operations of addition, multiplication, scalar multiplication, and transpose can be performed on any $n \times n$ matrices and the result is again an $n \times n$ matrix.

Example 4.1. Let $A = \begin{pmatrix} 1 & 2 & 3 \\ -4 & -4 & -4 \\ 5 & 6 & 7 \end{pmatrix}$ and $B = \begin{pmatrix} 2 & -5 & 1 \\ 0 & 3 & -2 \\ 1 & 2 & -4 \end{pmatrix}$. Then A and B are square matrices of order 3.

Also,

$$A + B = \begin{pmatrix} 3 & -3 & 4 \\ -4 & -1 & -6 \\ 6 & 8 & 3 \end{pmatrix} \qquad 2A = \begin{pmatrix} 2 & 4 & 6 \\ -8 & -8 & -8 \\ 10 & 12 & 14 \end{pmatrix} \qquad A^T = \begin{pmatrix} 1 & -4 & 5 \\ 2 & -4 & 6 \\ 3 & -4 & 7 \end{pmatrix}$$

and

$$AB = \begin{pmatrix} 2+0+3 & -5+6+6 & 1-4-12 \\ -8+0-4 & 20-12-8 & -4+8+16 \\ 10+0+7 & -25+18+14 & 5-12-28 \end{pmatrix} = \begin{pmatrix} 5 & 7 & -15 \\ -12 & 0 & 20 \\ 17 & 7 & -35 \end{pmatrix}$$

are matrices of order 3.

Remark: A nonempty collection **A** of matrices is called an *algebra* (of matrices) if **A** is closed under the operations of matrix addition, scalar multiplication of a matrix, and matrix multiplication. Thus the collection $\mathbf{M}_n$ of all n-square matrices forms an algebra of matrices.

Square Matrices as Functions

Let A be any n-square matrix. Then A may be viewed as a function $A : \mathbf{R}^n \to \mathbf{R}^n$ in two different ways:

(1) $A(u) = Au$ where u is a column vector;

(2) $A(u) = uA$ where u is a row vector.

This book adopts the first meaning of $A(u)$, that is, that the function defined by the matrix A will be $A(u) = Au$. Accordingly, unless otherwise stated or implied, vectors u in $\mathbf{R}^n$ are assumed to be column vectors (not row vectors). However, for typographical convenience, such column vectors u will often be presented as transposed row vectors.

Example 4.2. Let $A = \begin{pmatrix} 1 & -2 & 3 \\ 4 & 5 & -6 \\ 2 & 0 & -1 \end{pmatrix}$. If $u = (1, -3, 7)^T$, then

$$A(u) = Au = \begin{pmatrix} 1 & -2 & 3 \\ 4 & 5 & -6 \\ 2 & 0 & -1 \end{pmatrix}\begin{pmatrix} 1 \\ -3 \\ 7 \end{pmatrix} = \begin{pmatrix} 1+6+21 \\ 4-15-42 \\ 2+0-7 \end{pmatrix} = \begin{pmatrix} 28 \\ -53 \\ -5 \end{pmatrix}$$

If $w = (2, -1, 4)^T$, then

$$A(w) = Aw = \begin{pmatrix} 1 & -2 & 3 \\ 4 & 5 & -6 \\ 2 & 0 & -1 \end{pmatrix}\begin{pmatrix} 2 \\ -1 \\ 4 \end{pmatrix} = \begin{pmatrix} 2+2+12 \\ 8-5-24 \\ 4+0-4 \end{pmatrix} = \begin{pmatrix} 16 \\ -21 \\ 0 \end{pmatrix}$$

Commuting Matrices

Matrices A and B are said to *commute* if $AB = BA$, a condition that applies only for square matrices of the same order. For example, suppose

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 5 & 4 \\ 6 & 11 \end{pmatrix}$$

Then

$$AB = \begin{pmatrix} 5+12 & 4+22 \\ 15+24 & 12+44 \end{pmatrix} = \begin{pmatrix} 17 & 26 \\ 39 & 56 \end{pmatrix}$$

and

$$BA = \begin{pmatrix} 5+12 & 10+16 \\ 6+33 & 12+44 \end{pmatrix} = \begin{pmatrix} 17 & 26 \\ 39 & 56 \end{pmatrix}$$

Since $AB = BA$, the matrices commute.

4.3 DIAGONAL AND TRACE, IDENTITY MATRIX

Let $A = (a_{ij})$ be an n-square matrix. The *diagonal* (or *main diagonal*) of A consists of the elements $a_{11}, a_{22}, \ldots, a_{nn}$. The *trace* of A, written tr A, is the sum of the diagonal elements, that is,

$$\text{tr } A = a_{11} + a_{22} + \cdots + a_{nn} \equiv \sum_{i=1}^{n} a_{ii}$$

The n-square matrix with 1's on the diagonal and 0's elsewhere, denoted by I_n or simply I, is called the *identity* (or *unit*) matrix. The matrix I is similar to the scalar 1 in that, for any matrix A (of the same order),

$$AI = IA = A$$

More generally, if B is an $m \times n$ matrix, then $BI_n = B$ and $I_m B = B$ (Problem 4.9).

For any scalar $k \in K$, the matrix kI which contains k's on the diagonal and 0's elsewhere is called the *scalar matrix* corresponding to the scalar k.

Example 4.3.

(*a*) The Kronecker delta δ_{ij} is defined by

$$\delta_{ij} = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases}$$

Thus the identity matrix may be defined by $I = (\delta_{ij})$.

(*b*) The scalar matrices of orders 2, 3, and 4 corresponding to the scalar $k = 5$ are, respectively,

$$\begin{pmatrix} 5 & 0 \\ 0 & 5 \end{pmatrix} \qquad \begin{pmatrix} 5 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 5 \end{pmatrix} \qquad \begin{pmatrix} 5 & & & \\ & 5 & & \\ & & 5 & \\ & & & 5 \end{pmatrix}$$

(It is common practice to omit blocks or patterns of 0's as in the third matrix.)

The following theorem is proved in Problem 4.10.

Theorem 4.1: Suppose $A = (a_{ij})$ and $B = (b_{ij})$ are n-square matrices and k is a scalar. Then

(i) $\operatorname{tr}(A + B) = \operatorname{tr} A + \operatorname{tr} B$, (ii) $\operatorname{tr} kA = k \cdot \operatorname{tr} A$, (iii) $\operatorname{tr} AB = \operatorname{tr} BA$

4.4 POWERS OF MATRICES, POLYNOMIALS IN MATRICES

Let A be an n-square matrix over a field K. Powers of A are defined as follows:

$$A^2 = AA \qquad A^3 = A^2A, \ldots, A^{n+1} = A^nA, \ldots \qquad \text{and} \qquad A^0 = I$$

Polynomials in the matrix A are also defined. Specifically, for any polynomial

$$f(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$$

where the a_i are scalars, $f(A)$ is defined to be the matrix

$$f(A) = a_0I + a_1A + a_2A^2 + \cdots + a_nA^n$$

[Note that $f(A)$ is obtained from $f(x)$ by substituting the matrix A for the variable x and substituting the scalar matrix a_0I for the scalar a_0.] In the case that $f(A)$ is the zero matrix, the matrix A is called a *zero* or *root* of the polynomial $f(x)$.

Example 4.4. Let $A = \begin{pmatrix} 1 & 2 \\ 3 & -4 \end{pmatrix}$. Then

$$A^2 = \begin{pmatrix} 1 & 2 \\ 3 & -4 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 3 & -4 \end{pmatrix} = \begin{pmatrix} 7 & -6 \\ -9 & 22 \end{pmatrix} \quad \text{and} \quad A^3 = A^2A = \begin{pmatrix} 7 & -6 \\ -9 & 22 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 3 & -4 \end{pmatrix} = \begin{pmatrix} -11 & 38 \\ 57 & -106 \end{pmatrix}$$

If $f(x) = 2x^2 - 3x + 5$, then

$$f(A) = 2\begin{pmatrix} 7 & -6 \\ -9 & 22 \end{pmatrix} - 3\begin{pmatrix} 1 & 2 \\ 3 & -4 \end{pmatrix} + 5\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 16 & -18 \\ -27 & 61 \end{pmatrix}$$

If $g(x) = x^2 + 3x - 10$, then

$$g(A) = \begin{pmatrix} 7 & -6 \\ -9 & 22 \end{pmatrix} + 3\begin{pmatrix} 1 & 2 \\ 3 & -4 \end{pmatrix} - 10\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Thus A is a zero of the polynomial $g(x)$.

The above map from the ring $K[x]$ of polynomials over K into the algebra $\mathbf{M}_n$ of n-square matrices defined by

$$f(x) \to f(A)$$

is called the *evaluation map at* A.

The following theorem (proved in Problem 4.11) applies.

Theorem 4.2: Let $f(x)$ and $g(x)$ be polynomials and let A be an n-square matrix (all over K). Then

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

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

(iii) $f(A)g(A) = g(A)f(A)$.

In other words, (i) and (ii) state that if we first add (multiply) the polynomials $f(x)$ and $g(x)$ and then evaluate the sum (product) at the matrix A, we get the same result as if we first evaluated $f(x)$ and $g(x)$ at A and then added (multiplied) the matrices $f(A)$ and $g(A)$. Part (iii) states that any two polynomials in A commute.

4.5 INVERTIBLE (NONSINGULAR) MATRICES

A square matrix A is said to be *invertible* (or *nonsingular*) if there exists a matrix B with the property that

$$AB = BA = I$$

where I is the identity matrix. Such a matrix B is unique; for

$$AB_1 = B_1 A = I \text{ and } AB_2 = B_2 A = I \quad \text{implies} \quad B_1 = B_1 I = B_1(AB_2) = (B_1 A)B_2 = IB_2 = B_2$$

We call such a matrix B the *inverse* of A and denote it by A^{-1}. Observe that the above relation is symmetric; that is, if B is the inverse of A, then A is the inverse of B.

Example 4.5

(*a*) Suppose $A = \begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix}$ and $B = \begin{pmatrix} 3 & -5 \\ -1 & 2 \end{pmatrix}$. Then

$$AB = \begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix}\begin{pmatrix} 3 & -5 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} 6-5 & -10+10 \\ 3-3 & -5+6 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I$$

$$BA = \begin{pmatrix} 3 & -5 \\ -1 & 2 \end{pmatrix}\begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix} = \begin{pmatrix} 6-5 & 15-15 \\ -2+2 & -5+6 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I$$

Thus A and B are invertible and are inverses of each other.

(*b*) Suppose $A = \begin{pmatrix} 1 & 0 & 2 \\ 2 & -1 & 3 \\ 4 & 1 & 8 \end{pmatrix}$ and $B = \begin{pmatrix} -11 & 2 & 2 \\ -4 & 0 & 1 \\ 6 & -1 & -1 \end{pmatrix}$. Then

$$AB = \begin{pmatrix} -11+0+12 & 2+0-2 & 2+0-2 \\ -22+4+18 & 4+0-3 & 4-1-3 \\ -44-4+48 & 8+0-8 & 8+1-8 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = I$$

By Problem 4.21, $AB = I$ if and only if $BA = I$; hence we do not need to test if $BA = I$. Thus A and B are inverses of each other.

Consider now a general 2×2 matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

We are able to determine when A is invertible and, in such a case, to give a formula for its inverse. First of all, we seek scalars x, y, z, t such that

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x & y \\ z & t \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} ax + bz & ay + bt \\ cx + dz & cy + dt \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

which reduces to solving the following two systems

$$\begin{cases} ax + bz = 1 \\ cx + dz = 0 \end{cases} \qquad \begin{cases} ay + bt = 0 \\ cy + dt = 1 \end{cases}$$

where the original matrix A is the coefficient matrix of each system. Set $|A| = ad - bc$ (the determinant of A). By Problems 1.60 and 1.61, the two systems are solvable, and A is invertible, if and only if $|A| \neq 0$. In that case, the first system has the unique solution $x = d/|A|$, $z = -c/|A|$, and the second system has the unique solution $y = -b/|A|$, $t = a/|A|$. Accordingly,

$$A^{-1} = \begin{pmatrix} d/|A| & -b/|A| \\ -c/|A| & a/|A| \end{pmatrix} = \frac{1}{|A|}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

In words: When $|A| \neq 0$, the inverse of a 2×2 matrix A is obtained by (i) interchanging the elements on the main diagonal, (ii) taking the negatives of the other elements, and (iii) multiplying the matrix by $1/|A|$.

Remark 1: The above property that A is invertible if and only if its determinant $|A| \neq 0$ is true for square matrices of any order. (See Chapter 7.)

Remark 2: Suppose A and B are invertible, then AB is invertible and $(AB)^{-1} = B^{-1}A^{-1}$. More generally, if $A_1, A_2, \ldots, A_k$ are invertible, then their product is invertible and

$$(A_1A_2 \cdots A_k)^{-1} = A_k^{-1} \cdots A_2^{-1}A_1^{-1}$$

the product of the inverses in the reverse order.

4.6 SPECIAL TYPES OF SQUARE MATRICES

This section describes a number of special kinds of square matrices which play an important role in linear algebra.

Diagonal Matrices

A square matrix $D = (d_{ij})$ is *diagonal* if its nondiagonal entries are all zero. Such a matrix is frequently notated as $D = \text{diag}\,(d_{11}, d_{22}, \ldots, d_{nn})$, where some or all of the d_{ii} may be zero. For example,

$$\begin{pmatrix} 3 & 0 & 0 \\ 0 & -7 & 0 \\ 0 & 0 & 2 \end{pmatrix} \qquad \begin{pmatrix} 4 & 0 \\ 0 & -5 \end{pmatrix} \qquad \begin{pmatrix} 6 & & & \\ & 0 & & \\ & & -9 & \\ & & & 1 \end{pmatrix}$$

are diagonal matrices which may be represented, respectively, by

$$\operatorname{diag}(3, -7, 2) \qquad \operatorname{diag}(4, -5) \qquad \text{and} \qquad \operatorname{diag}(6, 0, -9, 1)$$

(Observe that patterns of 0s in the third matrix have been omitted.)

Clearly, the sum, scalar product, and product of diagonal matrices are again diagonal. Thus all the $n \times n$ diagonal matrices form an algebra of matrices. In fact, the diagonal matrices form a commutative algebra since any two $n \times n$ diagonal matrices commute.

Triangular Matrices

A square matrix $A = (a_{ij})$ is an *upper triangular matrix* or simply a *triangular matrix* if all entries below the main diagonal are equal to zero; that is, if $a_{ij} = 0$ for $i > j$. Generic upper triangular matrices of orders 2, 3, and 4 are, respectively,

$$\begin{pmatrix} a_{11} & a_{12} \\ 0 & a_{22} \end{pmatrix} \qquad \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ & b_{22} & b_{23} \\ & & b_{33} \end{pmatrix} \qquad \begin{pmatrix} c_{11} & c_{12} & c_{13} & c_{14} \\ & c_{22} & c_{23} & c_{24} \\ & & c_{33} & c_{34} \\ & & & c_{44} \end{pmatrix}$$

(As in diagonal matrices, it is common practice to omit patterns of 0s.)

The upper triangular matrices also form an algebra of matrices. In fact,

Theorem 4.3: Suppose $A = (a_{ij})$ and $B = (b_{ij})$ are $n \times n$ upper triangular matrices. Then

(i) $A + B$ is upper triangular, with diagonal $(a_{11} + b_{11}, a_{22} + b_{22}, \ldots, a_{nn} + b_{nn})$.

(ii) kA is upper triangular, with diagonal $(ka_{11}, ka_{22}, \ldots, ka_{nn})$.

(iii) AB is upper triangular, with diagonal $(a_{11}b_{11}, a_{22}b_{22}, \ldots, a_{nn}b_{nn})$.

(iv) For any polynomial $f(x)$, the matrix $f(A)$ is upper triangular with diagonal $(f(a_{11}), f(a_{22}), \ldots, f(a_{nn}))$.

(v) A is invertible if and only if each diagonal element $a_{ii} \neq 0$.

Analogously, a *lower triangular matrix* is a square matrix whose entries above the diagonal are all zero, and a theorem analogous to Theorem 4.3 holds for such matrices.

Symmetric Matrices

A real matrix A is *symmetric* if $A^T = A$. Equivalently, $A = (a_{ij})$ is symmetric if *symmetric elements* (mirror images in the diagonal) are equal, i.e., if each $a_{ij} = a_{ji}$. (Note that A must be square in order for $A^T = A$.)

A real matrix A is *skew-symmetric* if $A^T = -A$. Equivalently, $A = (a_{ij})$ is skew-symmetric if each $a_{ij} = -a_{ji}$. Clearly, the diagonal elements of a skew-symmetric matrix must be zero since $a_{ii} = -a_{ii}$ implies $a_{ii} = 0$.

Example 4.6. Consider the following matrices:

$$A = \begin{pmatrix} 2 & -3 & 5 \\ -3 & 6 & 7 \\ 5 & 7 & -8 \end{pmatrix} \qquad B = \begin{pmatrix} 0 & 3 & -4 \\ -3 & 0 & 5 \\ 4 & -5 & 0 \end{pmatrix} \qquad C = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

(*a*) By inspection, the symmetric elements in A are equal, or $A^T = A$. Thus A is symmetric.

(*b*) By inspection, the diagonal elements of B are 0 and symmetric elements are negatives of each other. Thus B is skew-symmetric.

(*c*) Since C is not square, C is neither symmetric nor skew-symmetric.

If A and B are symmetric matrices, then $A + B$ and kA are symmetric. However, AB need not be symmetric. For example,

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} \text{ and } B = \begin{pmatrix} 4 & 5 \\ 5 & 6 \end{pmatrix} \text{ are symmetric, but } AB = \begin{pmatrix} 14 & 17 \\ 23 & 28 \end{pmatrix} \text{ is not symmetric.}$$

Thus the symmetric matrices do not form an algebra of matrices.

The following theorem is proved in Problem 4.29.

Theorem 4.4: If A is a square matrix, then (i) $A + A^T$ is symmetric; (ii) $A - A^T$ is skew-symmetric; (iii) $A = B + C$, for some symmetric matrix B and some skew-symmetric matrix C.

Orthogonal Matrices

A real matrix A is said to be *orthogonal* if $AA^T = A^TA = I$. Observe that an orthogonal matrix A is necessarily square and invertible, with inverse $A^{-1} = A^T$.

Example 4.7. Let $A = \begin{pmatrix} \frac{1}{9} & \frac{8}{9} & -\frac{4}{9} \\ \frac{4}{9} & -\frac{4}{9} & -\frac{7}{9} \\ \frac{8}{9} & \frac{1}{9} & \frac{4}{9} \end{pmatrix}$. Then

$$AA^T = \begin{pmatrix} \frac{1}{9} & \frac{8}{9} & -\frac{4}{9} \\ \frac{4}{9} & -\frac{4}{9} & -\frac{7}{9} \\ \frac{8}{9} & \frac{1}{9} & \frac{4}{9} \end{pmatrix} \begin{pmatrix} \frac{1}{9} & \frac{4}{9} & \frac{8}{9} \\ \frac{8}{9} & -\frac{4}{9} & \frac{1}{9} \\ -\frac{4}{9} & -\frac{7}{9} & \frac{4}{9} \end{pmatrix} = \frac{1}{81} \begin{pmatrix} 1 + 64 + 16 & 4 - 32 + 28 & 8 + 8 - 16 \\ 4 - 32 + 28 & 16 + 16 + 49 & 32 - 4 - 28 \\ 8 + 8 - 16 & 32 - 4 - 28 & 64 + 1 + 16 \end{pmatrix}$$

$$= \frac{1}{81} \begin{pmatrix} 81 & 0 & 0 \\ 0 & 81 & 0 \\ 0 & 0 & 81 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = I$$

This means $A^T = A^{-1}$ and so $A^TA = I$ as well. Thus A is orthogonal.

Consider now an arbitrary 3×3 matrix

$$A = \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix}$$

If A is orthogonal, then

$$AA^T = \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix} \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = I$$

This yields

$$\begin{array}{rrr} a_1^2 + a_2^2 + a_3^2 = 1 & a_1b_1 + a_2b_2 + a_3b_3 = 0 & a_1c_1 + a_2c_2 + a_3c_3 = 0 \\ b_1a_1 + b_2a_2 + b_3a_3 = 0 & b_1^2 + b_2^2 + b_3^2 = 1 & b_1c_1 + b_2c_2 + b_3c_3 = 0 \\ c_1a_1 + c_2a_2 + c_3a_3 = 0 & c_1b_1 + c_2b_2 + c_3b_3 = 0 & c_1^2 + c_2^2 + c_3^2 = 1 \end{array}$$

or, in other words,

$$\begin{array}{lll} u_1 \cdot u_1 = 1 & u_1 \cdot u_2 = 0 & u_1 \cdot u_3 = 0 \\ u_2 \cdot u_1 = 0 & u_2 \cdot u_2 = 1 & u_2 \cdot u_3 = 0 \\ u_3 \cdot u_1 = 0 & u_3 \cdot u_2 = 0 & u_3 \cdot u_3 = 1 \end{array}$$

where $u_1 = (a_1, a_2, a_3)$, $u_2 = (b_1, b_2, b_3)$, $u_3 = (c_1, c_2, c_3)$ are the rows of A. Thus the rows u_1, u_2, and u_3 are orthogonal to each other and have unit lengths, or, in other words, u_1, u_2, u_3 form an *orthonormal set of vectors*. The condition $A^T A = I$ similarly shows that the columns of A form an orthonormal set of vectors. Furthermore, since each step is reversible, the converse is true.

The above result for 3×3 matrices is true in general. That is,

Theorem 4.5: Let A be a real matrix. Then the following are equivalent: (*a*) A is orthogonal; (*b*) the rows of A form an orthonormal set; (*c*) the columns of A form an orthonormal set.

For $n = 2$, we have the following result, proved in Problem 4.33.

Theorem 4.6: Every 2×2 orthogonal matrix has the form $\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$ or $\begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}$ for some real number θ.

Remark: The condition that vectors $u_1, u_2, \ldots, u_m$ form an orthonormal set may be described simply by $u_i \cdot u_j = \delta_{ij}$, where δ_{ij} is the Kronecker delta [Example 4.3(*a*)].

Normal Matrices

A real matrix A is *normal* if A commutes with its transpose, that is, if $AA^T = A^T A$. Clearly, if A is symmetric, orthogonal or skew-symmetric, then A is normal. These, however, are not the only normal matrices.

Example 4.8. Let $A = \begin{pmatrix} 6 & -3 \\ 3 & 6 \end{pmatrix}$. Then

$$AA^T = \begin{pmatrix} 6 & -3 \\ 3 & 6 \end{pmatrix}\begin{pmatrix} 6 & 3 \\ -3 & 6 \end{pmatrix} = \begin{pmatrix} 45 & 0 \\ 0 & 45 \end{pmatrix} \quad \text{and} \quad A^T A = \begin{pmatrix} 6 & 3 \\ -3 & 6 \end{pmatrix}\begin{pmatrix} 6 & -3 \\ 3 & 6 \end{pmatrix} = \begin{pmatrix} 45 & 0 \\ 0 & 45 \end{pmatrix}$$

Since $AA^T = A^T A$, the matrix A is normal.

The following theorem, proved in Problem 4.35, completely characterizes real 2×2 normal matrices.

Theorem 4.7: Let A be a real 2×2 normal matrix. Then A is either symmetric or the sum of a scalar matrix and a skew-symmetric matrix.

4.7 COMPLEX MATRICES

Let A be a complex matrix, i.e., a matrix whose entries are complex numbers. Recall (Section 2.9) that if $z = a + bi$ is a complex number, then $\bar{z} = a - bi$ is its conjugate. The conjugate of the complex matrix A, written $\bar{A}$, is the matrix obtained from A by taking the conjugate of each entry in A, that is, if $A = (a_{ij})$ then $\bar{A} = (b_{ij})$ where $b_{ij} = \overline{a_{ij}}$. [We denote this fact by writing $\bar{A} = (\overline{a_{ij}})$.]

The two operations of transpose and conjugation commute for any complex matrix A, that is, $(\bar{A})^T = \overline{(A^T)}$. In fact, the special notation A^H is used for the conjugate transpose of A. (Note that if A is real then $A^H = A^T$.)

Example 4.9. Let $A = \begin{pmatrix} 2+8i & 5-3i & 4-7i \\ 6i & 1-4i & 3+2i \end{pmatrix}$. Then

$$A^H = \begin{pmatrix} \overline{2+8i} & \overline{6i} \\ \overline{5-3i} & \overline{1-4i} \\ \overline{4-7i} & \overline{3+2i} \end{pmatrix} = \begin{pmatrix} 2-8i & -6i \\ 5+3i & 1+4i \\ 4+7i & 3-2i \end{pmatrix}$$

Hermitian, Unitary, and Normal Complex Matrices

A square complex matrix A is said to be *Hermitian* or *skew-Hermitian* according as

$$A^H = A \qquad \text{or} \qquad A^H = -A$$

If $A = (a_{ij})$ is Hermitian, then $a_{ij} = \overline{a_{ji}}$ and hence each diagonal element a_{ii} must be real. Similarly, if A is skew-Hermitian then each diagonal element $a_{ii} = 0$.

A square complex matrix A is said to be *unitary* if

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

A complex matrix A is unitary if and only if its rows (columns) form an orthonormal set of vectors relative to the inner product of complex vectors. (See Problem 4.39.)

Note that when A is real, Hermitian is the same as symmetric, and unitary is the same as orthogonal.

A square complex matrix A is said to be *normal* if

$$AA^H = A^HA$$

This definition reduces to the one for real matrices when A is real.

Example 4.10. Consider the following matrices:

$$A = \frac{1}{2}\begin{pmatrix} 1 & -i & -1+i \\ i & 1 & 1+i \\ 1+i & -1+i & 0 \end{pmatrix} \qquad B = \begin{pmatrix} 3 & 1-2i & 4+7i \\ 1+2i & -4 & -2i \\ 4-7i & 2i & 2 \end{pmatrix} \qquad C = \begin{pmatrix} 2+3i & 1 \\ i & 1+2i \end{pmatrix}$$

(*a*) A is unitary if $A^H = A^{-1}$ or if $AA^H = A^HA = I$. As noted previously, we need only show that $AA^H = I$:

$$AA^H = A\bar{A}^T = \frac{1}{4}\begin{pmatrix} 1 & -i & -1+i \\ i & 1 & 1+i \\ 1+i & -1+i & 0 \end{pmatrix}\begin{pmatrix} 1 & -i & 1-i \\ i & 1 & -1-i \\ -1-i & 1-i & 0 \end{pmatrix}$$

$$= \frac{1}{4}\begin{pmatrix} 1+1+2 & -i-i+2i & 1-i+i-1+0 \\ i+i-2i & 1+1+2 & i+1-1-i \\ 1+i-i-1+0 & -i+1-1+i+0 & 2+2+0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = I$$

Accordingly, A is unitary.

(*b*) B is Hermitian since its diagonal elements 3, -4, and 2 are real, and the symmetric elements, $1-2i$ and $1+2i$, $4+7i$ and $4-7i$, and $-2i$ and $2i$, are conjugates.

(*c*) To show that C is normal, evaluate CC^H and C^HC:

$$CC^H = C\bar{C}^T = \begin{pmatrix} 2+3i & 1 \\ i & 1+2i \end{pmatrix}\begin{pmatrix} 2-3i & -i \\ 1 & 1-2i \end{pmatrix} = \begin{pmatrix} 14 & 4-4i \\ 4+4i & 6 \end{pmatrix}$$

$$C^HC = \bar{C}^TC = \begin{pmatrix} 2-3i & -i \\ 1 & 1-2i \end{pmatrix}\begin{pmatrix} 2+3i & 1 \\ i & 1+2i \end{pmatrix} = \begin{pmatrix} 14 & 4-4i \\ 4+4i & 6 \end{pmatrix}$$

Since $CC^H = C^HC$, the complex matrix C is normal.

4.8 SQUARE BLOCK MATRICES

A block matrix A is called a *square block matrix* if (i) A is a square matrix, (ii) the blocks form a square matrix, and (iii) the diagonal blocks are also square matrices. The latter two conditions will occur if and only if there are the same number of horizontal and vertical lines and they are placed symmetrically.

Consider the following two block matrices:

$$A = \left(\begin{array}{cc|cc|c} 1 & 2 & 3 & 4 & 5 \\ 1 & 1 & 1 & 1 & 1 \\ \hline 9 & 8 & 7 & 6 & 5 \\ \hline 4 & 4 & 4 & 4 & 4 \\ 3 & 5 & 3 & 5 & 3 \end{array}\right) \qquad B = \left(\begin{array}{cc|cc|c} 1 & 2 & 3 & 4 & 5 \\ 1 & 1 & 1 & 1 & 1 \\ \hline 9 & 8 & 7 & 6 & 5 \\ 4 & 4 & 4 & 4 & 4 \\ \hline 3 & 5 & 3 & 5 & 3 \end{array}\right)$$

Block matrix A is not a square block matrix since the second and third diagonal blocks are not square matrices. On the other hand, block matrix B is a square block matrix.

A *block diagonal matrix* M is a square block matrix where the nondiagonal blocks are all zero matrices. The importance of block diagonal matrices is that the algebra of the block matrix is frequently reduced to the algebra of the individual blocks. Specifically, suppose M is a block diagonal matrix and $f(x)$ is any polynomial. Then M and $f(M)$ have the following form:

$$M = \begin{pmatrix} A_{11} & & & \\ & A_{22} & & \\ & & \cdots & \\ & & & A_{rr} \end{pmatrix} \qquad f(M) = \begin{pmatrix} f(A_{11}) & & & \\ & f(A_{22}) & & \\ & & \cdots & \\ & & & f(A_{rr}) \end{pmatrix}$$

(As usual, we use blank spaces for patterns of zeros or zero blocks.)

Analogously, a square block matrix is called a *block upper triangular matrix* if the blocks below the diagonal are zero matrices, and a *block lower triangular matrix* if the blocks above the diagonal are zero matrices.

4.9 ELEMENTARY MATRICES AND APPLICATIONS

First recall (Section 1.8) the following operations on a matrix A, called *elementary row operations*:

[E_1] (Row-interchange) Interchange the ith row and the jth row:

$$R_i \leftrightarrow R_j$$

[E_2] (Row-scaling) Multiply the ith row by a nonzero scalar k:

$$kR_i \rightarrow R_i \quad (k \neq 0)$$

[E_3] (Row-addition) Replace the ith row by k times the jth row plus the ith row:

$$kR_j + R_i \rightarrow R_i$$

Each of the above operations has an inverse operation of the same type. Specifically (Problem 4.19):

(1) $R_j \rightarrow R_i$ is its own inverse.

(2) $kR_i \rightarrow R_i$ and $k^{-1}R_i \rightarrow R_i$ are inverses.

(3) $kR_j + R_i \rightarrow R_i$ and $-kR_j + R_i \rightarrow R_i$ are inverses.

Also recall (Section 1.8) that a matrix B is said to be *row equivalent* to a matrix A, written $A \sim B$, if B can be obtained from A by a finite sequence of elementary row operations. Since the elementary row operations are reversible, row equivalence is an equivalence relation; that is, (*a*) $A \sim A$; (*b*) if $A \sim B$, then

$B \sim A$; (c) if $A \sim B$ and $B \sim C$, then $A \sim C$. We also restate the following basic result on row equivalence:

Theorem 4.8: Every matrix A is row equivalent to a unique matrix in row canonical form.

Elementary Matrices

Let e denote an elementary row operation and let $e(A)$ denote the result of applying the operation e to a matrix A. The matrix E obtained by applying e to the identity matrix,

$$E = e(I)$$

is called the *elementary matrix* corresponding to the elementary row operation e.

Example 4.11. The 3-square elementary matrices corresponding to the elementary row operations $R_2 \leftrightarrow R_3$, $-6R_2 \to R_2$ and $-4R_1 + R_3 \to R_3$ are, respectively,

$$E_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \qquad E_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -6 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad E_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -4 & 0 & 1 \end{pmatrix}$$

The following theorem, proved in Problem 4.18, shows the fundamental relationship between the elementary row operations and their corresponding elementary matrices.

Theorem 4.9: Let e be an elementary row operation and E the corresponding m-square elementary matrix, i.e., $E = e(I_m)$. Then, for any $m \times n$ matrix A, $e(A) = EA$.

That is, the result of applying an elementary row operation e to a matrix A can be obtained by premultiplying A by the corresponding elementary matrix E.

Now suppose e' is the inverse of an elementary row operation e. Let E' and E be the corresponding matrices. We prove in Problem 4.19 that E is invertible and E' is its inverse. This means, in particular, that any product

$$P = E_k \cdots E_2 E_1$$

of elementary matrices is nonsingular.

Using Theorem 4.9, we are also able to prove (Problem 4.20) the following fundamental result on invertible matrices.

Theorem 4.10: Let A be a square matrix. Then the following are equivalent:

(i) A is invertible (nonsingular);

(ii) A is row equivalent to the identity matrix I;

(iii) A is a product of elementary matrices.

We also use Theorem 4.9 to prove the following theorems.

Theorem 4.11: If $AB = I$, then $BA = I$ and hence $B = A^{-1}$.

Theorem 4.12: B is row equivalent to A if and only if there exists a nonsingular matrix P such that $B = PA$.

Application to Finding Inverses

Suppose a matrix A is invertible and, say, it is row reducible to the identity matrix I by the sequence of elementary operations $e_1, e_2, \ldots, e_q$. Let E_i be the elementary matrix corresponding to the operation e_i. Then, by Theorem 4.9,

$$E_q \cdots E_2 E_1 A = I \qquad \text{or} \qquad (E_q \cdots E_2 E_1 I)A = I \qquad \text{so} \qquad A^{-1} = E_q \cdots E_2 E_1 I$$

In other words, A^{-1} can be obtained by applying the elementary row operations $e_1, e_2, \ldots, e_q$ to the identity matrix I.

The above discussion leads us to the following (Gaussian elimination) algorithm which either finds the inverse of an n-square matrix A or determines that A is not invertible.

Algorithm 4.9: Inverse of a matrix A

Step 1. Form the $n \times 2n$ [block] matrix $M = (A \,\vdots\, I)$; that is, A is in the left half of M and the identity matrix I is in the right half of M.

Step 2. Row reduce M to echelon form. If the process generates a zero row in the A-half of M, STOP (A is not invertible). Otherwise, the A-half will assume triangular form.

Step 3. Further row reduce M to the row canonical form $(I \,\vdots\, B)$, where I has replaced A in the left half of the matrix.

Step 4. Set $A^{-1} = B$.

Example 4.12. Suppose we want to find the inverse of $A = \begin{pmatrix} 1 & 0 & 2 \\ 2 & -1 & 3 \\ 4 & 1 & 8 \end{pmatrix}$. First we form the block matrix $M = (A \,\vdots\, I)$ and reduce M to echelon form:

$$M = \left(\begin{array}{ccc:ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 2 & -1 & 3 & 0 & 1 & 0 \\ 4 & 1 & 8 & 0 & 0 & 1 \end{array}\right) \sim \left(\begin{array}{ccc:ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & -1 & -1 & -2 & 1 & 0 \\ 0 & 1 & 0 & -4 & 0 & 1 \end{array}\right) \sim \left(\begin{array}{ccc:ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & -1 & -1 & -2 & 1 & 0 \\ 0 & 0 & -1 & -6 & 1 & 1 \end{array}\right)$$

In echelon form, the left half of M is in triangular form; hence A is invertible. Next we further reduce M to its row canonical form:

$$M \sim \left(\begin{array}{ccc:ccc} 1 & 0 & 0 & -11 & 2 & 2 \\ 0 & -1 & 0 & 4 & 0 & -1 \\ 0 & 0 & 1 & 6 & -1 & -1 \end{array}\right) \sim \left(\begin{array}{ccc:ccc} 1 & 0 & 0 & -11 & 2 & 2 \\ 0 & 1 & 0 & -4 & 0 & 1 \\ 0 & 0 & 1 & 6 & -1 & -1 \end{array}\right)$$

The identity matrix is in the left half of the final matrix; hence the right half is A^{-1}. In other words,

$$A^{-1} = \begin{pmatrix} -11 & 2 & 2 \\ -4 & 0 & 1 \\ 6 & -1 & -1 \end{pmatrix}$$

4.10 ELEMENTARY COLUMN OPERATIONS, MATRIX EQUIVALENCE

This section repeats some of the discussion of the preceding section using the columns of a matrix instead of the rows. (The choice of first using rows comes from the fact that the row operations are closely related to the operations with linear equations.) We also show the relationship between the row and column operations and their elementary matrices.

The elementary column operations which are analogous to the elementary row operations are as follows:

$[F_1]$ (Column-interchange) Interchange the ith column and the jth column:

$$C_i \leftrightarrow C_j$$

$[F_2]$ (Column-scaling) Multiply the ith column by a nonzero scalar k:

$$kC_i \to C_i \quad (k \neq 0)$$

$[F_3]$ (Column-addition) Replace the ith column by k times the jth column plus the ith column:

$$kC_j + C_i \to C_i$$

Each of the above operations has an inverse operation of the same type just like the corresponding row operations.

Let f denote an elementary column operation. The matrix F, obtained by applying f to the identity matrix I, that is,

$$F = f(I)$$

is called the *elementary matrix* corresponding to the elementary column operation f.

Example 4.13. The 3-square elementary matrices corresponding to the elementary column operations $C_3 \leftrightarrow C_1$, $-2C_3 \to C_3$, and $-5C_2 + C_3 \to C_3$ are, respectively,

$$F_1 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \qquad F_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} \qquad F_3 = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -5 \\ 0 & 0 & 1 \end{pmatrix}$$

Throughout the discussion below, e and f will denote, respectively, corresponding elementary row and column operations, and E and F will denote the corresponding elementary matrices.

Lemma 4.13: Suppose A is any matrix. Then

$$f(A) = [e(A^T)]^T$$

that is, applying the column operation f to a matrix A gives the same result as applying the corresponding row operation e to A^T and then taking the transpose.

The proof of the lemma follows directly from the fact that the columns of A are the rows of A^T, and vice versa. The lemma shows that

$$F = f(I) = [e(I^T)]^T = [e(I)]^T = E^T$$

In other words,

Corollary 4.14: F is the transpose of E.

(Thus F is invertible since E is invertible.) Also, by the above lemma,

$$f(A) = [e(A^T)]^T = [EA^T]^T = (A^T)^T E^T = AF$$

This proves the following theorem (which is analogous to Theorem 4.9 for the elementary row operations):

Theorem 4.15: For any matrix A, $f(A) = AF$.

That is, the result of applying an elementary column operation f on a matrix A can be obtained by postmultiplying A by the corresponding elementary matrix F.

A matrix B is said to be *column equivalent* to a matrix A if B can be obtained from A by a sequence of elementary column operations. Using the argument that is analogous to that for Theorem 4.12 yields:

Theorem 4.16: B is column equivalent to A if and only if there exists a nonsingular matrix Q such that $B = AQ$.

Matrix Equivalence

A matrix B is said to be *equivalent* to a matrix A if B can be obtained from A by a finite sequence of elementary row and column operations. Alternatively (Problem 4.23), B is equivalent to A if there exist nonsingular matrices P and Q such that $B = PAQ$. Just like row equivalence and column equivalence, equivalence of matrices is an equivalence relation.

The main result of this subsection, proved in Problem 4.25, is as follows:

Theorem 4.17: Every $m \times n$ matrix A is equivalent to a unique block matrix of the form

$$\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$$

where I_r is the $r \times r$ identity matrix. (The nonnegative integer r is called the *rank* of A.)

4.11 CONGRUENT SYMMETRIC MATRICES, LAW OF INERTIA

A matrix B is said to be *congruent* to a matrix A if there exists a nonsingular (invertible) matrix P such that

$$B = P^T A P$$

By Problem 4.123, congruence is an equivalence relation. Suppose A is symmetric, i.e., $A^T = A$. Then

$$B^T = (P^T A P)^T = P^T A^T P^{TT} = P^T A P = B$$

and so B is symmetric. Since diagonal matrices are special symmetric matrices, it follows that only symmetric matrices are congruent to diagonal matrices.

The next theorem plays an important role in linear algebra.

Theorem 4.18 (Law of Inertia): Let A be a real symmetric matrix. Then there exists a nonsingular matrix P such that $B = P^T A P$ is diagonal. Moreover, every such diagonal matrix B has the same number $\mathbf{p}$ of positive entries and the same number $\mathbf{n}$ of negative entries.

The *rank* and *signature* of the above real symmetric matrix A are denoted and defined, respectively, by

$$\text{rank } A = \mathbf{p} + \mathbf{n} \qquad \text{and} \qquad \text{sig } A = \mathbf{p} - \mathbf{n}$$

These are uniquely defined by Theorem 4.18. [The notion of rank is actually defined for any matrix (Section 5.7), and the above definition agrees with the general definition.]

Diagonalization Algorithm

The following is an algorithm which diagonalizes (under congruence) a real symmetric matrix $A = (a_{ij})$.

Algorithm 4.11: Congruence diagonalization of a symmetric matrix

Step 1. Form the $n \times 2n$ [block] matrix $M = (A \mid I)$; that is, A is the left half of M and the identity matrix I is the right half of M.

Step 2. Examine the entry a_{11}.

Case I: $a_{11} \neq 0$. Apply the row operations $-a_{i1}R_1 + a_{11}R_i \to R_i$, $i = 2, \ldots, n$, and then apply the corresponding column operations $-a_{i1}C_1 + a_{11}C_i \to C_i$ (to the left half of M) to reduce the matrix M to the form

$$M = \left(\begin{array}{cc|cc} a_{11} & 0 & * & * \\ 0 & B & * & * \end{array}\right) \tag{I}$$

Case II: $a_{11} = 0$ but $a_{ii} \neq 0$, for some $i > 1$. Apply the row operation $R_1 \leftrightarrow R_i$ and then the corresponding column operation $C_1 \leftrightarrow C_i$ to bring a_{ii} into the first diagonal position. This reduces the matrix to Case I.

Case III: All diagonal entries $a_{ii} = 0$. Choose i, j such that $a_{ij} \neq 0$ and apply the row operations $R_j + R_i \to R_i$ and the corresponding column operation $C_j + C_i \to C_i$ to bring $2a_{ij} \neq 0$ into the ith diagonal position. This reduces the matrix to Case II.

In each of the cases, we finally reduce M to the form (I) where B is a symmetric matrix of order less than A.

Remark: The row operations will change both halves of M, but the column operations will only change the left half of M.

Step 3. Repeat Step 2 with each new matrix (neglecting the first row and column of the preceding matrix) until A is diagonalized, that is, until M is transformed into the form $M' = (D, Q)$ where D is diagonal.

Step 4. Set $P = Q^T$. Then $D = P^TAP$.

The justification of the above algorithm is as follows. Let $e_1, e_2, \ldots, e_k$ be all the elementary row operations in the algorithm and let $f_1, f_2, \ldots, f_k$ be the corresponding elementary column operations. Suppose E_i and F_i are the corresponding elementary matrices. By Corollary 4.14,

$$F_i = E_i^T$$

By the above algorithm,

$$Q = E_k \cdots E_2 E_1 I = E_k \cdots E_2 E_1$$

since the right half I of M is only changed by the row operations. On the other hand, the left half A of M is changed by both the row and column operations; therefore,

$$\begin{aligned} D &= E_k \cdots E_2 E_1 A F_1 F_2 \cdots F_k \\ &= (E_k \cdots E_2 E_1) A (E_k \cdots E_2 E_1)^T \\ &= QAQ^T = P^TAP \end{aligned}$$

where $P = Q^T$.

Example 4.14. Suppose $A = \begin{pmatrix} 1 & 2 & -3 \\ 2 & 5 & -4 \\ -3 & -4 & 8 \end{pmatrix}$, a symmetric matrix. To find a nonsingular matrix P such that

$B = P^T AP$ is diagonal, first form the block matrix $(A \mid I)$:

$$(A \mid I) = \begin{pmatrix} 1 & 2 & -3 & 1 & 0 & 0 \\ 2 & 5 & -4 & 0 & 1 & 0 \\ -3 & -4 & 8 & 0 & 0 & 1 \end{pmatrix}$$

Apply the operations $-2R_1 + R_2 \to R_2$ and $3R_1 + R_3 \to R_3$ to $(A \mid I)$ and then the corresponding operations $-2C_1 + C_2 \to C_2$ and $3C_1 + C_3 \to C_3$ to A to obtain

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

Next apply the operation $-2R_2 + R_3 \to R_3$ and then the corresponding operation $-2C_2 + C_3 \to C_3$ to obtain

$$\begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 2 & -2 & 1 & 0 \\ 0 & 0 & -5 & 7 & -2 & 1 \end{pmatrix} \quad \text{and then} \quad \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & -2 & 1 & 0 \\ 0 & 0 & -5 & 7 & -2 & 1 \end{pmatrix}$$

Now A has been diagonalized. Set

$$P = \begin{pmatrix} 1 & -2 & 7 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{and then} \quad B = P^T AP = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -5 \end{pmatrix}$$

Note that B has $\mathbf{p} = 2$ positive entries and $\mathbf{n} = 1$ negative entry.

4.12 QUADRATIC FORMS

A *quadratic form* q in variables $x_1, x_2, \ldots, x_n$ is a polynomial

$$q(x_1, x_2, \ldots, x_n) = \sum_{i<j} c_{ij} x_i x_j \tag{4.1}$$

(where each term has degree two). The quadratic form q is said to be *diagonalized* if

$$q(x_1, x_2, \ldots, x_n) = c_{11}x_1^2 + c_{22}x_2^2 + \cdots + c_{nn}x_n^2$$

that is, if q has no *cross product* terms $x_i x_j$ (where $i \neq j$).

The quadratic form (4.1) may be expressed uniquely in the matrix form

$$q(X) = X^T AX \tag{4.2}$$

where $X = (x_1, x_2, \ldots, x_n)^T$ and $A = (a_{ij})$ is a symmetric matrix. The entries of A can be obtained from (4.1) by setting

$$a_{ii} = c_{ii} \qquad \text{and} \qquad a_{ij} = a_{ji} = c_{ij}/2 \qquad (\text{for } i \neq j)$$

that is, A has diagonal entry a_{ii} equal to the coefficient of x_i^2 and has entries a_{ij} and a_{ji} each equal to half the coefficient of $x_i x_j$. Thus

$$q(X) = (x_1, \ldots, x_n)\begin{pmatrix} a_{11} & a_{22} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

$$= \sum_{i,j} a_{ij}x_i x_j = a_{11}x_1^2 + a_{22}x_2^2 + \cdots + a_{nn}x_n^2 + 2\sum_{i<j} a_{ij}x_i x_j$$

The above symmetric matrix A is called the *matrix representation* of the quadratic form q. Although many matrices A in (4.2) will yield the same quadratic form q, only one such matrix is symmetric.

Conversely, any symmetric matrix A defines a quadratic form q by (*4.2*). Thus there is a one-to-one correspondence between quadratic forms q and symmetric matrices A. Furthermore, a quadratic form q is diagonalized if and only if the corresponding symmetric matrix A is diagonal.

Example 4.15

(*a*) The quadratic form

$$q(x, y, z) = x^2 - 6xy + 8y^2 - 4xz + 5yz + 7z^2$$

may be expressed in the matrix form

$$q(x, y, z) = (x, y, z)\begin{pmatrix} 1 & -3 & -2 \\ -3 & 8 & \frac{5}{2} \\ -2 & \frac{5}{2} & 7 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

where the defining matrix is symmetric. The quadratic form may also be expressed in the matrix form

$$q(x, y, z) = (x, y, z)\begin{pmatrix} 1 & -6 & -4 \\ 0 & 8 & 5 \\ 0 & 0 & 7 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

where the defining matrix is upper triangular.

(*b*) The symmetric matrix $\begin{pmatrix} 2 & 3 \\ 3 & 5 \end{pmatrix}$ determines the quadratic form

$$q(x, y) = (x, y)\begin{pmatrix} 2 & 3 \\ 3 & 5 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = 2x^2 + 6xy + 5y^2$$

Remark: For theoretical reasons, we will always assume a quadratic form q is represented by a symmetric matrix A. Since A is obtained from q by division by 2, we must also assume $1 + 1 \neq 0$ in our field K. This is always true when K is the real field **R** or the complex field **C**.

Change-of-Variable Matrix

Consider a change of variables, say from $x_1, x_2, \ldots, x_n$ to $y_1, y_2, \ldots, y_n$, by means of an invertible linear substitution of the form

$$x_i = p_{i1}y_1 + p_{i2}y_2 + \cdots + p_{in}y_n \qquad (i = 1, 2, \ldots, n)$$

(Here *invertible* means that one can solve for each of the y's uniquely in terms of the x's.) Such a linear substitution can be expressed in the matrix form

$$X = PY \tag{4.3}$$

where

$$X = (x_1, x_2, \ldots, x_n)^T \qquad Y = (y_1, y_2, \ldots, y_n)^T \qquad \text{and} \qquad P = (p_{ij})$$

The matrix P is called the *change-of-variable matrix*; it is nonsingular since the linear substitution is invertible.

Conversely, any nonsingular matrix P defines an invertible linear substitution of variables, $X = PY$. Furthermore,

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

yields the formula for the y's in terms of the x's.

There is a geometrical interpretation of the change-of-variable matrix P which is illustrated in the next example.

Example 4.16. Consider the cartesian plane $\mathbf{R}^2$ with the usual x and y axes, and consider the 2×2 nonsingular matrix

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

The columns $u_1 = (2, 1)^T$ and $u_2 = (-1, 1)^T$ of P determine a new coordinate system of the plane, say with s and t axes, as shown in Fig. 4-1. That is

(1) The s axis is in the direction of u_1 and its unit length is the length of u_1.

(2) The t axis is in the direction of u_2 and its unit length is the length of u_2.

Any point Q in the plane will have coordinates relative to each coordinate system, say $Q(a, b)$ relative to the x and y axes and $Q(a', b')$ relative to the s and t axes. These coordinate vectors are related by the matrix P. Specifically,

$$\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} a' \\ b' \end{pmatrix} \quad \text{or} \quad X = PY$$

where $X = (a, b)^T$ and $Y = (a', b')^T$.

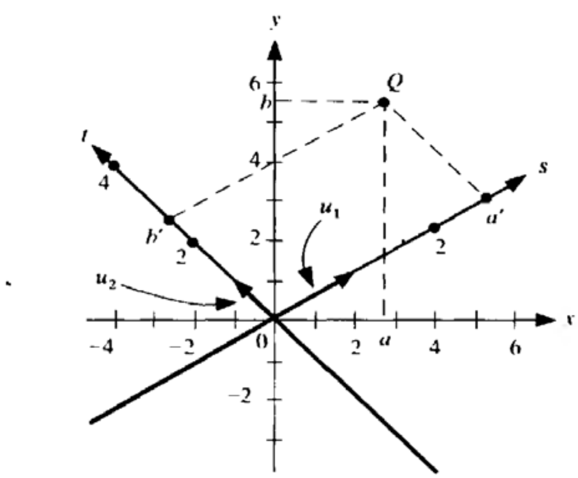

Fig. 4-1

Diagonalizing a Quadratic Form

Consider a quadratic form q in variables $x_1, x_2, \ldots, x_n$, say $q(X) = X^TAX$ (where A is a symmetric matrix). Suppose a change of variables is made in q using the linear substitution (*4.3*). Setting $X = PY$ in q yields the quadratic form $q(Y) = (PY)^TA(PY) = Y^T(P^TAP)Y$. Thus $B = P^TAP$ is the matrix representation of the quadratic form in the new variables $y_1, y_2, \ldots, y_n$. Observe that the new matrix B is congruent to the original matrix A representing q.

The above linear substitution $X = PY$ is said to *diagonalize* the quadratic form $q(X)$ if $q(Y)$ is diagonal, i.e., if $B = P^TAP$ is a diagonal matrix. Since B is congruent to A and A is a symmetric matrix, Theorem 4.18 may be restated as follows.

Theorem 4.19 (Law of Inertia): Let $q(X) = X^TAX$ be a real quadratic form. Then there exists an invertible linear substitution $Y = PX$ which diagonalizes q. Moreover, every such diagonal representation of q has the same number $\mathbf{p}$ of positive entries and the same number $\mathbf{n}$ of negative entries.

The *rank* and *signature* of a real quadratic form q are denoted and defined by

$$\text{rank } q = \mathbf{p} + \mathbf{n} \qquad \text{and} \qquad \text{sig } q = \mathbf{p} - \mathbf{n}$$

These are uniquely defined by Theorem 4.19.

Since diagonalizing a quadratic form is the same as diagonalizing under congruence a symmetric matrix, Algorithm 4.11 may be used here.

Example 4.17. Consider the quadratic form

$$q(x, y, z) = x^2 + 4xy + 5y^2 - 6xz - 8yz + 8z^2 \tag{1}$$

The (symmetric) matrix A which represents q is as follows:

$$A = \begin{pmatrix} 1 & 2 & -3 \\ 2 & 5 & -4 \\ -3 & -4 & 8 \end{pmatrix}$$

From Example 4.14, the following nonsingular matrix P diagonalizes the matrix A under congruence:

$$P = \begin{pmatrix} 1 & -2 & 7 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix} \qquad \text{and} \qquad B = P^T AP = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -5 \end{pmatrix}$$

Accordingly, q may be diagonalized by the following linear substitution:

$$\begin{aligned} x &= r - 2s + 7t \\ y &= \phantom{r - {}} s - 2t \\ z &= t \end{aligned}$$

Specifically, substituting for x, y, and z in (*1*) yields the quadratic form

$$q(r, s, t) = r^2 + s^2 - 5t^2 \tag{2}$$

Here $\mathbf{p} = 2$ and $\mathbf{n} = 1$; hence

$$\text{rank } q = 3 \qquad \text{and} \qquad \text{sig } q = 1$$

Remark: There is a geometrical interpretation of the Law of Inertia (Theorem 4.19) which we give here using the quadratic form q in Example 4.17. Consider the following surface S in $\mathbf{R}^3$:

$$q(x, y, z) = x^2 + 4xy + 5y^2 - 6xz - 8yz + 8z^2 = 25$$

Under the change of variables,

$$\begin{aligned} x &= r - 2s + 7t \\ y &= s - 2t \\ z &= t \end{aligned}$$

or, equivalently, relative to a new coordinate system with r, s, and t axes, the equation of S becomes

$$q(r, s, t) = r^2 + s^2 - 5t^2 = 25$$

Accordingly, S is a hyperboloid of one sheet, since there are two positive and one negative entry on the diagonal. Furthermore, S will always be a hyperboloid of one sheet regardless of the coordinate system. Thus any diagonal representation of the quadratic form $q(x, y, z)$ will contain two positive and one negative entries on the diagonal.

Positive Definite Symmetric Matrices and Quadratic Forms

A real symmetric matrix A is said to be *positive definite* if

$$X^TAX > 0$$

for every nonzero (column) vector X in $\mathbf{R}^n$. Analogously, a quadratic form q is said to be *positive definite* if $q(v) > 0$ for every nonzero vector in $\mathbf{R}^n$.

Alternatively, a real symmetric matrix A or its quadratic form q is *positive definite* if any diagonal representation has only positive diagonal entries. Such matrices and quadratic forms play a very important role in linear algebra. They are considered in Problems 4.54–4.60.

4.13 SIMILARITY

A function $f: \mathbf{R}^n \to \mathbf{R}^n$ may be viewed geometrically as "sending" or "mapping" each point Q into a point $f(Q)$ in the space $\mathbf{R}^n$. Suppose the function f can be represented in the form

$$f(Q) = AQ$$

where A is an $n \times n$ matrix and the coordinates of Q are written as a column vector. Furthermore, suppose P is a nonsingular matrix which may be viewed as introducing a new coordinate system in the space $\mathbf{R}^n$. (See Example 4.16.) Relative to this new coordinate system, we prove that f is represented by the matrix

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

that is,

$$f(Q') = BQ'$$

where Q' denotes the column vector of the coordinates of Q relative to the new coordinate system.

Example 4.18. Consider the function $f: \mathbf{R}^2 \to \mathbf{R}^2$ defined by

$$f(x, y) = (3x - 4y, 5x + 2y)$$

or, equivalently,

$$f\begin{pmatrix} x \\ y \end{pmatrix} = A\begin{pmatrix} x \\ y \end{pmatrix} \quad \text{where} \quad A = \begin{pmatrix} 3 & -4 \\ 5 & 2 \end{pmatrix}$$

Suppose a new coordinate system with s and t axes is introduced in $\mathbf{R}^2$ by means of the nonsingular matrix

$$P = \begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix} \quad \text{and so} \quad P^{-1} = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} \end{pmatrix}$$

(See Fig. 4-1.) Relative to this new coordinate system of $\mathbf{R}^2$, the function f may be represented as

$$f\begin{pmatrix} s \\ t \end{pmatrix} = B\begin{pmatrix} s \\ t \end{pmatrix}$$

where

$$B = P^{-1}AP = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} \end{pmatrix}\begin{pmatrix} 3 & -4 \\ 5 & 2 \end{pmatrix}\begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} \frac{14}{3} & -\frac{10}{3} \\ \frac{22}{3} & \frac{1}{3} \end{pmatrix}$$

In other words,

$$f(s, t) = \left(\tfrac{14}{3}s - \tfrac{10}{3}t, \tfrac{22}{3}s + \tfrac{1}{3}t\right)$$

The above discussion leads us to the following:

Definition: A matrix B is *similar* to a matrix A if there exists a nonsingular matrix P such that

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

Similarity, like congruence, is an equivalence relation (Problem 4.125); hence we say that A and B are similar matrices when $B = P^{-1}AP$.

A matrix A is said to be *diagonalizable* if there exists a nonsingular matrix P such that $B = P^{-1}AP$ is a diagonal matrix. The question of whether or not a given matrix A is diagonalizable and of finding the matrix P when A is diagonalizable plays an important role in linear algebra. These questions will be addressed in Chapter 8.

4.14 LU FACTORIZATION

Suppose A is a nonsingular matrix which can be brought into (upper) triangular form U without using any row-interchange operations, that is, suppose A can be triangularized by the following algorithm which we write using computer algorithmic notation.

Algorithm 4.14: Triangularizing matrix $A = (a_{ij})$

Step 1. Repeat for $i = 1, 2, \ldots, n-1$;

Step 2. Repeat for $j = i+1, \ldots, n$
- (*a*) Set $m_{ij} := a_{ij}/a_{ii}$
- (*b*) Set $R_j := m_{ij}R_i + R_j$

[End of Step 2 inner loop.]

[End of Step 1 outer loop.]

Step 3. Exit.

The numbers m_{ij} are called *multipliers*. Sometimes we keep track of these multipliers by means of the following lower triangular matrix L:

$$L = \begin{pmatrix} 1 & 0 & 0 & \ldots & 0 & 0 \\ -m_{21} & 1 & 0 & \ldots & 0 & 0 \\ -m_{31} & -m_{32} & 1 & \ldots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ -m_{n1} & -m_{n2} & -m_{n3} & \ldots & -m_{n,n-1} & 1 \end{pmatrix}$$

That is, L has 1s on the diagonal, 0s above the diagonal, and the negative of m_{ij} as its ij-entry below the diagonal.

The above lower triangular matrix L may be alternatively described as follows. Let $e_1, e_2, \ldots, e_k$ denote the sequence of elementary row operations in the above algorithm. The inverses of these operations are as follows. For $i = 1, 2, \ldots, n-1$, we have

$$-m_{ij}R_i + R_j \to R_j \qquad (j = i+1, \ldots, n)$$

Applying these inverse operations in reverse order to the identity matrix I yields the matrix L Thus

$$L = E_1^{-1}E_2^{-1} \cdots E_k^{-1}I$$

where $E_1, \ldots, E_k$ are the elementary matrices corresponding to the elementary operations $e_1, \ldots, e_k$.

On the other hand, the elementary operations $e_1, \ldots, e_k$ transform the original matrix A into the upper triangular matrix U. Thus $E_k \cdots E_2E_1A = U$. Accordingly,

$$A = (E_1^{-1}E_2^{-1} \cdots E_k^{-1})U = (E_1^{-1}E_2^{-1} \cdots E_k^{-1}I)U = LU$$

This gives us the classical LU factorization of such a matrix A. We formally state this result as a theorem.

Theorem 4.20: Let A be a matrix as above. Then $A = LU$ where L is a lower triangular matrix with 1s on the diagonal and U is an upper triangular matrix with no 0s on the diagonal.

Remark: We emphasize that the above theorem only applies to nonsingular matrices A which can be brought into triangular form without any row interchanges. Such matrices are said to be *LU-factorable* or to have an LU factorization.

Example 4.19. Let $A = \begin{pmatrix} 1 & 2 & -3 \\ -3 & -4 & 13 \\ 2 & 1 & -5 \end{pmatrix}$. Then A may be reduced to triangular form by the operations $3R_1 + R_2 \to R_2$ and $-2R_1 + R_3 \to R_3$, and then $(\frac{3}{2})R_2 + R_3 \to R_3$:

$$A \sim \begin{pmatrix} 1 & 2 & -3 \\ 0 & 2 & 4 \\ 0 & -3 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & -3 \\ 0 & 2 & 4 \\ 0 & 0 & 7 \end{pmatrix}$$

This gives us the factorization $A = LU$ where

$$L = \begin{pmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 2 & -\frac{3}{2} & 1 \end{pmatrix} \qquad U = \begin{pmatrix} 1 & 2 & -3 \\ 0 & 2 & 4 \\ 0 & 0 & 7 \end{pmatrix}$$

Note that the entries -3, 2, and $-\frac{3}{2}$ in L come from the above elementary row operations, and that U is the triangular form of A.

Applications to Linear Equations

Consider a computer algorithm M. Let $C(n)$ denote the running time of the algorithm as a function of the size n of the input data. [The function $C(n)$ is sometimes called the *time complexity* or simply the *complexity* of the algorithm M.] Frequently, $C(n)$ simply counts the number of multiplications and divisions executed by M, but does not count the number of additions and subtractions since they take much less time to execute.

Now consider a square system of linear equations

$$AX = B$$

where $A = (a_{ij})$ has an LU factorization and

$$X = (x_1, \ldots, x_n)^T \qquad \text{and} \qquad B = (b_1, \ldots, b_n)^T$$

Then the system may be brought into triangular form (in order to apply back-substitution) by applying the above algorithm to the augmented matrix $M = (A, b)$ of the system. The time complexity of the above algorithm and back-substitution are, respectively,

$$C(n) \approx n^3/2 \qquad \text{and} \qquad C(n) \approx n^2/2$$

where n is the number of equations.

On the other hand, suppose we already have the factorization $A = LU$. Then to triangularize the system we need only apply the row operations in the algorithm (retained by the matrix L) to the column vector B. In this case, the time complexity is

$$C(n) \approx n^2/2$$

Of course, to obtain the factorization $A = LU$ requires the original algorithm where $C(n) \approx n^3/2$. Thus nothing may be gained by first finding the LU factorization when a single system is involved. However, there are situations, illustrated below, where the LU factorization is useful.

Suppose that for a given matrix A we need to solve the system

$$AX = B$$

repeatedly for a sequence of different constant vectors, say $B_1, B_2, \ldots, B_k$. Also, suppose some of the B_i depend upon the solution of the system obtained while using preceding vectors B_j. In such a case, it is more efficient to first find the LU factorization of A, and then to use this factorization to solve the system for each new B.

Example 4.20. Consider the system

$$\begin{aligned} x - 2y - z &= k_1 \\ 2x - 5y - z &= k_2 \\ -3x + 10y - 3z &= k_3 \end{aligned} \qquad \text{or} \qquad AX = B \qquad (1)$$

where $A = \begin{pmatrix} 1 & -2 & -1 \\ 2 & -5 & -1 \\ -3 & 10 & -3 \end{pmatrix}$ and $B = \begin{pmatrix} k_1 \\ k_2 \\ k_3 \end{pmatrix}$.

Suppose we want to solve the system for B_1, B_2, B_3, B_4, where $B_1 = (1, 2, 3)^T$ and

$$B_{j+1} = B_j + X_j \qquad (\text{for } j > 1)$$

where X_j is the solution of (1) obtained using B_j. Here it is more efficient to first obtain the LU factorization for A and then to use the LU factorization to solve the system for each of the B's. (See Problem 4.73.)

Solved Problems

ALGEBRA OF SQUARE MATRICES

4.1. Let $A = \begin{pmatrix} 1 & 3 & 6 \\ 2 & -5 & 8 \\ 4 & -2 & 7 \end{pmatrix}$. Find: (*a*) the diagonal and trace of A; (*b*) $A(u)$ where $u = (2, -3, 5)^T$; (*c*) $A(v)$ where $v = (1, 7, -2)$.

(*a*) The diagonal consists of the elements from the upper left corner to the lower right corner of the matrix, that is, the elements a_{11}, a_{22}, a_{33}. Thus the diagonal of A consists of the scalars 1, -5, and 7. The trace of A is the sum of the diagonal elements; hence $\operatorname{tr} A = 1 - 5 + 7 = 3$.

(*b*)
$$A(u) = Au = \begin{pmatrix} 1 & 3 & 6 \\ 2 & -5 & 8 \\ 4 & -2 & 7 \end{pmatrix}\begin{pmatrix} 2 \\ -3 \\ 5 \end{pmatrix} = \begin{pmatrix} 2 - 9 + 30 \\ 4 + 15 + 40 \\ 8 + 6 + 35 \end{pmatrix} = \begin{pmatrix} 23 \\ 59 \\ 49 \end{pmatrix}$$

(*c*) By our convention, $A(v)$ is not defined for a row vector v.

4.2. Let $A = \begin{pmatrix} 1 & 2 \\ 4 & -3 \end{pmatrix}$. (*a*) Find A^2 and A^3. (*b*) Find $f(A)$, where $f(x) = 2x^3 - 4x + 5$.

(*a*)
$$A^2 = AA = \begin{pmatrix} 1 & 2 \\ 4 & -3 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 4 & -3 \end{pmatrix} = \begin{pmatrix} 1+8 & 2-6 \\ 4-12 & 8+9 \end{pmatrix} = \begin{pmatrix} 9 & -4 \\ -8 & 17 \end{pmatrix}$$
$$A^3 = AA^2 = \begin{pmatrix} 1 & 2 \\ 4 & -3 \end{pmatrix}\begin{pmatrix} 9 & -4 \\ -8 & 17 \end{pmatrix} = \begin{pmatrix} 9-16 & -4+34 \\ 36+24 & -16-51 \end{pmatrix} = \begin{pmatrix} -7 & 30 \\ 60 & -67 \end{pmatrix}$$

(*b*) To find $f(A)$, first substitute A for x and $5I$ for the constant 5 in the given polynomial $f(x) = 2x^3 - 4x + 5$:

$$f(A) = 2A^3 - 4A + 5I = 2\begin{pmatrix} -7 & 30 \\ 60 & -67 \end{pmatrix} - 4\begin{pmatrix} 1 & 2 \\ 4 & -3 \end{pmatrix} + 5\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Then multiply each matrix by its respective scalar:

$$f(A) = \begin{pmatrix} -14 & 60 \\ 120 & -134 \end{pmatrix} + \begin{pmatrix} -4 & -8 \\ -16 & 12 \end{pmatrix} + \begin{pmatrix} 5 & 0 \\ 0 & 5 \end{pmatrix}$$

Lastly, add the corresponding elements in the matrices:

$$f(A) = \begin{pmatrix} -14-4+5 & 60-8+0 \\ 120-16+0 & -134+12+5 \end{pmatrix} = \begin{pmatrix} -13 & 52 \\ 104 & -117 \end{pmatrix}$$

4.3. Let $A = \begin{pmatrix} 2 & 2 \\ 3 & -1 \end{pmatrix}$. Find $g(A)$, where $g(x) = x^2 - x - 8$.

$$A^2 = \begin{pmatrix} 2 & 2 \\ 3 & -1 \end{pmatrix}\begin{pmatrix} 2 & 2 \\ 3 & -1 \end{pmatrix} = \begin{pmatrix} 4+6 & 4-2 \\ 6-3 & 6+1 \end{pmatrix} = \begin{pmatrix} 10 & 2 \\ 3 & 7 \end{pmatrix}$$
$$g(A) = A^2 - A - 8I = \begin{pmatrix} 10 & 2 \\ 3 & 7 \end{pmatrix} - \begin{pmatrix} 2 & 2 \\ 3 & -1 \end{pmatrix} - 8\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$
$$= \begin{pmatrix} 10 & 2 \\ 3 & 7 \end{pmatrix} + \begin{pmatrix} -2 & -2 \\ -3 & 1 \end{pmatrix} + \begin{pmatrix} -8 & 0 \\ 0 & -8 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

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

4.4. Given $A = \begin{pmatrix} 1 & 3 \\ 4 & -3 \end{pmatrix}$. Find a *nonzero* column vector $u = \begin{pmatrix} x \\ y \end{pmatrix}$ such that $A(u) = 3u$.

First set up the matrix equation $Au = 3u$:

$$\begin{pmatrix} 1 & 3 \\ 4 & -3 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = 3\begin{pmatrix} x \\ y \end{pmatrix}$$

Write each side as a single matrix (column vector):

$$\begin{pmatrix} x+3y \\ 4x-3y \end{pmatrix} = \begin{pmatrix} 3x \\ 3y \end{pmatrix}$$

Set corresponding elements equal to each other to obtain the system of equations, and reduce it to echelon form:

$$\left.\begin{aligned} x+3y &= 3x \\ 4x-3y &= 3y \end{aligned}\right\} \to \left\{\begin{aligned} 2x-3y &= 0 \\ 4x-6y &= 0 \end{aligned}\right\} \to \left\{\begin{aligned} 2x-3y &= 0 \\ 0 &= 0 \end{aligned}\right\} \to 2x - 3y = 0$$

The system reduces to one homogeneous equation in two unknowns, and so has an infinite number of solutions. To obtain a nonzero solution let, say, $y = 2$; then $x = 3$. That is, $u = (3, 2)^T$ has the desired property.

4.5. Let $A = \begin{pmatrix} 1 & 2 & -3 \\ 2 & 5 & -1 \\ 5 & 12 & -5 \end{pmatrix}$. Find all vectors $u = (x, y, z)^T$ such that $A(u) = 0$.

Set up the equation $Au = 0$ and then write each side as a single matrix:

$$\begin{pmatrix} 1 & 2 & -3 \\ 2 & 5 & -1 \\ 5 & 12 & -5 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} x + 2y - 3z \\ 2x + 5y - z \\ 5x + 12y - 5z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Set corresponding elements equal to each other to obtain a homogeneous system, and reduce the system to echelon form:

$$\left.\begin{array}{r} x + 2y - 3z = 0 \\ 2x + 5y - z = 0 \\ 5x + 12y - 5z = 0 \end{array}\right\} \rightarrow \left\{\begin{array}{r} x + 2y - 3z = 0 \\ y + 5z = 0 \\ 2y + 10z = 0 \end{array}\right\} \rightarrow \left\{\begin{array}{r} x + 2y - 3z = 0 \\ y + 5z = 0 \end{array}\right.$$

In the echelon form, z is the free variable. To obtain the general solution, set $z = a$, where a is a parameter. Back-substitution yields $y = -5a$, and then $x = 13a$. Thus, $u = (13a, -5a, a)^T$ represents all vectors such that $Au = 0$.

4.6. Show that the collection **M** of all 2×2 matrices of the form $\begin{pmatrix} s & t \\ t & s \end{pmatrix}$ is a commutative algebra of matrices.

Clearly, **M** is nonempty. If $A = \begin{pmatrix} a & b \\ b & a \end{pmatrix}$ and $B = \begin{pmatrix} c & d \\ d & c \end{pmatrix}$ belong to **M**, then

$$A + B = \begin{pmatrix} a + c & d + b \\ b + d & a + c \end{pmatrix} \qquad kA = \begin{pmatrix} ka & kb \\ kb & ka \end{pmatrix} \qquad AB = \begin{pmatrix} ac + bd & ad + bc \\ bc + ad & bd + ac \end{pmatrix}$$

also belong to **M**. Thus **M** is an algebra of matrices. Furthermore,

$$BA = \begin{pmatrix} ca + db & cb + da \\ da + cb & db + ca \end{pmatrix}$$

Thus $BA = AB$ and so **M** is a commutative algebra of matrices.

4.7. Find all matrices $M = \begin{pmatrix} x & y \\ z & t \end{pmatrix}$ that commute with $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$.

First find

$$AM = \begin{pmatrix} x + z & y + t \\ z & t \end{pmatrix} \quad \text{and} \quad MA = \begin{pmatrix} x & x + y \\ z & z + t \end{pmatrix}$$

Then set $AM = MA$ to obtain the four equations

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

From the first or last equation, $z = 0$; from the second equation, $x = t$. Thus M is any matrix of the form $\begin{pmatrix} x & y \\ 0 & x \end{pmatrix}$.

4.8. Let $e_i = (0, \ldots, 1, \ldots, 0)^T$, where $i = 1, \ldots, n$, be the (column) vector in $\mathbf{R}^n$ with 1 in the ith position and 0 elsewhere, and let A and B be $m \times n$ matrices.

(*a*) Show that Ae_i is the ith column of A.

(*b*) Suppose $Ae_i = Be_i$ for each i. Show that $A = B$.

(*c*) Suppose $Au = Bu$ for every vector u in $\mathbf{R}^n$. Show that $A = B$.

(*a*) Let $A = (a_{ij})$ and let $Ae_i = (b_1, \ldots, b_n)^T$. Then

$$b_k = R_k e_i = (a_{k1}, \ldots, a_{kn})(0, \ldots, 1, \ldots, 0)^T = a_{ki}$$

where R_k is the kth row of A. Thus

$$Ae_i = (a_{1i}, a_{2i}, \ldots, a_{ni})^T$$

the ith column of A.

(*b*) $Ae_i = Be_i$ means A and B have the same ith column for each i. Thus $A = B$.

(*c*) If $Au = Bu$ for every vector u in $\mathbf{R}^n$, then $Ae_i = Be_i$ for each i. Thus $A = B$.

4.9. Suppose A is an $m \times n$ matrix. Show that: (*a*) $I_m A = A$, (*b*) $AI_n = A$. (Thus $AI = IA = A$ when A is a square matrix.)

We use the fact that $I = (\delta_{ij})$ where δ_{ij} is the Kronecker delta (Example 4.3).

(*a*) Suppose $I_m A = (f_{ij})$. Then

$$f_{ij} = \sum_{k=1}^{m} \delta_{ik} a_{kj} = \delta_{ii} a_{ij} = a_{ij}$$

Thus $I_m A = A$, since corresponding entries are equal.

(*b*) Suppose $AI_n = (g_{ij})$. Then

$$g_{ij} = \sum_{k=1}^{n} a_{ik} \delta_{kj} = a_{ij} \delta_{jj} = a_{ij}$$

Thus $AI_n = A$, since corresponding entries are equal.

4.10. Prove Theorem 4.1. (i) $\text{tr}\,(A + B) = \text{tr}\,(A) + \text{tr}\,(B)$, (ii) $\text{tr}\,(kA) = k\,\text{tr}\,(A)$, (iii) $\text{tr}\,(AB) = \text{tr}\,(BA)$.

(i) Let $A + B = (c_{ij})$. Then $c_{ij} = a_{ij} + b_{ij}$, so that

$$\text{tr}\,(A + B) = \sum_{k=1}^{n} c_{kk} = \sum_{k=1}^{n} (a_{kk} + b_{kk}) = \sum_{k=1}^{n} a_{kk} + \sum_{k=1}^{n} b_{kk} = \text{tr}\,A + \text{tr}\,B$$

(ii) Let $kA = (c_{ij})$. Then $c_{ij} = ka_{ij}$, and

$$\text{tr}\,kA = \sum^{n} ka_{jj} = k \sum^{n} a_{jj} = k \cdot \text{tr}\,A$$

(iii) Let $AB = (c_{ij})$ and $BA = (d_{ij})$. Then $c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$ and $d_{ij} = \sum_{k=1}^{n} b_{ik} a_{kj}$, whence

$$\text{tr}\,AB = \sum_{i=1}^{n} c_{ii} = \sum_{i=1}^{n} \sum_{k=1}^{n} a_{ik} b_{ki} = \sum_{k=1}^{n} \sum_{i=1}^{n} b_{ki} a_{ik} = \sum_{k=1}^{n} d_{kk} = \text{tr}\,BA$$

4.11. Prove Theorem 4.2. (i) $(f + g)(A) = f(A) + g(A)$, (ii) $f(A)g(A) = (fg)(A)$, (iii) $f(A)g(A) = g(A)f(A)$.

Suppose $f(x) = \sum_{i=1}^{r} a_i x^i$ and $g(x) = \sum_{j=1}^{s} b_j x^j$.

(i) We can assume $r = s = n$ by adding powers of x with 0 as their coefficients. Then

$$f(x) + g(x) = \sum_{i=1}^{n} (a_i + b_i) x^i$$

Hence

$$(f+g)(A) = \sum_{i=1}^{n} (a_i + b_i)A^i = \sum_{i=1}^{n} a_i A^i + \sum_{i=1}^{n} b_i A^i = f(A) + g(A)$$

(ii) We have $f(x)g(x) = \sum_{i,j} a_i b_j x^{i+j}$. Then

$$f(A)g(A) = \left(\sum_i a_i A^i\right)\left(\sum_j b_j A^j\right) = \sum_{i,j} a_i b_j A^{i+j} = (fg)(A)$$

(iii) Using $f(x)g(x) = g(x)f(x)$, we have

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

4.12. Let $D_k = kI$, the scalar matrix belonging to the scalar k. Show that (*a*) $D_k A = kA$, (*b*) $BD_k = kB$, (*c*) $D_k + D_{k'} = D_{k+k'}$, and (*d*) $D_k D_{k'} = D_{kk'}$.

(*a*) $D_k A = (kI)A = k(IA) = kA$

(*b*) $BD_k = B(kI) = k(BI) = kB$

(*c*) $D_k + D_{k'} = kI + k'I = (k + k')I = D_{k+k'}$

(*d*) $D_k D_{k'} = (kI)(k'I) = kk'(II) = kk'I = D_{kk'}$

INVERTIBLE MATRICES, INVERSES

4.13. Find the inverse of $\begin{pmatrix} 3 & 5 \\ 2 & 3 \end{pmatrix}$.

Method 1. We seek scalars x, y, z, and w for which

$$\begin{pmatrix} 3 & 5 \\ 2 & 3 \end{pmatrix}\begin{pmatrix} x & y \\ z & w \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} 3x + 5z & 3y + 5w \\ 2x + 3z & 2y + 3w \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

or which satisfy

$$\begin{cases} 3x + 5z = 1 \\ 2x + 3z = 0 \end{cases} \quad \text{and} \quad \begin{cases} 3y + 5w = 0 \\ 2y + 3w = 1 \end{cases}$$

The solution of the first system is $x = -3$, $z = 2$, and of the second system is $y = 5$, $w = -3$. Thus the inverse of the given matrix is $\begin{pmatrix} -3 & 5 \\ 2 & -3 \end{pmatrix}$.

Method 2. The general formula for the inverse of the 2×2 matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is

$$A^{-1} = \begin{pmatrix} d/|A| & -b/|A| \\ -c/|A| & a/|A| \end{pmatrix} = \frac{1}{|A|}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \quad \text{where} \quad |A| = ad - bc$$

Thus if $A = \begin{pmatrix} 3 & 5 \\ 2 & 3 \end{pmatrix}$, then first find $|A| = (3)(3) - (5)(2) = -1 \neq 0$. Next interchange the diagonal elements, take the negatives of the other elements, and multiply by $1/|A|$:

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

4.14. Find the inverse of (*a*) $A = \begin{pmatrix} 1 & 2 & -4 \\ -1 & -1 & 5 \\ 2 & 7 & -3 \end{pmatrix}$ and (*b*) $B = \begin{pmatrix} 1 & 3 & -4 \\ 1 & 5 & -1 \\ 3 & 13 & -6 \end{pmatrix}$.

(a) Form the block matrix $M = (A \mid I)$ and row reduce M to echelon form:

$$M = \begin{pmatrix} 1 & 2 & -4 & \vdots & 1 & 0 & 0 \\ -1 & -1 & 5 & \vdots & 0 & 1 & 0 \\ 2 & 7 & -3 & \vdots & 0 & 0 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & -4 & \vdots & 1 & 0 & 0 \\ 0 & 1 & 1 & \vdots & 1 & 1 & 0 \\ 0 & 3 & 5 & \vdots & -2 & 0 & 1 \end{pmatrix}$$

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

The left half of M is now in triangular form; hence A has an inverse. Further row reduce M to row canonical form:

$$M \sim \begin{pmatrix} 1 & 2 & 0 & \vdots & -9 & -6 & 2 \\ 0 & 1 & 0 & \vdots & \frac{7}{2} & \frac{5}{2} & -\frac{1}{2} \\ 0 & 0 & 1 & \vdots & -\frac{5}{2} & -\frac{3}{2} & \frac{1}{2} \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & \vdots & -16 & -11 & 3 \\ 0 & 1 & 0 & \vdots & \frac{7}{2} & \frac{5}{2} & -\frac{1}{2} \\ 0 & 0 & 1 & \vdots & -\frac{5}{2} & -\frac{3}{2} & \frac{1}{2} \end{pmatrix} = (I \mid A^{-1})$$

Thus $A^{-1} = \begin{pmatrix} -16 & -11 & 3 \\ \frac{7}{2} & \frac{5}{2} & -\frac{1}{2} \\ -\frac{5}{2} & -\frac{3}{2} & \frac{1}{2} \end{pmatrix}$.

(b) Form the block matrix $M = (B \mid I)$ and row reduce to echelon form:

$$\begin{pmatrix} 1 & 3 & -4 & \vdots & 1 & 0 & 0 \\ 1 & 5 & -1 & \vdots & 0 & 1 & 0 \\ 3 & 13 & -6 & \vdots & 0 & 0 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 3 & -4 & \vdots & 1 & 0 & 0 \\ 0 & 2 & 3 & \vdots & -1 & 1 & 0 \\ 0 & 4 & 6 & \vdots & -3 & 0 & 1 \end{pmatrix}$$

$$\sim \begin{pmatrix} 1 & 3 & -4 & \vdots & 1 & 0 & 0 \\ 0 & 2 & 3 & \vdots & -1 & 1 & 0 \\ 0 & 0 & 0 & \vdots & -1 & -2 & 1 \end{pmatrix}$$

In echelon form, M has a zero row in its left half; that is, B is not row reducible to triangular form. Accordingly, B is not invertible.

4.15. Prove the following:

(a) If A and B are invertible, then AB is invertible and $(AB)^{-1} = B^{-1}A^{-1}$.

(b) If $A_1, A_2, \ldots, A_n$ are invertible, then $(A_1 A_2 \cdots A_n)^{-1} = A_n^{-1} \cdots A_2^{-1} A_1^{-1}$.

(c) A is invertible if and only if A^T is invertible.

(d) The operations of inversion and transposing commute: $(A^T)^{-1} = (A^{-1})^T$.

(a) We have

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I$$
$$(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}IB = B^{-1}B = I$$

Thus $B^{-1}A^{-1}$ is the inverse of AB.

(b) By induction on n and using Part (a), we have

$$(A_1 \cdots A_{n-1}A_n)^{-1} = [(A_1 \cdots A_{n-1})A_n]^{-1} = A_n^{-1}(A_1 \cdots A_{n-1})^{-1} = A_n^{-1} \cdots A_2^{-1}A_1^{-1}$$

(c) If A is invertible, then there exists a matrix B such that $AB = BA = I$. Then

$$(AB)^T = (BA)^T = I^T \qquad \text{and so} \qquad B^T A^T = A^T B^T = I$$

Hence A^T is invertible, with inverse B^T. The converse follows from the fact that $(A^T)^T = A$.

(d) By Part (c), B^T is the inverse of A^T; that is $B^T = (A^T)^{-1}$. But $B = A^{-1}$; hence $(A^{-1})^T = (A^T)^{-1}$.

4.16. Show that, if A has a zero row or a zero column, then A is not invertible.

By Problem 3.20, if A has a zero row, then AB would have a zero row. Thus if A were invertible, then $AA^{-1} = I$ would imply that I has a zero row. Therefore, A is not invertible. On the other hand, if A has a zero column, then A^T would have a zero row; and so A^T would not be invertible. Thus, again, A is not invertible.

ELEMENTARY MATRICES

4.17. Find the 3-square elementary matrices E_1, E_2, E_3 which correspond, respectively, to the row operations $R_1 \leftrightarrow R_2$, $-7R_3 \rightarrow R_3$ and $-3R_1 + R_2 \rightarrow R_2$.

Apply the operations to the identity matrix $I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ to obtain

$$E_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad E_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -7 \end{pmatrix} \qquad E_3 = \begin{pmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

4.18. Prove Theorem 4.9.

Let R_i be the ith row of A; we denote this by writing $A = (R_1, \ldots, R_m)$. If B is a matrix for which AB is defined, then it follows directly from the definition of matrix multiplication that $AB = (R_1 B, \ldots, R_m B)$. We also let

$$e_i = (0, \ldots, 0, \hat{1}, 0, \ldots, 0), \qquad \hat{\ } = i$$

Here $\hat{\ } = i$ means that 1 is the ith component. By Problem 4.8, $e_i A = R_i$. We also remark that $I = (e_1, \ldots, e_m)$ is the identity matrix.

(i) Let e be the elementary row operation $R_i \leftrightarrow R_j$. Then, for $\hat{\ } = i$ and $\hat{\hat{\ }} = j$,

$$E = e(I) = (e_1, \ldots, \widehat{e_j}, \ldots, \hat{\hat{e_i}}, \ldots, e_m)$$

and

$$e(A) = (R_1, \ldots, \widehat{R_j}, \ldots, \hat{\hat{R_i}}, \ldots, R_m)$$

Thus

$$EA = (e_1 A, \ldots, \widehat{e_j A}, \ldots, \hat{\hat{e_i A}}, \ldots, e_m A) = (R_1, \ldots, \widehat{R_j}, \ldots, \hat{\hat{R_i}}, \ldots, R_m) = e(A)$$

(ii) Now let e be the elementary row operation $kR_i \rightarrow R_i$, $k \neq 0$. Then, for $\hat{\ } = i$,

$$E = e(I) = (e_1, \ldots, \widehat{ke_i}, \ldots, e_m) \quad \text{and} \quad e(A) = (R_1, \ldots, \widehat{kR_i}, \ldots, R_m)$$

Thus

$$EA = (e_1 A, \ldots, \widehat{ke_i A}, \ldots, e_m A) = (R_1, \ldots, \widehat{kR_i}, \ldots, R_m) = e(A)$$

(iii) Last, let e be the elementary row operation $kR_j + R_i \rightarrow R_i$. Then, for $\hat{\ } = i$,

$$E = e(I) = (e_1, \ldots, \widehat{ke_j + e_i}, \ldots, e_m) \quad \text{and} \quad e(A) = (R_1, \ldots, \widehat{kR_j + R_i}, \ldots, R_m)$$

Using $(ke_j + e_i)A = k(e_j A) + e_i A = kR_j + R_i$, we have

$$EA = (e_1 A, \ldots, \widehat{(ke_j + e_i)A}, \ldots, e_m A) = (R_1, \ldots, \widehat{kR_j + R_i}, \ldots, R_m) = e(A)$$

Thus we have proven the theorem.

4.19. Prove each of the following:

(a) Each of the following elementary row operations has an inverse operation of the same type.

[E_1] Interchange the ith row and the jth row: $R_i \leftrightarrow R_j$.

[E_2] Multiply the ith row by a nonzero scalar k: $kR_i \to R_i$, $k \neq 0$.

[E_3] Replace the ith row by k times the jth row plus the ith row: $kR_j + R_i \to R_i$.

(*b*) Every elementary matrix E is invertible, and its inverse is an elementary matrix.

(*a*) Each operation is treated separately.

(1) Interchanging the same two rows twice, we obtain the original matrix; that is, this operation is its own inverse.

(2) Multiplying the ith row by k and then by k^{-1}, or by k^{-1} and then by k, we obtain the original matrix. In other words, the operations $kR_i \to R_i$ and $k^{-1}R_i \to R_i$ are inverses.

(3) Applying the operation $kR_j + R_i \to R_i$ and then the operation $-kR_j + R_i \to R_i$, or applying the operation $-kR_j + R_i \to R_i$ and then the operation $kR_j + R_i \to R_i$, we obtain the original matrix. In other words, the operations $kR_j + R_i \to R_i$ and $-kR_j + R_i \to R_i$ are inverses.

(*b*) Let E be the elementary matrix corresponding to the elementary row operation e: $e(I) = E$. Let e' be the inverse operation of e and let E' be its corresponding elementary matrix. Then

$$I = e'(e(I)) = e'(E) = E'E \qquad \text{and} \qquad I = e(e'(I)) = e(E') = EE'$$

Therefore E' is the inverse of E.

4.20. Prove Theorem 4.10.

Suppose A is invertible and suppose A is row equivalent to a matrix B in row canonical form. Then there exist elementary matrices $E_1, E_2, \ldots, E_s$ such that $E_s \cdots E_2 E_1 A = B$. Since A is invertible and each elementary matrix E_i is invertible, B is invertible. But if $B \neq I$, then B has a zero row; hence B is not invertible. Thus $B = I$, and (*a*) implies (*b*).

If (*b*) holds, then there exist elementary matrices $E_1, E_2, \ldots, E_s$ such that $E_s \cdots E_2 E_1 A = I$, and so $A = (E_s \cdots E_2 E_1)^{-1} = E_1^{-1} E_2^{-1} \cdots E_s^{-1}$. But the E_i^{-1} are also elementary matrices. Thus (*b*) implies (*c*).

If (*c*) holds, then $A = E_1 E_2 \cdots E_s$. The E_i are invertible matrices; hence their product, A, is also invertible. Thus (*c*) implies (*a*). Accordingly, the theorem is proved.

4.21. Prove Theorem 4.11. If $AB = I$, then $BA = I$ and hence $B = A^{-1}$.

Suppose A is not invertible. Then A is not row equivalent to the identity matrix I, and so A is row equivalent to a matrix with a zero row. In other words, there exist elementary matrices $E_1, \ldots, E_s$ such that $E_s \cdots E_2 E_1 A$ has a zero row. Hence $E_s \cdots E_2 E_1 AB = E_s \cdots E_2 E_1$, an invertible matrix, also has a zero row. But invertible matrices cannot have zero rows; hence A is invertible, with inverse A^{-1}. Then also,

$$B = IB = (A^{-1}A)B = A^{-1}(AB) = A^{-1}I = A^{-1}$$

4.22. Prove Theorem 4.12. $B \sim A$ iff there exists nonsingular P such that $B = PA$.

If $B \sim A$, then $B = e_s(\ldots(e_2(e_1(A)))\ldots) = E_s \cdots E_2 E_1 A = PA$, where $P = E_s \cdots E_2 E_1$ is nonsingular. Conversely, suppose $B = PA$ where P is nonsingular. By Theorem 4.10, P is a product of elementary matrices and hence B can be obtained from A by a sequence of elementary row operations, i.e., $B \sim A$. Thus the theorem is proved.

4.23. Show that B is equivalent to A if and only if there exist invertible matrices P and Q such that $B = PAQ$.

If B is equivalent to A, then $B = E_s \cdots E_2 E_1 A F_1 F_2 \cdots F_t \equiv PAQ$, where $P = E_s \cdots E_2 E_1$ and $Q = F_1 F_2 \cdots F_t$ are invertible. The converse follows from the fact that each step is reversible.

4.24. Show that equivalence of matrices, written $\approx$, is an equivalence relation: (*a*) $A \approx A$, (*b*) If $A \approx B$, then $B \approx A$, (*c*) If $A \approx B$ and $B \approx C$, then $A \approx C$.

(*a*) $A = IAI$ where I is nonsingular; hence $A \approx A$.

(b) If $A \approx B$ then $A = PBQ$ where P and Q are nonsingular. Then $B = P^{-1}AQ^{-1}$ where P^{-1} and Q^{-1} are nonsingular. Hence $B \approx A$.

(c) If $A \approx B$ and $B \approx C$, then $A = PBQ$ and $B = P'CQ'$ where P, Q, P', Q' are nonsingular. Then

$$A = P(P'CQ')Q = (PP')C(QQ')$$

where PP' and QQ' are nonsingular. Hence $A \approx C$.

4.25. Prove Theorem 4.17.

The proof is constructive, in the form of an algorithm.

Step 1. Row reduce A to row canonical form, with leading nonzero entries $a_{11}, a_{2j_2}, \ldots, a_{rj_r}$.

Step 2. Interchange C_2 and C_{j_2}, interchange C_3 and $C_{j_3}, \ldots$, and interchange C_r and C_{j_r}. This gives a matrix in the form $\begin{pmatrix} I_r & B \\ 0 & 0 \end{pmatrix}$, with leading nonzero entries $a_{11}, a_{22}, \ldots, a_{rr}$.

Step 3. Use column operations, with the a_{ii} as pivots, to replace each entry in B with a zero; i.e., for

$$i = 1, 2, \ldots, r \qquad \text{and} \qquad j = r + 1, r + 2, \ldots, n,$$

apply the operation $-b_{ij}C_i + C_j \to C_j$.

The final matrix has the desired form $\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$.

SPECIAL TYPES OF MATRICES

4.26. Find an upper triangular matrix A such that $A^3 = \begin{pmatrix} 8 & -57 \\ 0 & 27 \end{pmatrix}$.

Set $A = \begin{pmatrix} x & y \\ 0 & z \end{pmatrix}$. Then A^3 has the form $\begin{pmatrix} x^3 & * \\ 0 & z^3 \end{pmatrix}$. Thus $x^3 = 8$, so $x = 2$; $z^3 = 27$, so $z = 3$. Next calculate A^3 using $x = 2$ and $z = 3$:

$$A^2 = \begin{pmatrix} 2 & y \\ 0 & 3 \end{pmatrix}\begin{pmatrix} 2 & y \\ 0 & 3 \end{pmatrix} = \begin{pmatrix} 4 & 5y \\ 0 & 9 \end{pmatrix} \qquad \text{and} \qquad A^3 = \begin{pmatrix} 2 & y \\ 0 & 3 \end{pmatrix}\begin{pmatrix} 4 & 5y \\ 0 & 9 \end{pmatrix} = \begin{pmatrix} 8 & 19y \\ 0 & 27 \end{pmatrix}$$

Thus $19y = -57$, or $y = -3$. Accordingly, $A = \begin{pmatrix} 2 & -3 \\ 0 & 3 \end{pmatrix}$.

4.27. Prove Theorem 4.3(iii).

Let $AB = (c_{ij})$. Then

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj} \qquad \text{and} \qquad c_{ii} = \sum_{k=1}^{n} a_{ik} b_{ki}$$

Suppose $i > j$. Then, for any k, either $i > k$ or $k > j$, so that either $a_{ik} = 0$ or $b_{kj} = 0$. Thus, $c_{ik} = 0$, and AB is upper triangular. Suppose $i = j$. Then, for $k < i$, $a_{ik} = 0$; and, for $k > i$, $b_{ki} = 0$. Hence $c_{ii} = a_{ii}b_{ii}$, as claimed.

4.28. What kinds of matrices are both upper triangular and lower triangular?

If A is both upper and lower triangular, then every entry off the main diagonal must be zero. Hence A is diagonal.

4.29. Prove Theorem 4.4.

(i) $(A + A^T)^T = A^T + (A^T)^T = A^T + A = A + A^T$

(ii) $(A - A^T)^T = A^T - (A^T)^T = A^T - A = -(A - A^T)$

(iii) Choose $B \equiv \frac{1}{2}(A + A^T)$ and $C \equiv \frac{1}{2}(A - A^T)$, and appeal to (i) and (ii). Observe that no other choice is possible.

4.30. Write $A = \begin{pmatrix} 2 & 3 \\ 7 & 8 \end{pmatrix}$ as the sum of a symmetric matrix B and a skew-symmetric matrix C.

Calculate $A^T = \begin{pmatrix} 2 & 7 \\ 3 & 8 \end{pmatrix}$, $A + A^T = \begin{pmatrix} 4 & 10 \\ 10 & 16 \end{pmatrix}$, and $A - A^T = \begin{pmatrix} 0 & -4 \\ 4 & 0 \end{pmatrix}$. Then

$$B = \frac{1}{2}(A + A^T) = \begin{pmatrix} 2 & 5 \\ 5 & 8 \end{pmatrix} \qquad C = \frac{1}{2}(A - A^T) = \begin{pmatrix} 0 & -2 \\ 2 & 0 \end{pmatrix}$$

4.31. Find x, y, z, s, t if $A = \begin{pmatrix} x & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} & y \\ z & s & t \end{pmatrix}$ is orthogonal.

Let R_1, R_2, R_3 denote the rows of A, and let C_1, C_2, C_3 denote the columns of A. Since R_1 is a unit vector, $x^2 + \frac{4}{9} + \frac{4}{9} = 1$, or $x = \pm\frac{1}{3}$. Since R_2 is a unit vector, $\frac{4}{9} + \frac{1}{9} + y^2 = 1$, or $y = \pm\frac{2}{3}$. Since $R_1 \cdot R_2 = 0$, we get $2x/3 + \frac{2}{9} + 2y/3 = 0$, or $3x + 3y = -1$. The only possibility is that $x = \frac{1}{3}$ and $y = -\frac{2}{3}$. Thus

$$A = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} & -\frac{2}{3} \\ z & s & t \end{pmatrix}$$

Since the columns are unit vectors,

$$\tfrac{1}{9} + \tfrac{4}{9} + z^2 = 1 \qquad \tfrac{4}{9} + \tfrac{1}{9} + s^2 = 1 \qquad \tfrac{4}{9} + \tfrac{4}{9} + t^2 = 1$$

Thus $z = \pm\frac{2}{3}$, $s = \pm\frac{2}{3}$, and $t = \pm\frac{1}{3}$.

Case (i): $z = \frac{2}{3}$. Since C_1 and C_2 are orthogonal, $s = -\frac{2}{3}$; since C_1 and C_3 are orthogonal, $t = \frac{1}{3}$.

Case (ii): $z = -\frac{2}{3}$. Since C_1 and C_2 are orthogonal, $s = \frac{2}{3}$; since C_1 and C_3 are orthogonal, $t = -\frac{1}{3}$.

Hence there are exactly two possible solutions:

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

4.32. Suppose $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is orthogonal. Show that $a^2 + b^2 = 1$ and

$$A = \begin{pmatrix} a & b \\ b & -a \end{pmatrix} \quad \text{or} \quad A = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$$

Since A is orthogonal, the rows of A form an orthonormal set. Hence

$$a^2 + b^2 = 1 \qquad c^2 + d^2 = 1 \qquad ac + bd = 0$$

Similarly, the columns form an orthonormal set, so

$$a^2 + c^2 = 1 \qquad b^2 + d^2 = 1 \qquad ab + cd = 0$$

Therefore, $c^2 = 1 - a^2 = b^2$, whence $c = \pm b$.

Case (i): $c = +b$. Then $b(a + d) = 0$, or $d = -a$; the corresponding matrix is $\begin{pmatrix} a & b \\ b & -a \end{pmatrix}$.

Case (ii): $c = -b$. Then $b = (d - a) = 0$, or $d = a$; the corresponding matrix is $\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$.

4.33. Prove Theorem 4.6.

Let a and b be any real numbers such that $a^2 + b^2 = 1$. Then there exists a real number θ such that $a = \cos\theta$ and $b = \sin\theta$. The result now follows from Problem 4.32.

4.34. Find a 3×3 orthogonal matrix P whose first row is a multiple of $u_1 = (1, 1, 1)$ and whose second row is a multiple of $u_2 = (0, -1, 1)$.

First find a vector u_3 orthogonal to u_1 and u_2, say (cross product) $u_3 = u_1 \times u_2 = (2, -1, -1)$. Let A be the matrix whose rows are u_1, u_2, u_3; and let P be the matrix obtained from A by normalizing the rows of A. Thus

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & -1 & 1 \\ 2 & -1 & -1 \end{pmatrix} \quad \text{and} \quad P = \begin{pmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \\ 0 & -1/\sqrt{2} & 1/\sqrt{2} \\ 2/\sqrt{6} & -1/\sqrt{6} & -1/\sqrt{6} \end{pmatrix}$$

4.35. Prove Theorem 4.7.

Suppose $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Then

$$AA^T = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} a & c \\ b & d \end{pmatrix} = \begin{pmatrix} a^2 + b^2 & ac + bd \\ ac + bd & c^2 + d^2 \end{pmatrix}$$

$$A^TA = \begin{pmatrix} a & c \\ b & d \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a^2 + c^2 & ab + cd \\ ab + cd & b^2 + d^2 \end{pmatrix}$$

Since $AA^T = A^TA$, we get

$$a^2 + b^2 = a^2 + c^2 \qquad c^2 + d^2 = b^2 + d^2 \qquad ac + bd = ab + cd$$

The first equation yields $b^2 = c^2$; hence $b = c$ or $b = -c$.

Case (i): $b = c$ (which includes the case $b = -c = 0$). Then we obtain the symmetric matrix $A = \begin{pmatrix} a & b \\ b & d \end{pmatrix}$.

Case (ii): $b = -c \neq 0$. Then $ac + bd = b(d - a)$ and $ab + cd = b(a - d)$. Thus $b(d - a) = b(a - d)$, and so $2b(d - a) = 0$. Since $b \neq 0$, we get $a = d$. Thus A has the form

$$A = \begin{pmatrix} a & b \\ -b & a \end{pmatrix} = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} + \begin{pmatrix} 0 & b \\ -b & 0 \end{pmatrix}$$

which is the sum of a scalar matrix and a skew-symmetric matrix.

COMPLEX MATRICES

4.36. Find the conjugate of the matrix $A = \begin{pmatrix} 2 + i & 3 - 5i & 4 + 8i \\ 6 - i & 2 - 9i & 5 + 6i \end{pmatrix}$.

Take the conjugate of each element of the matrix (where $\overline{a + bi} = a - bi$):

$$\bar{A} = \overline{\begin{pmatrix} 2 + i & 3 - 5i & 4 + 8i \\ 6 - i & 2 - 9i & 5 + 6i \end{pmatrix}} = \begin{pmatrix} 2 - i & 3 + 5i & 4 - 8i \\ 6 + i & 2 + 9i & 5 - 6i \end{pmatrix}$$

4.37. Find A^H when $A = \begin{pmatrix} 2-3i & 5+8i \\ -4 & 3-7i \\ -6-i & 5i \end{pmatrix}$.

$A^H = \bar{A}^T$, the conjugate transpose of A. Hence,

$$A^H = \begin{pmatrix} \overline{2-3i} & \overline{-4} & \overline{-6-i} \\ \overline{5+8i} & \overline{3-7i} & \overline{5i} \end{pmatrix} = \begin{pmatrix} 2+3i & -4 & -6+i \\ 5-8i & 3+7i & -5i \end{pmatrix}$$

4.38. Write $A = \begin{pmatrix} 2+6i & 5+3i \\ 9-i & 4-2i \end{pmatrix}$ in the form $A = B + C$, where B is Hermitian and C is skew-Hermitian.

First find

$$A^H = \begin{pmatrix} 2-6i & 9+i \\ 5-3i & 4+2i \end{pmatrix} \qquad A + A^H = \begin{pmatrix} 4 & 14+4i \\ 14-4i & 8 \end{pmatrix} \qquad A - A^H = \begin{pmatrix} 12i & -4+2i \\ 4+2i & -4i \end{pmatrix}$$

Then the required matrices are

$$B = \frac{1}{2}(A + A^H) = \begin{pmatrix} 2 & 7+2i \\ 7-2i & 4 \end{pmatrix} \quad \text{and} \quad C = \frac{1}{2}(A - A^H) = \begin{pmatrix} 6i & -2+i \\ 2+i & -2i \end{pmatrix}$$

4.39. Define an orthonormal set of vectors in $\mathbf{C}^n$ and prove the following complex analogue of Theorem 4.5:

Theorem: Let A be a complex matrix. Then the following are equivalent: (*a*) A is unitary; (*b*) the rows of A form an orthonormal set; (*c*) the columns of A form an orthonormal set.

The vectors $u_1, u_2, \ldots, u_r$ in $\mathbf{C}^n$ form an orthonormal set if $u_i \cdot u_j = \delta_{ij}$ where the dot product in $\mathbf{C}^n$ is defined by

$$(a_1, a_2, \ldots, a_n) \cdot (b_1, b_2, \ldots, b_n) = a_1\overline{b_1} + a_2\overline{b_2} + \cdots + a_n\overline{b_n}$$

and δ_{ij} is the Kronecker delta [see Example 4.3(*a*)].

Let $R_1, \ldots, R_n$ denote the rows of A; then $\bar{R}_1^T, \ldots, \bar{R}_n^T$ are the columns of A^H. Let $AA^H = (c_{ij})$. By matrix multiplication, $c_{ij} = R_i \bar{R}_j^T = R_i \cdot R_j$. Then $AA^H = I$ iff $R_i \cdot R_j = \delta_{ij}$ iff $R_1, R_2, \ldots, R_n$ form an orthonormal set. Thus (*a*) and (*b*) are equivalent. Similarly, A is unitary iff A^H is unitary iff the rows of A^H are orthonormal iff the conjugates of the columns of A are orthonormal iff the columns of A are orthonormal. Thus (*a*) and (*c*) are equivalent, and the theorem is proved.

4.40. Show that $A = \begin{pmatrix} \frac{1}{2} - \frac{2}{3}i & \frac{2}{3}i \\ -\frac{2}{3}i & -\frac{1}{3} - \frac{2}{3}i \end{pmatrix}$ is unitary.

The rows form an orthonormal set:

$$(\tfrac{1}{3} - \tfrac{2}{3}i, \tfrac{2}{3}i) \cdot (\tfrac{1}{3} - \tfrac{2}{3}i, \tfrac{2}{3}i) = (\tfrac{1}{9} + \tfrac{4}{9}) + \tfrac{4}{9} = 1$$

$$(\tfrac{1}{3} - \tfrac{2}{3}i, \tfrac{2}{3}i) \cdot (-\tfrac{2}{3}i, -\tfrac{1}{3} - \tfrac{2}{3}i) = (\tfrac{2}{9}i + \tfrac{4}{9}) + (-\tfrac{2}{9}i - \tfrac{4}{9}) = 0$$

$$(-\tfrac{2}{3}i, -\tfrac{1}{3} - \tfrac{2}{3}i) \cdot (-\tfrac{2}{3}i, -\tfrac{1}{3} - \tfrac{2}{3}i) = \tfrac{4}{9} + (\tfrac{1}{9} + \tfrac{4}{9}) = 1$$

Thus A is unitary.

SQUARE BLOCK MATRICES

4.41. Determine which matrix is a square block matrix:

$$A = \left(\begin{array}{cc|c|cc} 1 & 2 & 3 & 4 & 5 \\ 1 & 1 & 1 & 1 & 1 \\ \hline 9 & 8 & 7 & 6 & 5 \\ 3 & 3 & 3 & 3 & 3 \\ \hline 1 & 3 & 5 & 7 & 9 \end{array}\right) \qquad B = \left(\begin{array}{cc|c|cc} 1 & 2 & 3 & 4 & 5 \\ 1 & 1 & 1 & 1 & 1 \\ \hline 9 & 8 & 7 & 6 & 5 \\ \hline 3 & 3 & 3 & 3 & 3 \\ 1 & 3 & 5 & 7 & 9 \end{array}\right)$$

Although A is a 5×5 square matrix and is a 3×3 block matrix, the second and third diagonal blocks are not square matrices. Thus A is not a square block matrix.

B is a square block matrix.

4.42. Complete the partitioning of $C = \left(\begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \\ 1 & 1 & 1 & 1 & 1 \\ \hline 9 & 8 & 7 & 6 & 5 \\ 3 & 3 & 3 & 3 & 3 \\ \hline 1 & 3 & 5 & 7 & 9 \end{array}\right)$ into a square block matrix.

One horizontal line is between the second and third rows; hence add a vertical line between the second and third columns. The other horizontal line is between the fourth and fifth rows; hence add a vertical line between the fourth and fifth columns. [The horizontal lines and the vertical lines must be symmetrically placed to obtain a square block matrix.] This yields the square block matrix

$$C = \left(\begin{array}{cc|cc|c} 1 & 2 & 3 & 4 & 5 \\ 1 & 1 & 1 & 1 & 1 \\ \hline 9 & 8 & 7 & 6 & 5 \\ 3 & 3 & 3 & 3 & 3 \\ \hline 1 & 3 & 5 & 7 & 9 \end{array}\right)$$

4.43. Determine which of the following square block matrices are lower triangular, upper triangular, or diagonal:

$$A = \left(\begin{array}{cc|c} 1 & 2 & 0 \\ 3 & 4 & 5 \\ \hline 0 & 0 & 6 \end{array}\right) \qquad B = \left(\begin{array}{c|cc|c} 1 & 0 & 0 & 0 \\ \hline 2 & 3 & 4 & 0 \\ 5 & 0 & 6 & 0 \\ \hline 0 & 7 & 8 & 9 \end{array}\right) \qquad C = \left(\begin{array}{c|cc} 1 & 0 & 0 \\ \hline 0 & 2 & 3 \\ 0 & 4 & 5 \end{array}\right) \qquad D = \left(\begin{array}{cc|c} 1 & 2 & 0 \\ 3 & 4 & 5 \\ \hline 0 & 6 & 7 \end{array}\right)$$

A is upper triangular since the block below the diagonal is a zero block.

B is lower triangular since all blocks above the diagonal are zero blocks.

C is diagonal since the blocks above and below the diagonal are zero blocks.

D is neither upper triangular nor lower triangular. Furthermore, no other partitioning of D will make it into either a block upper triangular matrix or a block lower triangular matrix.

4.44. Consider the following block diagonal matrices of which corresponding diagonal blocks have the same size:

$$M = \operatorname{diag}(A_1, A_2, \ldots, A_r) \qquad \text{and} \qquad N = \operatorname{diag}(B_1, B_2, \ldots, B_r)$$

Find: (*a*) $M + N$, (*b*) kM, (*c*) MN, (*d*) $f(M)$ for a given polynomial $f(x)$.

(*a*) Simply add the diagonal blocks: $M + N = \operatorname{diag}(A_1 + B_1, A_2 + B_2, \ldots, A_r + B_r)$.

(*b*) Simply multiply the diagonal blocks by k: $kM = \operatorname{diag}(kA_1, kA_2, \ldots, kA_r)$.

(c) Simply multiply corresponding diagonal blocks: $MN = \text{diag}(A_1B_1, A_2B_2, \ldots, A_rB_r)$.

(d) Find $f(A_i)$ for each diagonal block A_i. Then $f(M) = \text{diag}(f(A_1), f(A_2), \ldots, f(A_r))$.

4.45. Find M^2 where $M = \begin{pmatrix} 1 & 2 & & & \\ 3 & 4 & & & \\ & & 5 & & \\ & & & 1 & 3 \\ & & & 5 & 7 \end{pmatrix}$.

Since M is block diagonal, square each block:

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 7 & 10 \\ 15 & 22 \end{pmatrix}$$

$$(5)(5) = (25)$$

$$\begin{pmatrix} 1 & 3 \\ 5 & 7 \end{pmatrix}\begin{pmatrix} 1 & 3 \\ 5 & 7 \end{pmatrix} = \begin{pmatrix} 16 & 24 \\ 40 & 64 \end{pmatrix}$$

Then $M^2 = \begin{pmatrix} 7 & 10 & & & \\ 15 & 22 & & & \\ & & 25 & & \\ & & & 16 & 24 \\ & & & 40 & 64 \end{pmatrix}$.

CONGRUENT SYMMETRIC MATRICES, QUADRATIC FORMS

4.46. Let $A = \begin{pmatrix} 1 & -3 & 2 \\ -3 & 7 & -5 \\ 2 & -5 & 8 \end{pmatrix}$, a symmetric matrix. Find (a) a nonsingular matrix P such that P^TAP is diagonal, i.e., the diagonal matrix $B = P^TAP$, and (b) the signature of A.

(a) First form the block matrix $(A \mid I)$:

$$(A \mid I) = \left(\begin{array}{ccc|ccc} 1 & -3 & 2 & 1 & 0 & 0 \\ -3 & 7 & -5 & 0 & 1 & 0 \\ 2 & -5 & 8 & 0 & 0 & 1 \end{array}\right)$$

Apply the row operations $3R_1 + R_2 \to R_2$ and $-2R_1 + R_3 \to R_3$ to $(A \mid I)$ and then the corresponding column operations $3C_1 + C_2 \to C_2$ and $-2C_1 + C_3 \to C_3$ to A to obtain

$$\left(\begin{array}{ccc|ccc} 1 & -3 & 2 & 1 & 0 & 0 \\ 0 & -2 & 1 & 3 & 1 & 0 \\ 0 & 1 & 4 & -2 & 0 & 1 \end{array}\right) \quad \text{and then} \quad \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & -2 & 1 & 3 & 1 & 0 \\ 0 & 1 & 4 & -2 & 0 & 1 \end{array}\right)$$

Next apply the row operation $R_2 + 2R_3 \to R_3$ and then the corresponding column operation $C_2 + 2C_3 \to C_3$ to obtain

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & -2 & 1 & 3 & 1 & 0 \\ 0 & 0 & 9 & -1 & 1 & 2 \end{array}\right) \quad \text{and then} \quad \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & -2 & 0 & 3 & 1 & 0 \\ 0 & 0 & 18 & -1 & 1 & 2 \end{array}\right)$$

Now A has been diagonalized. Set

$$P = \begin{pmatrix} 1 & 3 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{pmatrix} \quad \text{and then} \quad B = P^TAP = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 18 \end{pmatrix}$$

(b) B has $\mathbf{p} = 2$ positive and $\mathbf{n} = 1$ negative diagonal elements. Hence sig $A = 2 - 1 = 1$.

QUADRATIC FORMS

4.47. Find the quadratic form $q(x, y)$ corresponding to the symmetric matrix $A = \begin{pmatrix} 5 & -3 \\ -3 & 8 \end{pmatrix}$.

$$q(x, y) = (x, y)\begin{pmatrix} 5 & -3 \\ -3 & 8 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = (5x - 3y, \; -3x + 8y)\begin{pmatrix} x \\ y \end{pmatrix}$$
$$= 5x^2 - 3xy - 3xy + 8y^2 = 5x^2 - 6xy + 8y^2$$

4.48. Find the symmetric matrix A which corresponds to the quadratic form

$$q(x, y, z) = 3x^2 + 4xy - y^2 + 8xz - 6yz + z^2$$

The symmetric matrix $A = (a_{ij})$ representing $q(x_1, \ldots, x_n)$ has the diagonal entry a_{ii} equal to the coefficient of x_i^2 and has the entries a_{ij} and a_{ji} each equal to half the coefficient of $x_i x_j$. Thus

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

4.49. Find the symmetric matrix B which corresponds to the quadratic forms

(a) $q(x, y) = 4x^2 + 5xy - 7y^2$ (b) $q(x, y, z) = 4xy + 5y^2$

(a) Here $B = \begin{pmatrix} 4 & \frac{5}{2} \\ \frac{5}{2} & -7 \end{pmatrix}$. (Division by 2 may introduce fractions even though the coefficients in q are integers.)

(b) Even though only x and y appears in the polynomial, the expression $q(x, y, z)$ indicates that there are three variables. In other words,

$$q(x, y, z) = 0x^2 + 4xy + 5y^2 + 0xz + 0yz + 0z^2$$

Thus

$$B = \begin{pmatrix} 0 & 2 & 0 \\ 2 & 5 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

4.50. Consider the quadratic form $q(x, y) = 3x^2 + 2xy - y^2$ and the linear substitution, $x = s - 3t$, $y = 2s + t$.

(a) Rewrite $q(x, y)$ in matrix notation, and find the matrix A representing the quadratic form.

(b) Rewrite the linear substitution using matrix notation, and find the matrix P corresponding to the substitution.

(c) Find $q(s, t)$ using direct substitution.

(d) Find $q(s, t)$ using matrix notation.

(a) Here $q(x, y) = (x, y)\begin{pmatrix} 3 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$. Hence $A = \begin{pmatrix} 3 & 1 \\ 1 & -1 \end{pmatrix}$ and $q(X) = X^TAX$ where $X = (x, y)^T$.

(b) We have $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & -3 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} s \\ t \end{pmatrix}$. Thus $P = \begin{pmatrix} 1 & -3 \\ 2 & 1 \end{pmatrix}$ and $X = PY$, where $X = (x, y)^T$ and $Y = (s, t)^T$.

(c) Substitute for x and y in q to obtain

$$q(s, t) = 3(s - 3t)^2 + 2(s - 3t)(2s + t) - (2s + t)^2$$
$$= 3(s^2 - 6st + 9t^2) + 2(2s^2 - 5st - 3t^2) - (s^2 + 4st + t^2) = 3s^2 - 32st + 20t^2$$

(d) Here $q(X) = X^TAX$ and $X = PY$. Thus $X^T = Y^TP^T$. Therefore,

$$q(s, t) = q(Y) = Y^TP^TAPY = (s, t)\begin{pmatrix} 1 & 2 \\ -3 & 1 \end{pmatrix}\begin{pmatrix} 3 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 1 & -3 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} s \\ t \end{pmatrix}$$
$$= (s, t)\begin{pmatrix} 3 & -16 \\ -16 & 20 \end{pmatrix}\begin{pmatrix} s \\ t \end{pmatrix} = 3s^2 - 32st + 20t^2$$

[As expected, the results in (c) and (d) are equal.]

4.51. Let L be a linear substitution $X = PY$, as in Problem 4.50. (a) When is L nonsingular? orthogonal? (b) Describe one main advantage of an orthogonal substitution over a nonsingular substitution. (c) Is the linear substitution in Problem 4.50 nonsingular? orthogonal?

(a) L is said to be nonsingular or orthogonal according as the matrix P representing the substitution is nonsingular or orthogonal.

(b) Recall that the columns of the matrix P representing the linear substitution introduces a new coordinate system. If P is orthogonal, then the new axes are perpendicular and have the same unit lengths as the original axes.

(c) The matrix $P = \begin{pmatrix} 1 & -3 \\ 2 & 1 \end{pmatrix}$ is nonsingular, but not orthogonal; hence the linear substitution is nonsingular, but not orthogonal.

4.52. Let $q(x, y, z) = x^2 + 4xy + 3y^2 - 8xz - 12yz + 9z^2$. Find a nonsingular linear substitution expressing the variables x, y, z in terms of the variables r, s, t so that $q(r, s, t)$ is diagonal. Also find the signature of q.

Form the block matrix $(A \mid I)$ where A is the matrix which corresponds to the quadratic form:

$$(A \mid I) = \left(\begin{array}{ccc|ccc} 1 & 2 & -4 & 1 & 0 & 0 \\ 2 & 3 & -6 & 0 & 1 & 0 \\ -4 & -6 & 9 & 0 & 0 & 1 \end{array}\right)$$

Apply $-2R_1 + R_2 \to R_2$ and $4R_1 + R_3 \to R_3$ and the corresponding column operations, and then $2R_2 + R_3 \to R_3$ and the corresponding column operation to obtain

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & -1 & 2 & -2 & 1 & 0 \\ 0 & 2 & -7 & 4 & 0 & 1 \end{array}\right) \quad \text{and then} \quad \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & -2 & 1 & 0 \\ 0 & 0 & -3 & 0 & 2 & 1 \end{array}\right)$$

Thus the linear substitution $x = r - 2s$, $y = s + 2t$, $z = t$ will yield the quadratic form

$$q(r, s, t) = r^2 - s^2 - 3t^2$$

By inspection, sig $q = 1 - 2 = -1$.

4.53. Diagonalize the following quadratic form q by the method known as "completing the square":

$$q(x, y) = 2x^2 - 12xy + 5y^2$$

First factor out the coefficient of x^2 from the x^2 term and the xy term to get

$$q(x, y) = 2(x^2 - 6xy \qquad) + 5y^2$$

Next complete the square inside the parentheses by adding an appropriate multiple of y^2 and then subtract the corresponding amount outside the parentheses to get

$$q(x, y) = 2(x^2 - 6xy + 9y^2) + 5y^2 - 18y^2 = 2(x - 3y)^2 - 13y^2$$

(The -18 comes from the fact that the $9y^2$ inside the parentheses is multiplied by 2.) Let $s = x - 3y$, $t = y$. Then $x = s + 3t$, $y = t$. This linear substitution yields the quadratic form $q(s, t) = 2s^2 - 13t^2$.

POSITIVE DEFINITE QUADRATIC FORMS

4.54. Let $q(x, y, z) = x^2 + 2y^2 - 4xz - 4yz + 7z^2$. Is q positive definite?

Diagonalize (under congruence) the symmetric matrix A corresponding to q (by applying $2R_1 + R_3 \to R_3$ and $2C_1 + C_3 \to C_3$, and then $R_2 + R_3 \to R_3$ and $C_2 + C_3 \to C_3$):

$$A = \begin{pmatrix} 1 & 0 & -2 \\ 0 & 2 & -2 \\ -2 & -2 & 7 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & -2 \\ 0 & -2 & 3 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

The diagonal representation of q only contains positive entries, 1, 2, and 1, on the diagonal; hence q is positive definite.

4.55. Let $q(x, y, z) = x^2 + y^2 + 2xz + 4yz + 3z^2$. Is q positive definite?

Diagonalize (under congruence) the symmetric matrix A corresponding to q:

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 1 & 2 & 3 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 2 & 2 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}$$

There is a negative entry -2 in the diagonal representation of q; hence q is not positive definite.

4.56. Show that $q(x, y) = ax^2 + bxy + cy^2$ is positive definite if and only if $a > 0$ and the discriminant $D = b^2 - 4ac < 0$.

Suppose $v = (x, y) \neq 0$, say $y \neq 0$. Let $t = x/y$. Then

$$q(v) = y^2[a(x/y)^2 + b(x/y) + c] = y^2(at^2 + bt + c)$$

However, $s = at^2 + bt + c$ lies above the t axis, i.e., is positive for every value of t if and only if $a > 0$ and $D = b^2 - 4ac < 0$. Thus q is positive definite if and only if $a > 0$ and $D < 0$.

4.57. Determine which quadratic form q is positive definite:

$$(a)\quad q(x, y) = x^2 - 4xy + 5y^2 \qquad (b)\quad q(x, y) = x^2 + 6xy + 3y^2$$

(*a*) **Method 1.** Diagonalize by completing the square:

$$q(x, y) = x^2 - 4xy + 4y^2 + 5y^2 - 4y^2 = (x - 2y)^2 + y^2 = s^2 + t^2$$

where $s = x - 2y$, $t = y$. Thus q is positive definite.

Method 2. Compute the discriminant $D = b^2 - 4ac = 16 - 20 = -4$. Since $D < 0$, q is positive definite.

(*b*) **Method 1.** Diagonalize by completing the square:

$$q(x, y) = x^2 + 6xy + 9y^2 + 3y^2 - 9y^2 = (x + 3y)^2 - 6y^2 = s^2 - 6t^2$$

where $s = x + 3y$, $t = y$. Since -6 is negative, q is not positive definite.

Method 2. Compute $D = b^2 - 4ac = 36 - 12 = 24$. Since $D > 0$, q is not positive definite.

4.58. Let B be any nonsingular matrix, and let $M = B^T B$. Show that (*a*) M is symmetric, and (*b*) M is positive definite.

(*a*) $M^T = (B^T B)^T = B^T B^{TT} = B^T B = M$; hence M is symmetric.

(*b*) Since B is nonsingular, $BX \neq 0$ for any nonzero $X \in \mathbf{R}^n$. Hence the dot product of BX with itself, $BX \cdot BX = (BX)^T(BX)$, is positive. Thus

$$q(X) = X^T M X = X^T(B^T B)X = (X^T B^T)(BX) = (BX)^T(BX) > 0$$

Thus M is positive definite.

4.59. Show that $q(X) = \| X \|^2$, the square of the norm of a vector X, is a positive definite quadratic form.

For $X = (x_1, x_2, \ldots, x_n)$, we have $q(X) = x_1^2 + x_2^2 + \cdots + x_n^2$. Now q is a polynomial with each term of degree two, and q is in diagonal form where all diagonal entries are positive. Thus q is a positive definite quadratic form.

4.60. Prove that the following two definitions of a positive definite quadratic form are equivalent:

(*a*) The diagonal entries are all positive in any diagonal representation of q.

(*b*) $q(Y) > 0$, for any nonzero vector Y in $\mathbf{R}^n$.

Suppose $q(Y) = a_1 y_1^2 + a_2 y_2^2 + \cdots + a_n y_n^2$. If all the coefficients a_i are positive, then clearly $q(Y) > 0$ for any nonzero vector Y. Thus (*a*) implies (*b*). Conversely, suppose $a_k \leq 0$. Let $e_k = (0, \ldots, 1, \ldots, 0)$ be the vector whose entries are all 0 except 1 in the kth position. Then $q(e_k) = a_k \leq 0$ for $e_k \neq 0$. Thus not = (*a*) implies not = (*b*). Accordingly, (*a*) and (*b*) are equivalent.

SIMILARITY OF MATRICES

4.61. Consider the cartesian plane $\mathbf{R}^2$ with the usual x and y axes. The 2×2 nonsingular matrix

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

determines a new coordinate system of the plane, say with s and t axes. (See Example 4.16.)

(*a*) Plot the new s and t axis in the plane $\mathbf{R}^2$.

(*b*) Find the coordinates of $Q(1, 5)$ in the new system.

(*a*) Plot the s axis in the direction of the first column $u_1 = (1, -1)^T$ of P with unit length equal to the length u_1. Similarly, plot the t axis in the direction of the second column $u_2 = (3, 2)^T$ of P with unit length equal to the length of u_2. See Fig. 4-2.

(*b*) Find $P^{-1} = \begin{pmatrix} \frac{2}{5} & -\frac{3}{5} \\ \frac{1}{5} & \frac{1}{5} \end{pmatrix}$, say by using the formula for the inverse of a 2×2 matrix. Then multiply the coordinate (column) vector of Q by P^{-1}:

$$P^{-1}Q = \begin{pmatrix} \frac{2}{5} & -\frac{3}{5} \\ \frac{1}{5} & \frac{1}{5} \end{pmatrix}\begin{pmatrix} 1 \\ 5 \end{pmatrix} = \begin{pmatrix} -\frac{13}{5} \\ \frac{6}{5} \end{pmatrix}$$

Thus $Q'(-\frac{13}{5}, \frac{6}{5})$ represents Q in the new system.

4.62. Let $f: \mathbf{R}^2 \to \mathbf{R}^2$ be defined by $f(x, y) = (2x - 5y, 3x + 4y)$.

(*a*) Using $X = (x, y)^T$, write f in matrix notation, i.e., find the matrix A such that $f(X) = AX$.

(*b*) Referring to the new coordinate s and t axes of $\mathbf{R}^2$ introduced in Problem 4.61, and using $Y = (s, t)^T$, find $f(s, t)$ by first finding the matrix B such that $f(Y) = BY$.

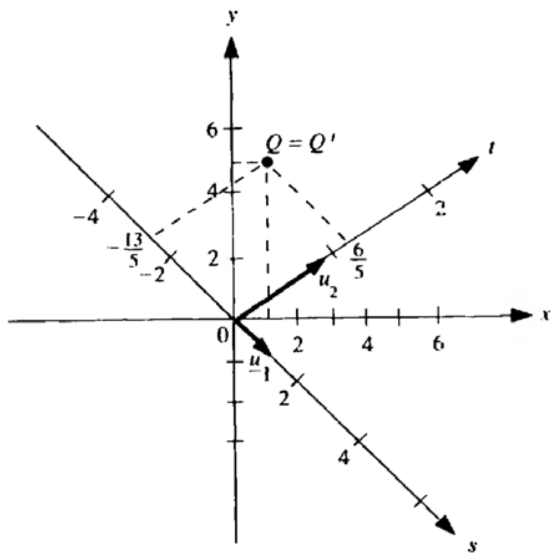

Fig. 4-2

(*a*) Here $f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 & -5 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$; hence $A = \begin{pmatrix} 2 & -5 \\ 3 & 4 \end{pmatrix}$.

(*b*) Find $B = P^{-1}AP = \begin{pmatrix} \frac{2}{5} & -\frac{3}{5} \\ \frac{1}{5} & \frac{1}{5} \end{pmatrix}\begin{pmatrix} 2 & -5 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 1 & 3 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} \frac{17}{5} & -\frac{59}{5} \\ \frac{6}{5} & \frac{13}{5} \end{pmatrix}$. Then

$$f\begin{pmatrix} s \\ t \end{pmatrix} = \begin{pmatrix} \frac{17}{5} & -\frac{59}{5} \\ \frac{6}{5} & \frac{13}{5} \end{pmatrix}\begin{pmatrix} s \\ t \end{pmatrix}$$

Thus $f(s, t) = (\frac{17}{5}s - \frac{59}{5}t, \frac{6}{5}s + \frac{13}{5}t)$.

4.63. Consider the space $\mathbf{R}^3$ with the usual x, y, z axes. The 3×3 nonsingular matrix

$$P = \begin{pmatrix} 1 & 3 & -2 \\ -2 & -5 & 2 \\ 1 & 2 & 1 \end{pmatrix}$$

determines a new coordinate system for $\mathbf{R}^3$, say with r, s, t axes. [Alternatively, P defines the linear substitution $X = PY$, where $X = (x, y, z)^T$ and $Y = (r, s, t)^T$.] Find the coordinates of the point $Q(1, 2, 3)$ in the new system.

First find P^{-1}. Form the block matrix $M = (P \mid I)$ and reduce M to row canonical form:

$$M = \left(\begin{array}{ccc|ccc} 1 & 3 & -2 & 1 & 0 & 0 \\ -2 & -5 & 2 & 0 & 1 & 0 \\ 1 & 2 & 1 & 0 & 0 & 1 \end{array}\right) \sim \left(\begin{array}{ccc|ccc} 1 & 3 & -2 & 1 & 0 & 0 \\ 0 & 1 & -2 & 2 & 1 & 0 \\ 0 & -1 & 3 & -1 & 0 & 1 \end{array}\right)$$

$$\sim \left(\begin{array}{ccc|ccc} 1 & 3 & -2 & 1 & 0 & 0 \\ 0 & 1 & -2 & 2 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 \end{array}\right) \sim \left(\begin{array}{ccc|ccc} 1 & 3 & 0 & 3 & 2 & 2 \\ 0 & 1 & 0 & 4 & 3 & 2 \\ 0 & 0 & 1 & 1 & 1 & 1 \end{array}\right)$$

$$\sim \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -9 & -7 & -4 \\ 0 & 1 & 0 & 4 & 3 & 2 \\ 0 & 0 & 1 & 1 & 1 & 1 \end{array}\right)$$

Accordingly,

$$P^{-1} = \begin{pmatrix} -9 & -7 & -4 \\ 4 & 3 & 2 \\ 1 & 1 & 1 \end{pmatrix} \quad \text{and} \quad P^{-1}Q = \begin{pmatrix} -9 & -7 & -4 \\ 4 & 3 & 2 \\ 1 & 1 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} -47 \\ 16 \\ 6 \end{pmatrix}$$

Thus $Q'(-47, 16, 6)$ represents Q in the new system.

4.64. Let $f: \mathbf{R}^3 \to \mathbf{R}^3$ be defined by

$$f(x, y, z) = (x + 2y - 3z, 2x + z, x - 3y + z)$$

and let P be the nonsingular change-of-variable matrix in Problem 4.63. [Thus, $X = PY$ where $X = (x, y, z)^T$ and $Y = (r, s, t)^T$.] Find: (*a*) the matrix A such that $f(X) = AX$, (*b*) the matrix B such that $f(Y) = BY$, and (*c*) $f(r, s, t)$.

(*a*) The coefficients of x, y and z give the matrix A:

$$f\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 & 2 & -3 \\ 2 & 0 & 1 \\ 1 & -3 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} \quad \text{and so} \quad A = \begin{pmatrix} 1 & 2 & -3 \\ 2 & 0 & 1 \\ 1 & -3 & 1 \end{pmatrix}$$

(*b*) Here B is similar to A with respect to P, that is,

$$B = P^{-1}AP = \begin{pmatrix} -9 & -7 & -4 \\ 4 & 3 & 2 \\ 1 & 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 & -3 \\ 2 & 0 & 1 \\ 1 & -3 & 1 \end{pmatrix}\begin{pmatrix} 1 & 3 & -2 \\ -2 & -5 & 2 \\ 1 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -19 & 58 \\ 1 & 12 & -27 \\ 5 & 15 & -11 \end{pmatrix}$$

(*c*) Use the matrix B to obtain

$$f(r, s, t) = (r - 19s + 58t, r + 12s - 27t, 5r + 15s - 11t)$$

4.65. Suppose B is similar to A. Prove $\operatorname{tr} B = \operatorname{tr} A$.

Since B is similar to A, there exists a nonsingular matrix P such that $B = P^{-1}AP$. Then, using Theorem 4.1,

$$\operatorname{tr} B = \operatorname{tr} P^{-1}AP = \operatorname{tr} PP^{-1}A = \operatorname{tr} A$$

LU FACTORIZATION

4.66. Find the LU factorization of $A = \begin{pmatrix} 1 & 3 & 2 \\ 2 & 5 & 6 \\ -3 & -2 & 7 \end{pmatrix}$.

Reduce A to triangular form by the operations $-2R_1 + R_2 \to R_2$ and $3R_1 + R_3 \to R_3$, and then $7R_2 + R_3 \to R_3$:

$$A \sim \begin{pmatrix} 1 & 3 & 2 \\ 0 & -1 & 2 \\ 0 & 7 & 13 \end{pmatrix} \sim \begin{pmatrix} 1 & 3 & 2 \\ 0 & -1 & 2 \\ 0 & 0 & 27 \end{pmatrix}$$

Use the negatives of the multipliers -2, 3, and 7 in the above row operations to form the matrix L, and use the triangular form of A to obtain the matrix U; that is,

$$L = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -3 & -7 & 1 \end{pmatrix} \quad \text{and} \quad U = \begin{pmatrix} 1 & 3 & 2 \\ 0 & -1 & 2 \\ 0 & 0 & 27 \end{pmatrix}$$

(As a check, multiply L and U to verify that $A = LU$.)

4.67. Find the LDU factorization of the matrix A in Problem 4.66.

The $A = LDU$ factorization refers to the situation where L is a lower triangular matrix with 1s on the diagonal (as in the LU factorization of A), D is a diagonal matrix, and U is an upper triangular matrix with 1s on the diagonal. Thus simply factor out the diagonal entries in the matrix U in the above LU factorization of A to obtain the matrices D and L. Hence

$$L = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -3 & -7 & 1 \end{pmatrix} \qquad D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & \frac{1}{27} \end{pmatrix} \qquad U = \begin{pmatrix} 1 & 3 & 2 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix}$$

4.68. Find the LU factorization of $B = \begin{pmatrix} 1 & 4 & -3 \\ 2 & 8 & 1 \\ -5 & -9 & 7 \end{pmatrix}$.

Reduce B to triangular form by first applying the operations $-2R_1 + R_2 \to R_2$ and $5R_1 + R_3 \to R_3$:

$$B \sim \begin{pmatrix} 1 & 4 & -3 \\ 0 & 0 & 7 \\ 0 & 11 & -8 \end{pmatrix}$$

Observe that the second diagonal entry is 0. Thus B cannot be brought into triangular form without row interchange operations. In other words, B is not LU-factorable.

4.69. Find the LU factorization of $A = \begin{pmatrix} 1 & 2 & -3 & 4 \\ 2 & 3 & -8 & 5 \\ 1 & 3 & 1 & 3 \\ 3 & 8 & -1 & 13 \end{pmatrix}$ by a direct method.

First form the following matrices L and U:

$$L = \begin{pmatrix} 1 & 0 & 0 & 0 \\ l_{21} & 1 & 0 & 0 \\ l_{31} & l_{32} & 1 & 0 \\ l_{41} & l_{42} & l_{43} & 1 \end{pmatrix} \qquad \text{and} \qquad U = \begin{pmatrix} u_{11} & u_{12} & u_{13} & u_{14} \\ 0 & u_{22} & u_{23} & u_{24} \\ 0 & 0 & u_{33} & u_{34} \\ 0 & 0 & 0 & u_{44} \end{pmatrix}$$

That part of the product LU which determines the first row of A yields the four equations

$$u_{11} = 1 \qquad u_{12} = 2 \qquad u_{13} = -3 \qquad u_{14} = 4$$

and that part of the product LU which determines the first column of A yields the equations

$$l_{21}u_{11} = 2, \qquad l_{31}u_{11} = 1, \qquad l_{41}u_{11} = 3 \qquad \text{or} \qquad l_{21} = 2, \qquad l_{31} = 1, \qquad l_{41} = 3$$

Thus, at this point, the matrices L and U have the form

$$L = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 \\ 1 & l_{32} & 1 & 0 \\ 3 & l_{42} & l_{43} & 1 \end{pmatrix} \qquad \text{and} \qquad U = \begin{pmatrix} 1 & 2 & -3 & 4 \\ 0 & u_{22} & u_{23} & u_{24} \\ 0 & 0 & u_{33} & u_{34} \\ 0 & 0 & 0 & u_{44} \end{pmatrix}$$

That part of the product LU which determines the remaining entries in the second row of A yields the equations

$$4 + u_{22} = 3 \qquad -6 + u_{23} = -8 \qquad 8 + u_{24} = 5$$

or

$$u_{22} = -1 \qquad u_{23} = -2 \qquad u_{24} = -3$$

and that part of the product LU which determines the remaining entries in the second column of A yields the equations

$$2 + l_{32}u_{22} = 3, \qquad 6 + l_{42}u_{22} = 8 \qquad \text{or} \qquad l_{32} = -1, \qquad l_{42} = -2$$

Thus L and U now have the form

$$L = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 1 & -1 & 1 & 0 \\ 3 & -2 & l_{43} & 1 \end{pmatrix} \quad \text{and} \quad U = \begin{pmatrix} 1 & 2 & -3 & 4 \\ 0 & -1 & -2 & -3 \\ 0 & 0 & u_{33} & u_{34} \\ 0 & 0 & 0 & u_{44} \end{pmatrix}$$

Continuing, using the third row, third column, and fourth row of A, we get

$$u_{33} = 2, \qquad u_{34} = -1, \qquad \text{then} \qquad l_{43} = 2, \qquad \text{and lastly} \qquad u_{44} = 3$$

Thus

$$L = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 1 & -1 & 1 & 0 \\ 3 & -2 & 2 & 1 \end{pmatrix} \quad \text{and} \quad U = \begin{pmatrix} 1 & 2 & -3 & 4 \\ 0 & -1 & -2 & -3 \\ 0 & 0 & 2 & -1 \\ 0 & 0 & 0 & 3 \end{pmatrix}$$

4.70. Find the LDU factorization of matrix A in the Problem 4.69.

Here U should have 1s on the diagonal and D is a diagonal matrix. Thus, using the above LU factorization of A, factor out the diagonal entries in that U to obtain

$$D = \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & 2 & \\ & & & 3 \end{pmatrix} \quad \text{and} \quad U = \begin{pmatrix} 1 & 2 & -3 & 4 \\ & 1 & 2 & 3 \\ & & 1 & -2 \\ & & & 1 \end{pmatrix}$$

The matrix L is the same as in Problem 4.69.

4.71. Given the factorization $A = LU$, where $L = (l_{ij})$ and $U = (u_{ij})$. Consider the system $AX = B$. Determine (*a*) the algorithm to find $L^{-1}B$, and (*b*) the algorithm that solves $UX = B$ by back-substitution.

(*a*) The entry l_{ij} in the matrix L corresponds to the elementary row operation $-l_{ij}R_i + R_j \rightarrow R_j$. Thus the algorithm which transforms B into B' is as follows:

***Algorithm** P4.71A: Evaluating $L^{-1}B$*

Step 1. Repeat for $j = 1$ to $n - 1$:

Step 2. Repeat for $i = j + 1$ to n:
$b_j := -l_{ij}b_i + b_j$
[End of Step 2 inner loop.]
[End of Step 1 outer loop.]

Step 3. Exit.

[The complexity of this algorithm is $C(n) \approx n^2/2$.]

(*b*) The back-substitution algorithm follows:

***Algorithm** P4.71B: Back-substitution for system $UX = B$*

Step 1. $x_n = b_n/u_{nn}$

Step 2. Repeat for $j = n - 1, n - 2, \ldots, 1$
$x_j = (b_j - u_{j,j+1}x_{j+1} - \cdots - u_{jn}x_n)/u_{jj}$

Step 3. Exit.

[The complexity here is also $C(n) \approx n^2/2$.]

4.72. Find the LU factorization of the matrix $A = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 3 & 3 \\ -3 & -10 & 2 \end{pmatrix}$.

Reduce A to triangular form by the operations

$$(1)\quad -2R_1 + R_2 \to R_2, \qquad (2)\quad 3R_1 + R_3 \to R_3, \qquad (3)\quad -4R_2 + R_3 \to R_3$$

$$A \sim \begin{pmatrix} 1 & 2 & 1 \\ 0 & -1 & 1 \\ 0 & -4 & 5 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & 1 \\ 0 & -1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

Thus
$$L = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -3 & 4 & 1 \end{pmatrix} \quad \text{and} \quad U = \begin{pmatrix} 1 & 2 & 1 \\ 0 & -1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

The entries 2, -3, and 4 in L are the negatives of the multipliers in the above row operations.

4.73. Solve the system $AX = B$ for B_1, B_2, B_3, where A is the matrix in Problem 4.72 and where $B_1 = (1, 1, 1)$, $B_2 = B_1 + X_1$, $B_3 = B_2 + X_2$ (here X_j is the solution when $B = B_j$).

(*a*) Find $L^{-1}B_1$ or, equivalently, apply the row operations (1), (2), and (3) to B_1 to yield

$$B_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \xrightarrow{(1)\text{ and }(2)} \begin{pmatrix} 1 \\ -1 \\ 4 \end{pmatrix} \xrightarrow{(3)} \begin{pmatrix} 1 \\ -1 \\ 8 \end{pmatrix}$$

Solve $UX = B$ for $B = (1, -1, 8)$ by back-substitution to obtain $X_1 = (-25, 9, 8)$.

(*b*) Find $B_2 = B_1 + X_1 = (1, 1, 1) + (-25, 9, 8) = (-24, 10, 9)$. Apply the operations (1), (2), and (3) to B_2 to obtain $(-24, 58, -63)$, and then $B = (-24, 58, -295)$.
Solve $UX = B$ by back-substitution to obtain $X_2 = (943, -353, -295)$.

(*c*) Find $B_3 = B_2 + X_2 = (-24, 10, 9) + (943, -353, -295) = (919, -343, -286)$. Apply the operations (1), (2), and (3) to B_3 to obtain $(919, -2181, 2671)$, and then $B = (919, -2181, 11395)$.
Solve $UX = B$ by back-substitution to obtain $X_3 = (-37{,}628, 13{,}576, 11{,}395)$.

Supplementary Problems

ALGEBRA OF MATRICES

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

4.75. Suppose the 2×2 matrix B commutes with every 2×2 matrix A. Show that $B = \begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix}$ for some scalar k, i.e., B is a scalar matrix.

4.76. Let $A = \begin{pmatrix} 5 & 2 \\ 0 & k \end{pmatrix}$. Find all numbers k for which A is a root of the polynomial

(*a*) $f(x) = x^2 - 7x + 10$, (*b*) $g(x) = x^2 - 25$, (*c*) $h(x) = x^2 - 4$

4.77. Let $B = \begin{pmatrix} 1 & 0 \\ 26 & 27 \end{pmatrix}$. Find a matrix A such that $A^3 = B$.

4.78. Let $A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$ Find: (*a*) A^n and (*b*) B^n, for all positive integers n.

4.79. Find conditions on matrices A and B so that $A^2 - B^2 = (A + B)(A - B)$.

INVERTIBLE MATRICES, INVERSES, ELEMENTARY MATRICES

4.80. Find the inverse of each matrix: (*a*) $\begin{pmatrix} 1 & 3 & -2 \\ 2 & 8 & -3 \\ 1 & 7 & 1 \end{pmatrix}$, (*b*) $\begin{pmatrix} 2 & 1 & -1 \\ 5 & 2 & -3 \\ 0 & 2 & 1 \end{pmatrix}$, (*c*) $\begin{pmatrix} 1 & -2 & 0 \\ 2 & -3 & 1 \\ 1 & 1 & 5 \end{pmatrix}$.

4.81. Find the inverse of each matrix: (*a*) $\begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$, (*b*) $\begin{pmatrix} 1 & 2 & 1 & 0 \\ 0 & 1 & -1 & 1 \\ 1 & 3 & 1 & -2 \\ 1 & 4 & -2 & 4 \end{pmatrix}$.

4.82. Express each matrix as a product of elementary matrices: (*a*) $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$, (*b*) $\begin{pmatrix} 3 & -6 \\ -2 & 4 \end{pmatrix}$.

4.83. Express $A = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 3 \\ 3 & 8 & 7 \end{pmatrix}$ as a product of elementary matrices.

4.84. Suppose A is invertible. Show that if $AB = AC$ then $B = C$. Give an example of a nonzero matrix A such that $AB = AC$ but $B \neq C$.

4.85. If A is invertible, show that kA is invertible when $k \neq 0$, with inverse $k^{-1}A^{-1}$.

4.86. Suppose A and B are invertible and $A + B \neq 0$. Show, by an example, that $A + B$ need not be invertible.

SPECIAL TYPES OF SQUARE MATRICES

4.87. Using only the elements 0 and 1, find all 3×3 nonsingular upper triangular matrices.

4.88. Using only the elements 0 and 1, find the number of: (*a*) 4×4 diagonal matrices, (*b*) 4×4 upper triangular matrices, (*c*) 4×4 nonsingular upper triangular matrices. Generalize to $n \times n$ matrices.

4.89. Find all real matrices A such that $A^2 = B$ where (*a*) $B = \begin{pmatrix} 4 & 21 \\ 0 & 25 \end{pmatrix}$, (*b*) $B = \begin{pmatrix} 1 & 4 \\ 0 & -9 \end{pmatrix}$.

4.90. Let $B = \begin{pmatrix} 1 & 8 & 5 \\ 0 & 9 & 5 \\ 0 & 0 & 4 \end{pmatrix}$. Find a matrix A with positive diagonal entries such that $A^2 = B$.

4.91. Suppose $AB = C$ where A and C are upper triangular.

(*a*) Show, by an example, that B need not be upper triangular even when A and C are nonzero matrices.

(*b*) Show that B is upper triangular when A is invertible.

4.92. Show that AB need not be symmetric, even though A and B are symmetric.

4.93. Let A and B be symmetric matrices. Show that AB is symmetric if and only if A and B commute.

4.94. Suppose A is a symmetric matrix. Show that (*a*) A^2 and, in general, A^n is symmetric; (*b*) $f(A)$ is symmetric for any polynomial $f(x)$; (*c*) P^TAP is symmetric.

4.95. Find a 2×2 orthogonal matrix P whose first row is (*a*) $(2/\sqrt{29}, 5/\sqrt{29})$, (*b*) a multiple of $(3, 4)$.

4.96. Find a 3×3 orthogonal matrix P whose first two rows are multiples of (*a*) $(1, 2, 3)$ and $(0, -2, 3)$, respectively; (*b*) $(1, 3, 1)$ and $(1, 0, -1)$, respectively.

4.97. Suppose A and B are orthogonal. Show that A^T, A^{-1}, and AB are also orthogonal.

4.98. Which of the following matrices are normal?

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

4.99. Suppose A is a normal matrix. Show that: (*a*) A^T, (*b*) A^2 and, in general A^n, (*c*) $B = kI + A$ are also normal.

4.100. A matrix E is *idempotent* if $E^2 = E$. Show that $E = \begin{pmatrix} 2 & -2 & -4 \\ -1 & 3 & 4 \\ 1 & -2 & -3 \end{pmatrix}$ is idempotent.

4.101. Show that if $AB = A$ and $BA = B$, then A and B are idempotent.

4.102. A matrix A is *nilpotent of class p* if $A^p = 0$ but $A^{p-1} \neq 0$. Show that $A = \begin{pmatrix} 1 & 1 & 3 \\ 5 & 2 & 6 \\ -2 & -1 & -3 \end{pmatrix}$ is nilpotent of class 3.

4.103. Suppose A is nilpotent of class p. Show that $A^q = 0$ for $q > p$ but $A^q \neq 0$ for $q < p$.

4.104. A square matrix is *tridiagonal* if the nonzero entries occur only on the diagonal directly above the main diagonal (on the superdiagonal), or directly below the main diagonal (on the subdiagonal). Display the generic tridiagonal matrices of orders 4 and 5.

4.105. Show that the product of tridiagonal matrices need not be tridiagonal.

COMPLEX MATRICES

4.106. Find real numbers x, y, and z so that A is Hermitian, where

$$(a)\quad A = \begin{pmatrix} x + yi & 3 \\ 3 + zi & 0 \end{pmatrix}, \quad (b) = \begin{pmatrix} 3 & x + 2i & yi \\ 3 - 2i & 0 & 1 + zi \\ yi & 1 - xi & -1 \end{pmatrix}$$

4.107. Suppose A is any complex matrix. Show that AA^H and A^HA are both Hermitian.

4.108. Suppose A is any complex square matrix. Show that $A + A^H$ is Hermitian and $A - A^H$ is skew-Hermitian.

4.109. Which of the following matrices are unitary?

$$A = \begin{pmatrix} i/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & -i/2 \end{pmatrix}, \quad B = \frac{1}{2}\begin{pmatrix} 1+i & 1-i \\ 1-i & 1+i \end{pmatrix}, \quad C = \frac{1}{2}\begin{pmatrix} 1 & -i & -1+i \\ i & 1 & 1+i \\ 1+i & -1+i & 0 \end{pmatrix}$$

4.110. Suppose A and B are unitary matrices. Show that: (*a*) A^H is unitary, (*b*) A^{-1} is unitary, (*c*) AB is unitary.

4.111. Determine which of the following matrices are normal: $A = \begin{pmatrix} 3+4i & 1 \\ i & 2+3i \end{pmatrix}$, $B = \begin{pmatrix} 1 & 0 \\ 1-i & i \end{pmatrix}$.

4.112. Suppose A is a normal matrix and U is a unitary matrix. Show that $B = U^H AU$ is also normal.

4.113. Recall the following elementary row operations:

$$[E_1] \quad R_i \leftrightarrow R_j, \quad [E_2] \quad kR_i \to R_i, \quad k \neq 0, \quad [E_3] \quad kR_j + R_i \to R_i$$

For complex matrices, the respective corresponding Hermitian column operations are as follows:

$$[G_1] \quad C_i \leftrightarrow C_j, \quad [G_2] \quad \bar{k}C_i \to C_i, \quad k \neq 0, \quad [G_3] \quad \bar{k}C_j + C_i \to C_i$$

Show that the elementary matrix corresponding to $[G_i]$ is the conjugate transpose of the elementary matrix corresponding to $[E_i]$.

SQUARE BLOCK MATRICES

4.114. Using vertical lines, complete the partitioning of each matrix so that it is a square block matrix:

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ \hline 1 & 1 & 1 & 1 & 1 \\ 9 & 8 & 7 & 6 & 5 \\ \hline 2 & 2 & 2 & 2 & 2 \\ 3 & 3 & 3 & 3 & 3 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 1 & 1 & 1 & 1 \\ \hline 9 & 8 & 7 & 6 & 5 \\ \hline 2 & 2 & 2 & 2 & 2 \\ 3 & 3 & 3 & 3 & 3 \end{pmatrix}.$$

4.115. Partition each of the following matrices so that it becomes a block diagonal matrix with as many diagonal blocks as possible:

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 3 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 2 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 4 & 0 & 0 \\ 0 & 0 & 5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 6 \end{pmatrix}, \qquad C = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 2 & 0 \end{pmatrix}$$

4.116. Find M^2 and M^3 for each matrix M:

$$(a) \quad M = \left(\begin{array}{c|cc|c} 2 & 0 & 0 & 0 \\ \hline 0 & 1 & 4 & 0 \\ 0 & 2 & 1 & 0 \\ \hline 0 & 0 & 0 & 3 \end{array}\right), \qquad (b) \quad M = \left(\begin{array}{cc|cc} 1 & 1 & 0 & 0 \\ 2 & 3 & 0 & 0 \\ \hline 0 & 0 & 1 & 2 \\ 0 & 0 & 4 & 5 \end{array}\right).$$

4.117. Let $M = \operatorname{diag}(A_1, \ldots, A_k)$ and $N = \operatorname{diag}(B_1, \ldots, B_k)$ be block diagonal matrices where each pair of blocks A_i, B_i have the same size. Prove MN is block diagonal and

$$MN = \operatorname{diag}(A_1 B_1, A_2 B_2, \ldots, A_k B_k)$$

REAL SYMMETRIC MATRICES AND QUADRATIC FORMS

4.118. Let $A = \begin{pmatrix} 1 & 1 & -2 & -3 \\ 1 & 2 & -5 & -1 \\ -2 & -5 & 6 & 9 \\ -3 & -1 & 9 & 11 \end{pmatrix}$. Find a nonsingular matrix P such that $B = P^T AP$ is diagonal. Also, find B and sig A.

4.119. For each quadratic form $q(x, y, z)$, find a nonsingular linear substitution expressing the variables x, y, z in terms of variables r, s, t such that $q(r, s, t)$ is diagonal.

(a) $q(x, y, z) = x^2 + 6xy + 8y^2 - 4xz + 2yz - 9z^2$

(b) $q(x, y, z) = 2x^2 - 3y^2 + 8xz + 12yz + 25z^2$

4.120. Find those values of k so that the given quadratic form is positive definite:

(a) $q(x, y) = 2x^2 - 5xy + ky^2$

(b) $q(x, y) = 3x^2 - kxy + 12y^2$

(c) $q(x, y, z) = x^2 + 2xy + 2y^2 + 2xz + 6yz + kz^2$

4.121. Give an example of a quadratic form $q(x, y)$ such that $q(u) = 0$ and $q(v) = 0$ but $q(u + v) \neq 0$.

4.122. Show that any real symmetric matrix A is congruent to a diagonal matrix with only 1s, -1s, and 0s on the diagonal.

4.123. Show that congruence of matrices is an equivalence relation.

SIMILARITY OF MATRICES

4.124. Consider the space $\mathbf{R}^3$ with the usual x, y, z axes. The nonsingular matrix $P = \begin{pmatrix} 1 & -2 & -2 \\ 2 & -3 & -6 \\ 1 & 1 & -7 \end{pmatrix}$ determines a new coordinate system for $\mathbf{R}^3$, say with r, s, t axes. Find:

(a) The coordinates of the point $Q(1, 1, 1)$ in the new system,

(b) $f(r, s, t)$ when $f(x, y, z) = (x + y, y + 2z, x - z)$,

(c) $g(r, s, t)$ when $g(x, y, z) = (x + y - z, x - 3z, 2x + y)$.

4.125. Show that similarity of matrices is an equivalence relation.

LU FACTORIZATION

4.126. Find the LU and LDU factorization of each matrix:

(a) $A = \begin{pmatrix} 1 & 3 & -1 \\ 2 & 5 & 1 \\ 3 & 4 & 2 \end{pmatrix}$, (b) $B = \begin{pmatrix} 2 & 3 & 6 \\ 4 & 7 & 9 \\ 3 & 5 & 4 \end{pmatrix}$.

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

(a) Find the LU factorization of A.

(b) Let X_k denote the solution of $AX = B_k$. Find X_1, X_2, X_3, X_4 when $B_1 = (1, 1, 1)^T$ and $B_{k+1} = B_k + X_k$ for $k > 0$.

Answers to Supplementary Problems

4.74. $\begin{pmatrix} 1 & 2^n \\ 0 & 1 \end{pmatrix}$

4.76. (a) $k = 2$, (b) $k = -5$, (c) None

4.77. $\begin{pmatrix} 1 & 0 \\ 2 & 3 \end{pmatrix}$

4.78. (a) $A^2 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$, $A^3 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$, $A^k = 0$ for $k > 3$ (b) $B^n = \begin{pmatrix} 1 & n & n(n-1)/2 \\ 0 & 1 & n \\ 0 & 0 & 1 \end{pmatrix}$

4.79. $AB = BA$

4.80. (a) $\begin{pmatrix} \frac{29}{2} & -\frac{17}{2} & \frac{7}{2} \\ -\frac{5}{2} & \frac{3}{2} & -\frac{1}{2} \\ 3 & -2 & 1 \end{pmatrix}$, (b) $\begin{pmatrix} 8 & -3 & -1 \\ -5 & 2 & 1 \\ 10 & -4 & -1 \end{pmatrix}$, (c) $\begin{pmatrix} -8 & 5 & -1 \\ -\frac{9}{2} & \frac{5}{2} & -\frac{1}{2} \\ \frac{5}{2} & -\frac{3}{2} & \frac{1}{2} \end{pmatrix}$

4.81. (a) $\begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$, (b) $\begin{pmatrix} -10 & -20 & 4 & 7 \\ 3 & 6 & -1 & -2 \\ 5 & 8 & -2 & -3 \\ 2 & 3 & -1 & -1 \end{pmatrix}$

4.82. (a) $\begin{pmatrix} 1 & 0 \\ 3 & 1 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -2 \end{pmatrix}$ or $\begin{pmatrix} 1 & 0 \\ 3 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -2 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$, (b) No product: matrix has no inverse.

4.83. $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 3 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

4.84. $A = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}$, $C = \begin{pmatrix} 2 & 2 \\ 0 & 0 \end{pmatrix}$

4.86. $A = \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}$, $B = \begin{pmatrix} 4 & 3 \\ 3 & 0 \end{pmatrix}$

4.87. All diagonal entries must be 1 to be nonsingular. There are eight possible choices for the entries above the diagonal:

$$\begin{pmatrix} 0 & 0 \\ * & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ * & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ * & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ * & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ * & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ * & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ * & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ * & 1 \end{pmatrix}$$

4.88. (*a*) $2^4[2^n]$, (*b*) $2^{10}[2^{n(n+1)/2}]$, (*c*) $2^6[2^{n(n-1)/2}]$

4.89. (*a*) $\begin{pmatrix} 2 & 3 \\ 0 & 5 \end{pmatrix}, \begin{pmatrix} 2 & -7 \\ 0 & -5 \end{pmatrix}, \begin{pmatrix} -2 & 7 \\ 0 & 5 \end{pmatrix}, \begin{pmatrix} -2 & -3 \\ 0 & -5 \end{pmatrix}$, (*b*) None

4.90. $\begin{pmatrix} 1 & 2 & 1 \\ & 3 & 1 \\ & & 2 \end{pmatrix}$

4.91. (*a*) $A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, B = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, C = \begin{pmatrix} 4 & 6 \\ 0 & 0 \end{pmatrix}$

4.92. $\begin{pmatrix} 1 & 2 \\ 2 & 2 \end{pmatrix}\begin{pmatrix} 3 & 3 \\ 3 & 1 \end{pmatrix} = \begin{pmatrix} 9 & 5 \\ 12 & 8 \end{pmatrix}$

4.95. (*a*) $\begin{pmatrix} 2/\sqrt{29} & 5/\sqrt{29} \\ -5/\sqrt{29} & 2/\sqrt{29} \end{pmatrix}$, (*b*) $\begin{pmatrix} \frac{3}{5} & \frac{4}{5} \\ -\frac{4}{5} & \frac{3}{5} \end{pmatrix}$

4.96. (*a*) $\begin{pmatrix} 1/\sqrt{14} & 2/\sqrt{14} & 3/\sqrt{14} \\ 0 & -2/\sqrt{13} & 3/\sqrt{13} \\ 12/\sqrt{157} & -3/\sqrt{157} & -2/\sqrt{157} \end{pmatrix}$, (*b*) $\begin{pmatrix} 1/\sqrt{11} & 3/\sqrt{11} & 1/\sqrt{11} \\ 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 3/\sqrt{22} & -2/\sqrt{22} & 3/\sqrt{22} \end{pmatrix}$

4.98. A, C

4.104. $\begin{pmatrix} a_{11} & a_{21} & & \\ a_{21} & a_{22} & a_{32} & \\ & a_{32} & a_{33} & a_{34} \\ & & a_{43} & a_{44} \end{pmatrix}$ $\begin{pmatrix} b_{11} & b_{21} & & & \\ b_{21} & b_{22} & b_{23} & & \\ & b_{32} & b_{33} & b_{34} & \\ & & b_{43} & b_{44} & b_{45} \\ & & & b_{54} & b_{55} \end{pmatrix}$

4.105. $\begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 2 & 1 \\ 2 & 3 & 2 \\ 1 & 2 & 2 \end{pmatrix}$

4.106. (*a*) $x = a$ (parameter), $y = 0, z = 0$; (*b*) $x = 3, y = 0, z = 3$

4.109. A, B, C

4.111. A

4.114. $A = \left(\begin{array}{c|cc|cc} 1 & 2 & 3 & 4 & 5 \\ \hline 1 & 1 & 1 & 1 & 1 \\ 9 & 8 & 7 & 6 & 5 \\ \hline 2 & 2 & 2 & 2 & 2 \\ 3 & 3 & 3 & 3 & 3 \end{array}\right), B = \left(\begin{array}{cc|c|cc} 1 & 2 & 3 & 4 & 5 \\ 1 & 1 & 1 & 1 & 1 \\ \hline 9 & 8 & 7 & 6 & 5 \\ \hline 2 & 2 & 2 & 2 & 2 \\ 3 & 3 & 3 & 3 & 3 \end{array}\right)$

4.115. $A = \left(\begin{array}{c|cc} 1 & 0 & 0 \\ \hline 0 & 0 & 2 \\ 0 & 0 & 3 \end{array}\right), B = \left(\begin{array}{cc|cc|c} 1 & 2 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 4 & 0 & 0 \\ 0 & 0 & 5 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 6 \end{array}\right), C = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 2 & 0 \end{pmatrix}$

(C, itself, is a block diagonal matrix; no further partitioning of C is possible.)

4.116. (*a*) $M^2 = \begin{pmatrix} 4 & & & \\ & 9 & 8 & \\ & 4 & 9 & \\ & & & 9 \end{pmatrix}$, $M^3 = \begin{pmatrix} 8 & & & \\ & 25 & 44 & \\ & 22 & 25 & \\ & & & 27 \end{pmatrix}$

(*b*) $M^2 = \begin{pmatrix} 3 & 4 & & \\ 8 & 11 & & \\ & & 9 & 12 \\ & & 24 & 33 \end{pmatrix}$, $M^3 = \begin{pmatrix} 11 & 15 & & \\ 30 & 41 & & \\ & & 57 & 78 \\ & & 156 & 213 \end{pmatrix}$

4.118. $P = \begin{pmatrix} 1 & -1 & -1 & 26 \\ 0 & 1 & 3 & 13 \\ 0 & 0 & 1 & 9 \\ 0 & 0 & 0 & 7 \end{pmatrix}$, $B = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & -7 & \\ & & & 469 \end{pmatrix}$, sig $A = 2$

4.119. (*a*) $x = r - 3s - 19t$, $y = s + 7t$, $z = t$,
$q(r, s, t) = r^2 - s^2 + 36t^2$, rank $q = 3$, sig $q = 1$

(*b*) $x = r - 2t$, $y = s + 2t$, $z = t$,
$q(r, s, t) = 2r^2 - 3s^2 + 29t^2$, rank $q = 3$, sig $q = 1$

(*c*) $x = r - 2s + 18t$, $y = s - 7t$, $z = t$,
$q(r, s, t) = r^2 + s^2 - 62t^2$, rank $q = 3$, sig $q = 1$

(*d*) $x = r - s - t$, $y = s - t$, $z = t$,
$q(x, y, z) = r^2 + 2s^2$, rank $q = 2$, sig $q = 2$

4.120. (*a*) $k > \frac{25}{8}$; (*b*) $k < -12$ or $k > 12$; (*c*) $k > 5$

4.121. $q(x, y) = x^2 - y^2$, $u = (1, 1)$, $v = (1, -1)$

4.122. Suppose A has been diagonalized to $P^T AP = \text{diag}\,(a_i)$. Let $Q = \text{diag}\,(b_i)$ be defined by
$b_i = \begin{cases} 1/\sqrt{|a_i|} & \text{if } a_i \neq 0 \\ 1 & \text{if } a_i = 0 \end{cases}$. Then $B = Q^T P^T APQ = (PQ)^T A(PQ)$ has the required form.

4.124. (*a*) $Q(17, 5, 3)$, (*b*) $f(r, s, t) = (17r - 61s + 134t, 4r - 41s + 46t, 3r - 25s + 25t)$,
(*c*) $g(r, s, t) = (61r + s - 330t, 16r + 3s - 91t, 9r - 4s - 4t)$

4.126. (*a*) $A = \begin{pmatrix} 1 & & \\ 2 & 1 & \\ 3 & 5 & 1 \end{pmatrix} \begin{pmatrix} 1 & & \\ & -1 & \\ & & -10 \end{pmatrix} \begin{pmatrix} 1 & 3 & -1 \\ & 1 & -3 \\ & & 1 \end{pmatrix}$

(*b*) $B = \begin{pmatrix} 1 & & \\ 2 & 1 & \\ \frac{3}{2} & \frac{1}{2} & 1 \end{pmatrix} \begin{pmatrix} 2 & & \\ & 1 & \\ & & -\frac{7}{2} \end{pmatrix} \begin{pmatrix} 1 & \frac{3}{2} & 3 \\ & 1 & -3 \\ & & 1 \end{pmatrix}$

4.127. (*a*) $A = \begin{pmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 2 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 & -1 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{pmatrix}$

(*b*) $X_1 = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$, $B_2 = \begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix}$, $X_2 = \begin{pmatrix} 6 \\ 4 \\ 0 \end{pmatrix}$, $B_3 = \begin{pmatrix} 8 \\ 6 \\ 0 \end{pmatrix}$, $X_3 = \begin{pmatrix} 22 \\ 16 \\ -2 \end{pmatrix}$, $B_4 = \begin{pmatrix} 30 \\ 22 \\ -2 \end{pmatrix}$, $X_4 = \begin{pmatrix} 86 \\ 62 \\ -6 \end{pmatrix}$