PEARSON NEW INTERNATIONAL EDITION

Elementary Linear Algebra with Applications

Bernard Kolman David Hill

Ninth Edition

ALWAYS LEARNING

PEARSON

Pearson New International Edition

Elementary Linear Algebra
with Applications
Bernard Kolman David Hill
Ninth Edition

PEARSON

Pearson Education Limited
Edinburgh Gate
Harlow
Essex CM20 2JE
England and Associated Companies throughout the world

Visit us on the World Wide Web at: www.pearsoned.co.uk

PEARSON

ISBN 10: 1-292-02365-1
ISBN 13: 978-1-292-02365-6

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library

Printed in the United States of America

Table of Contents

GLOSSARY FOR LINEAR ALGEBRA

Additive inverse of a matrix: The additive inverse of an $m \times n$ matrix A is an $m \times n$ matrix B such that $A + B = O$. Such a matrix B is the negative of A, denoted $-A$, which is equal to $(-1)A$.

Adjoint: For an $n \times n$ matrix $A = \begin{bmatrix} a_{ij} \end{bmatrix}$ the adjoint of A, denoted adj A is the transpose of the matrix formed by replacing each entry by its cofactor A_{ij}; that is, adj $A = \begin{bmatrix} A_{ji} \end{bmatrix}$.

Angle between vectors: For nonzero vectors $\mathbf{u}$ and $\mathbf{v}$ in R^n the angle θ between $\mathbf{u}$ and $\mathbf{v}$ is determined from the expression

$$\cos(\theta) = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\| \|\mathbf{v}\|}.$$

Augmented matrix: For the linear system $A\mathbf{x} = \mathbf{b}$, the augmented matrix is formed by adjoining to the coefficient matrix A the right side vector $\mathbf{b}$. We expressed the augmented matrix as $\left[A \mid \mathbf{b} \right]$.

Back substitution: If $U = \begin{bmatrix} u_{ij} \end{bmatrix}$ is an upper triangular matrix all of whose diagonal entries are not zero, then the linear system $U\mathbf{x} = \mathbf{b}$ can be solved by back substitution. The process starts with the last equation and computes

$$x_n = \frac{b_n}{u_{nn}};$$

we use the next to last equation and compute

$$x_{n-1} = \frac{b_{n-1} - u_{n-1\,n}x_n}{u_{n-1\,n-1}};$$

continuing in this fashion using the jth equation we compute

$$x_j = \frac{b_j - \sum_{k=n}^{j+1} u_{jk}x_k}{u_{jj}}.$$

Basis: A set of vectors $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ from a vector space V is called a basis for V provided S spans V and S is a linearly independent set.

Cauchy–Schwarz inequality: For vectors $\mathbf{v}$ and $\mathbf{u}$ in R^n, the Cauchy–Schwarz inequality says that the absolute value of the dot product of $\mathbf{v}$ and $\mathbf{u}$ is less than or equal to the product of the lengths of $\mathbf{v}$ and $\mathbf{u}$; that is, $|\mathbf{v} \cdot \mathbf{u}| \leq \|\mathbf{v}\| \|\mathbf{u}\|$.

Characteristic equation: For a square matrix A, its characteristic equation is given by $f(t) = \det(A - tI) = 0$.

Characteristic polynomial: For a square matrix A, its characteristic polynomial is given by $f(t) = \det(A - tI)$.

Closure properties: Let V be a given set, with members that we call vectors, and two operations, one called vector addition, denoted $\oplus$, and the second called scalar multiplication, denoted $\odot$. We say that V is closed under $\oplus$, provided for $\mathbf{u}$ and $\mathbf{v}$ in V, $\mathbf{u} \oplus \mathbf{v}$ is a member of V. We say that V is closed under $\odot$, provided for any real number k, $k \odot \mathbf{u}$ is a member of V.

Coefficient matrix: A linear system of m equations in n unknowns has the form

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \vdots \qquad \vdots \qquad\qquad \vdots \quad &\;\; \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m. \end{aligned}$$

The matrix

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

is called the coefficient matrix of the linear system.

Cofactor: For an $n \times n$ matrix $A = \begin{bmatrix} a_{ij} \end{bmatrix}$ the cofactor A_{ij} of a_{ij} is defined as $A_{ij} = (-1)^{i+j} \det(M_{ij})$, where M_{ij} is the ij-minor of A.

Column rank: The column rank of a matrix A is the dimension of the column space of A or equivalently the number of linearly independent columns of A.

Column space: The column space of a real $m \times n$ matrix A is the subspace of R^m spanned by the columns of A.

Complex vector space: A complex vector space V is a set, with members that we call vectors, and two operations: one called vector addition, denoted $\oplus$, and the second called scalar multiplication, denoted $\odot$. We require that V be closed under $\oplus$, that is, for $\mathbf{u}$ and $\mathbf{v}$ in V, $\mathbf{u} \oplus \mathbf{v}$ is a member of V; in addition we require that V be closed under $\odot$, that is, for any complex number k, $k \odot \mathbf{u}$ is a member of V. There are 8 other properties that must be satisfied before V with the two operations $\oplus$ and $\odot$ is called a complex vector space.

Complex vector subspace: A subset W of a complex vector space V that is closed under addition and scalar multiplication is called a complex subspace of V.

From *Elementary Linear Algebra with Applications*, Ninth Edition. Bernard Kolman, David R. Hill.

Components of a vector: The components of a vector $\mathbf{v}$ in R^n are its entries;

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}.$$

Composite linear transformation: Let L_1 and L_2 be linear transformations with $L_1 \colon V \to W$ and $L_2 \colon W \to U$. Then the composition $L_2 \circ L_1 \colon V \to U$ is a linear transformation and for $\mathbf{v}$ in V, we compute $(L_2 \circ L_1)(\mathbf{v}) = L_2(L_1(v))$.

Computation of a determinant via reduction to triangular form: For an $n \times n$ matrix A the determinant of A, denoted $\det(A)$ or $|A|$, can be computed with the aid of elementary row operations as follows. Use elementary row operations on A, keeping track of the operations used, to obtain an upper triangular matrix. Using the changes in the determinant as the result of applying a row operation and the fact that the determinant of an upper triangular matrix is the product of its diagonal entries, we can obtain an appropriate expression for $\det(A)$.

Consistent linear system: A linear system $A\mathbf{x} = \mathbf{b}$ is called consistent if the system has at least one solution.

Coordinates: The coordinates of a vector $\mathbf{v}$ in a vector space V with ordered basis $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ are the coefficients $c_1, c_2, \ldots, c_n$ such that $\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n$. We denote the coordinates of $\mathbf{v}$ relative to the basis S by $\left[\mathbf{v}\right]_S$ and write

$$\left[\mathbf{v}\right]_S = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}.$$

Cross product: The cross product of a pair of vectors $\mathbf{u}$ and $\mathbf{v}$ from R^3 is denoted $\mathbf{u} \times \mathbf{v}$, and is computed as the determinant

$$\begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix},$$

where $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$ are the unit vectors in the x-, y-, and z-directions, respectively.

Defective matrix: A square matrix A is called defective if it has an eigenvalue of multiplicity $m > 1$ for which the associated eigenspace has a basis with fewer than m vectors.

Determinant: For an $n \times n$ matrix A the determinant of A, denoted $\det(A)$ or $|A|$, is a scalar that is computed as the sum of all possible products of n entries of A each with its appropriate sign, with exactly one entry from each row and exactly one entry from each column.

Diagonal matrix: A square matrix $A = \left[a_{ij}\right]$ is called diagonal provided $a_{ij} = 0$ whenever $i \neq j$.

Diagonalizable: A square matrix A is called diagonalizable provided it is similar to a diagonal matrix D; that is, there exists a nonsingular matrix P such that $P^{-1}AP = D$.

Difference of matrices: The difference of the $m \times n$ matrices A and B is denoted $A - B$ and is equal to the sum $A + (-1)B$. The difference $A - B$ is the $m \times n$ matrix whose entries are the difference of corresponding entries of A and B.

Difference of vectors: The difference of the vectors $\mathbf{v}$ and $\mathbf{w}$ in a vector space V is denoted $\mathbf{v} - \mathbf{w}$, which is equal to the sum $\mathbf{v} + (-1)\mathbf{w}$. If $V = R^n$, then $\mathbf{v} - \mathbf{w}$ is computed as the difference of corresponding entries.

Dilation: The linear transformation $L \colon R^n \to R^n$ given by $L(\mathbf{v}) = k\mathbf{v}$, for $k > 1$, is called a dilation.

Dimension: The dimension of a nonzero vector space V is the number of vectors in a basis for V. The dimension of the vector space $\{\mathbf{0}\}$ is defined as zero.

Distance between points (or vectors): The distance between the points $(u_1, u_2, \ldots, u_n)$ and $(v_1, v_2, \ldots, v_n)$ is the length of the vector $\mathbf{u} - \mathbf{v}$, where $\mathbf{u} = (u_1, u_2, \ldots, u_n)$ and $\mathbf{v} = (v_1, v_2, \ldots, v_n)$, and is given by

$$\|\mathbf{u} - \mathbf{v}\| = \sqrt{(u_1 - v_1)^2 + (u_2 - v_2)^2 + \cdots + (u_n - v_n)^2}.$$

Thus we see that the distance between vectors in R^n is also $\|\mathbf{u} - \mathbf{v}\|$.

Dot product: For vectors $\mathbf{v}$ and $\mathbf{w}$ in R^n the dot product of $\mathbf{v}$ and $\mathbf{w}$ is also called the standard inner product or just the inner product of $\mathbf{v}$ and $\mathbf{w}$. The dot product of $\mathbf{v}$ and $\mathbf{w}$ in R^n is denoted $\mathbf{v} \cdot \mathbf{w}$ and is computed as

$$\mathbf{v} \cdot \mathbf{w} = v_1w_1 + v_2w_2 + \cdots + v_nw_n.$$

Eigenspace: The set of all eigenvectors of a square matrix A associated with a specified eigenvalue λ of A, together with the zero vector, is called the eigenspace associated with the eigenvalue λ.

Eigenvalue: An eigenvalue of an $n \times n$ matrix A is a scalar λ for which there exists a nonzero n-vector $\mathbf{x}$ such that $A\mathbf{x} = \lambda\mathbf{x}$. The vector $\mathbf{x}$ is an eigenvector associated with the eigenvalue λ.

Eigenvector: An eigenvector of an $n \times n$ matrix A is a nonzero n-vector $\mathbf{x}$ such that $A\mathbf{x}$ is a scalar multiple of $\mathbf{x}$; that is, there exists some scalar λ such that $A\mathbf{x} = \lambda\mathbf{x}$. The scalar is an eigenvalue of the matrix A.

Elementary row operations: An elementary row operation on a matrix is any of the following three operations: (1) an interchange of rows, (2) multiplying a row by a nonzero scalar, and (3) replacing a row by adding a scalar multiple of a different row to it.

Equal matrices: The $m \times n$ matrices A and B are equal provided corresponding entries are equal; that is, $A = B$ if $a_{ij} = b_{ij}$, $i = 1, 2, \ldots, m$, $j = 1, 2, \ldots, n$.

Equal vectors: Vectors $\mathbf{v}$ and $\mathbf{w}$ in R^n are equal provided corresponding entries are equal; that is, $\mathbf{v} = \mathbf{w}$ if their corresponding components are equal.

Finite-dimensional vector space: A vector space V that has a basis that is a finite subset of V is said to be finite dimensional.

Forward substitution: If $L = \left[l_{ij} \right]$ is a lower triangular matrix all of whose diagonal entries are not zero, then the linear system $L\mathbf{x} = \mathbf{b}$ can be solved by forward substitution. The process starts with the first equation and computes

$$x_1 = \frac{b_1}{l_{11}};$$

next we use the second equation and compute

$$x_2 = \frac{b_2 - l_{21}x_1}{l_{22}};$$

continuing in this fashion using the jth equation we compute

$$x_j = \frac{b_j - \displaystyle\sum_{k=1}^{j-1} l_{jk}x_k}{l_{jj}}.$$

Fundamental vector spaces associated with a matrix: If A is an $m \times n$ matrix there are four fundamental subspaces associated with A: (1) the null space of A, a subspace of R^n; (2) the row space of A, a subspace of R^n; (3) the null space of A^T, a subspace of R^m; and (4) the column space of A, a subspace of R^m.

Gaussian elimination: For the linear system $A\mathbf{x} = \mathbf{b}$ form the augmented matrix $\left[A \;\vdots\; \mathbf{b} \right]$. Compute the row echelon form of the augmented matrix; then the solution can be computed using back substitution.

Gauss–Jordan reduction: For the linear system $A\mathbf{x} = \mathbf{b}$ form the augmented matrix $\left[A \;\vdots\; \mathbf{b} \right]$. Compute the reduced row echelon form of the augmented matrix; then the solution can be computed using back substitution.

General solution: The general solution of a consistent linear system $A\mathbf{x} = \mathbf{b}$ is the set of all solutions to the system. If $\mathbf{b} = \mathbf{0}$, then the general solution is just the set of all solutions to the homogeneous system $A\mathbf{x} = \mathbf{0}$, denoted $\mathbf{x}_h$. If $\mathbf{b} \neq \mathbf{0}$, then the general solution of the nonhomogeneous system consists of a particular solution of $A\mathbf{x} = \mathbf{b}$, denoted $\mathbf{x}_p$, together with $\mathbf{x}_h$; that is, the general solution is expressed as $\mathbf{x}_p + \mathbf{x}_h$.

Gram–Schmidt process: The Gram–Schmidt process converts a basis for a subspace into an orthonormal basis for the same subspace.

Hermitian matrix: An $n \times n$ complex matrix A is called Hermitian provided $\overline{A}^T = A$.

Homogeneous system: A homogeneous system is a linear system in which the right side of each equation is zero. We denote a homogeneous system by $A\mathbf{x} = \mathbf{0}$.

Identity matrix: The $n \times n$ identity matrix, denoted I_n, is a diagonal matrix with diagonal entries of all 1s.

Inconsistent linear system: A linear system $A\mathbf{x} = \mathbf{b}$ that has no solution is called inconsistent.

Infinite-dimensional vector space: A vector space V for which there is no finite subset of vectors that form a basis for V is said to be infinite dimensional.

Inner product: For vectors $\mathbf{v}$ and $\mathbf{w}$ in R^n the inner product of $\mathbf{v}$ and $\mathbf{w}$ is also called the dot product or standard inner product of $\mathbf{v}$ and $\mathbf{w}$. The inner product of $\mathbf{v}$ and $\mathbf{w}$ in R^n is denoted $\mathbf{v} \cdot \mathbf{w}$ and is computed as

$$\mathbf{v} \cdot \mathbf{w} = v_1w_1 + v_2w_2 + \cdots + v_nw_n.$$

Invariant subspace: A subspace W of a vector space V is said to be invariant under the linear transformation $L \colon V \to V$, provided $L(\mathbf{v})$ is in W for all vectors $\mathbf{v}$ in W.

Inverse linear transformation: See invertible linear transformation.

Inverse of a matrix: An $n \times n$ matrix A is said to have an inverse provided there exists an $n \times n$ matrix B such that $AB = BA = I$. We call B the inverse of A and denote it as A^{-1}. In this case, A is also called nonsingular.

Invertible linear transformation: A linear transformation $L \colon V \to W$ is called invertible if there exists a linear transformation, denoted L^{-1}, such that $L^{-1}(L(\mathbf{v})) = \mathbf{v}$, for all vectors $\mathbf{v}$ in V and $L(L^{-1}(\mathbf{w})) = \mathbf{w}$, for all vectors $\mathbf{w}$ in W.

Isometry: An isometry is a linear transformation L that preserves the distance between pairs of vectors; that is, $\|L(\mathbf{v}) - L(\mathbf{u})\| = \|\mathbf{v} - \mathbf{u}\|$, for all vectors $\mathbf{u}$ and $\mathbf{v}$. Since an isometry preserves distances, it also preserves lengths; that is, $\|L(\mathbf{v})\| = \|\mathbf{v}\|$, for all vectors $\mathbf{v}$.

Length (or magnitude) of a vector: The length of a vector $\mathbf{v}$ in R^n is denoted $\|\mathbf{v}\|$ and is computed as the expression

$$\sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}.$$

For a vector $\mathbf{v}$ in a vector space V on which an inner product (dot product) is defined, the length of $\mathbf{v}$ is computed as $\|\mathbf{v}\| = \sqrt{\mathbf{v} \cdot \mathbf{v}}$.

Linear combination: A linear combination of vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ from a vector space V is an expression of the form $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k$, where the $c_1, c_2, \ldots, c_k$ are scalars. A linear combination of the $m \times n$ matrices $A_1, A_2, \ldots, A_k$ is given by $c_1A_1 + c_2A_2 + \cdots + c_kA_k$.

Linear operator: A linear operator is a linear transformation L from a vector space to itself; that is, $L \colon V \to V$.

Linear system: A system of m linear equations in n unknowns $x_1, x_2, \ldots, x_n$ is a set of linear equations in the n unknowns. We express a linear system in matrix form as $A\mathbf{x} = \mathbf{b}$, where A is the matrix of coefficients, $\mathbf{x}$ is the vector of unknowns, and $\mathbf{b}$ is the vector of right sides of the linear equations. (See coefficient matrix.)

Linear transformation: A linear transformation $L\colon V \to W$ is a function assigning a unique vector $L(\mathbf{v})$ in W to each vector $\mathbf{v}$ in V such that two properties are satisfied: (1) $L(\mathbf{u}+\mathbf{v}) = L(\mathbf{u}) + L(\mathbf{v})$, for every $\mathbf{u}$ and $\mathbf{v}$ in V, and (2) $L(k\mathbf{v}) = kL(\mathbf{v})$, for every $\mathbf{v}$ in V and every scalar k.

Linearly dependent: A set of vectors $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is called linearly dependent provided there exists a linear combination $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n$ that produces the zero vector when not all the coefficients are zero.

Linearly independent: A set of vectors $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is called linearly independent provided the only linear combination $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n$ that produces the zero vector is when all the coefficients are zero, that is, only when $c_1 = c_2 = \cdots = c_n = 0$.

Lower triangular matrix: A square matrix with zero entries above its diagonal entries is called lower triangular.

LU-factorization (or LU-decomposition): An LU-factorization of a square matrix A expresses A as the product of a lower triangular matrix L and an upper triangular matrix U; that is, $A = LU$.

Main diagonal of a matrix: The main diagonal of an $n \times n$ matrix A is the set of entries $a_{11}, a_{22}, \ldots, a_{nn}$.

Matrix: An $m \times n$ matrix A is a rectangular array of mn entries arranged in m rows and n columns.

Matrix addition: For $m \times n$ matrices $A = \left[a_{ij}\right]$ and $B = \left[b_{ij}\right]$, the addition of A and B is performed by adding corresponding entries; that is, $A + B = \left[a_{ij}\right] + \left[b_{ij}\right]$. This is also called the sum of the matrices A and B.

Matrix representing a linear transformation: Let $L\colon V \to W$ be a linear transformation from an n-dimensional space V to an m-dimensional space W. For a basis $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ in V and a basis $T = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m\}$ in W there exists an $m \times n$ matrix A, with column j of $A = \left[L(\mathbf{v}_j)\right]_T$ such that the coordinates of $L(\mathbf{x})$, for any x in V, with respect to the T basis can be computed as $\left[L(\mathbf{x})\right]_T = A\left[\mathbf{x}\right]_S$. We say A is the matrix representing the linear transformation L.

Matrix transformation: For an $m \times n$ matrix A the function f defined by $f(\mathbf{u}) = A\mathbf{u}$ for $\mathbf{u}$ in R^n is called the matrix transformation from R^n to R^m defined by the matrix A.

Minor: Let $A = \left[a_{ij}\right]$ be an $n \times n$ matrix and M_{ij} the $(n-1) \times (n-1)$ submatrix of A obtained by deleting the ith row and jth column of A. The determinant $\det(M_{ij})$ is called the minor of a_{ij}.

Multiplicity of an eigenvalue: The multiplicity of an eigenvalue λ of a square matrix A is the number of times λ is a root of the characteristic polynomial of A.

Natural (or standard) basis: The natural basis for R^n is the set of vectors $\mathbf{e}_j$ = column j (or, equivalently, row j) of the $n \times n$ identity matrix, $j = 1, 2, \ldots, n$.

Negative of a vector: The negative of a vector $\mathbf{u}$ is a vector $\mathbf{w}$ such that $\mathbf{u} + \mathbf{w} = \mathbf{0}$, the zero vector. We denote the negative of $\mathbf{u}$ as $-\mathbf{u} = (-1)\mathbf{u}$.

Nonhomogeneous system: A linear system $A\mathbf{x} = \mathbf{b}$ is called nonhomogeneous provided the vector $\mathbf{b}$ is not the zero vector.

Nonsingular (or invertible) matrix: An $n \times n$ matrix A is called nonsingular provided there exists an $n \times n$ matrix B such that $AB = BA = I$. We call B the inverse of A and denote it as A^{-1}.

Nontrivial solution: A nontrivial solution of a linear system $A\mathbf{x} = \mathbf{b}$ is any vector $\mathbf{x}$ containing at least one nonzero entry such that $A\mathbf{x} = \mathbf{b}$.

Normal matrix: An $n \times n$ complex matrix A is called normal provided $\left(\overline{A}^T\right) A = A \left(\overline{A}^T\right)$.

***n*-space**: The set of all n-vectors is called n-space. For vectors whose entries are real numbers we denote n-space as R^n. For a special case see 2-space.

Nullity: The nullity of the matrix A is the dimension of the null space of A.

***n*-vector**: A $1 \times n$ or an $n \times 1$ matrix is called an n-vector. When n is understood, we refer to n-vectors merely as vectors.

One-to-one: A function $f\colon S \to T$ is said to be one-to-one provided $f(s_1) \neq f(s_2)$ whenever s_1 and s_2 are distinct elements of S. A linear transformation $L\colon V \to W$ is called one-to-one provided L is a one-to-one function.

Onto: A function $f\colon S \to T$ is said to be onto provided for each member t of T there is some member s in S so that $f(s) = t$. A linear transformation $L\colon V \to W$ is called onto provided range $L = W$.

Ordered basis: A set of vectors $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ in a vector space V is called an ordered basis for V provided S is a basis for V and if we reorder the vectors in S, this new ordering of the vectors in S is considered a different basis for V.

Orthogonal basis: A basis for a vector space V that is also an orthogonal set is called an orthogonal basis for V.

Orthogonal complement: The orthogonal complement of a set S of vectors in a vector space V is the set of all vectors in V that are orthogonal to all vectors in S.

Orthogonal matrix: A square matrix P is called orthogonal provided $P^{-1} = P^T$.

Orthogonal projection: For a vector $\mathbf{v}$ in a vector space V, the orthogonal projection of $\mathbf{v}$ onto a subspace W of V with orthonormal basis $\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_k\}$ is the vector $\mathbf{w}$ in W, where

$\mathbf{w} = (\mathbf{v} \cdot \mathbf{w}_1)\mathbf{w}_1 + (\mathbf{v} \cdot \mathbf{w}_2)\mathbf{w}_2 + \cdots + (\mathbf{v} \cdot \mathbf{w}_k)\mathbf{w}_k$. Vector $\mathbf{w}$ is the vector in W that is closest to $\mathbf{v}$.

Orthogonal set: A set of vectors $S = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_k\}$ from a vector space V on which an inner product is defined is an orthogonal set provided none of the vectors is the zero vector and the inner product of any two different vectors is zero.

Orthogonal vectors: A pair of vectors is called orthogonal provided their dot (inner) product is zero.

Orthogonally diagonalizable: A square matrix A is said to be orthogonally diagonalizable provided there exists an orthogonal matrix P such that $P^{-1}AP$ is a diagonal matrix. That is, A is similar to a diagonal matrix using an orthogonal matrix P.

Orthonormal basis: A basis for a vector space V that is also an orthonormal set is called an orthonormal basis for V.

Orthonormal set: A set of vectors $S = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_k\}$ from a vector space V on which an inner product is defined is an orthonormal set provided each vector is a unit vector and the inner product of any two different vectors is zero.

Parallel vectors: Two nonzero vectors are said to be parallel if one is a scalar multiple of the other.

Particular solution: A particular solution of a consistent linear system $A\mathbf{x} = \mathbf{b}$ is a vector $\mathbf{x}_p$ containing no arbitrary constants such that $A\mathbf{x}_p = \mathbf{b}$.

Partitioned matrix: A matrix that has been partitioned into submatrices by drawing horizontal lines between rows and/or vertical lines between columns is called a partitioned matrix. There are many ways to partition a matrix.

Perpendicular (or orthogonal) vectors: A pair of vectors is said to be perpendicular or orthogonal provided their dot product is zero.

Pivot: When using row operations on a matrix A, a pivot is a nonzero entry of a row that is used to zero-out entries in the column in which the pivot resides.

Positive definite: Matrix A is positive definite provided A is symmetric and all of its eigenvalues are positive.

Powers of a matrix: For a square matrix A and nonnegative integer k, the kth power of A, denoted A^k, is the product of A with itself k times; $A^k = A \cdot A \cdot \cdots \cdot A$, where there are k factors.

Projection: The projection of a point P in a plane onto a line L in the same plane is the point Q obtained by intersecting the line L with the line through P that is perpendicular to L. The linear transformation $L\colon R^3 \to R^2$ defined by $L(x, y, z) = (x, y)$ is called a projection of R^3 onto R^2. (See also orthogonal projection.)

Range: The range of a function $f\colon S \to T$ is the set of all members t of T such that there is a member s in S with $f(s) = t$. The range of a linear transformation $L\colon V \to W$ is the set of all vectors in W that are images under L of vectors in V.

Rank: Since row rank A = column rank A, we just refer to the rank of the matrix A as rank A. Equivalently, rank A = the number of linearly independent rows (columns) of A = the number of leading 1s in the reduced row echelon form of A.

Real vector space: A real vector space V is a set, with members that we call vectors and two operations: one is called vector addition, denoted $\oplus$, and the second called scalar multiplication, denoted $\odot$. We require that V be closed under $\oplus$; that is, for $\mathbf{u}$ and $\mathbf{v}$ in V, $\mathbf{u} \oplus \mathbf{v}$ is a member of V. In addition we require that V be closed under $\odot$; that is, for any real number k, $k \odot \mathbf{u}$ is a member of V. There are 8 other properties that must be satisfied before V with the two operations $\oplus$ and $\odot$ is called a vector space.

Reduced row echelon form: A matrix is said to be in reduced row echelon form provided it satisfies the following properties: (1) All zero rows, if there are any, appear as bottom rows. (2) The first nonzero entry in a nonzero row is a 1; it is called a leading 1. (3) For each nonzero row, the leading 1 appears to the right and below any leading 1s in preceding rows. (4) If a column contains a leading 1, then all other entries in that column are zero.

Reflection: The linear transformation $L\colon R^2 \to R^2$ given by $L(x, y) = (x, -y)$ is called a reflection with respect to the x-axis. Similarly, $L(x, y) = (-x, y)$ is called a reflection with respect to the y-axis.

Roots of the characteristic polynomial: For a square matrix A, the roots of its characteristic polynomial $f(t) = \det(A - tI)$ are the eigenvalues of A.

Rotation: The linear transformation $L\colon R^2 \to R^2$ given by

$$L\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

is called a counterclockwise rotation in the plane by the angle θ.

Row echelon form: A matrix is said to be in row echelon form provided it satisfies the following properties: (1) All zero rows, if there are any, appear as bottom rows. (2) The first nonzero entry in a nonzero row is a 1; it is called a leading 1. (3) For each nonzero row, the leading 1 appears to the right and below any leading 1s in preceding rows.

Row equivalent: The $m \times n$ matrices A and B are row equivalent provided there exists a set of row operations that when performed on A yield B.

Row rank: The row rank of matrix A is the dimension of the row space of A or, equivalently, the number of linearly independent rows of A.

Row space: The row space of a real $m \times n$ matrix A is the subspace of R^n spanned by the rows of A.

Scalar matrix: Matrix A is a scalar matrix provided A is a diagonal matrix with equal diagonal entries.

Scalar multiple of a matrix: For an $m \times n$ matrix $A = \left[a_{ij}\right]$ and scalar r, the scalar multiple of A by r gives the $m \times n$ matrix $rA = \left[ra_{ij}\right]$.

Scalar multiple of a vector: If $\mathbf{v}$ is in real vector space V, then for any real number k, a scalar, the scalar multiple of $\mathbf{v}$ by k is denoted $k\mathbf{v}$. If $V = R^n$, then $k\mathbf{v} = (kv_1, kv_2, \ldots, kv_n)$.

Scalars: In a real vector space V the scalars are real numbers and are used when we form scalar multiples $k\mathbf{v}$, where $\mathbf{v}$ is in V. Also, when we form linear combinations of vectors the coefficients are scalars.

Shear: A shear in the x-direction is defined by the matrix transformation

$$L(\mathbf{u}) = \begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix},$$

where k is a scalar. Similarly, a shear in the y-direction is given by

$$L(\mathbf{u}) = \begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}.$$

Similar matrices: Matrices A and B are similar provided there exists a nonsingular matrix P such that $A = P^{-1}BP$.

Singular (or noninvertible) matrix: A matrix A that has no inverse matrix is said to be singular. Any square matrix whose reduced row echelon form is not the identity matrix is singular.

Skew symmetric matrix: A square real matrix A such that $A = -A^T$ is called skew symmetric.

Solution space: The solution space of an $m \times n$ real homogeneous system $A\mathbf{x} = \mathbf{0}$ is the set W of all n-vectors $\mathbf{x}$ such that A times $\mathbf{x}$ gives the zero vector. W is a subspace of R^n.

Solution to a homogeneous system: A solution to a homogeneous system $A\mathbf{x} = \mathbf{0}$ is a vector $\mathbf{x}$ such that A times $\mathbf{x}$ gives the zero vector.

Solution to a linear system: A solution to a linear system $A\mathbf{x} = \mathbf{b}$ is any vector $\mathbf{x}$ such that A times $\mathbf{x}$ gives the vector $\mathbf{b}$.

Span: The span of a set $W = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_k\}$, denoted by span W, from a vector space V is the set of all possible linear combinations of the vectors $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_k$. Span W is a subspace of V.

Square matrix: A matrix with the same number of rows as columns is called a square matrix.

Standard inner product: For vectors $\mathbf{v}$ and $\mathbf{w}$ in R^n the standard inner product of $\mathbf{v}$ and $\mathbf{w}$ is also called the dot product of $\mathbf{v}$ and $\mathbf{w}$, denoted $\mathbf{v} \cdot \mathbf{w} = v_1w_1 + v_2w_2 + \cdots + v_nw_n$.

Submatrix: A matrix obtained from a matrix A by deleting rows and/or columns is called a submatrix of A.

Subspace: A subset W of a vector space V that is closed under addition and scalar multiplication is called a subspace of V.

Sum of vectors: The sum of two vectors is also called vector addition. In R^n adding corresponding components of the vectors performs the sum of two vectors. In a vector space V, $\mathbf{u} \oplus \mathbf{v}$ is computed using the definition of the operation $\oplus$.

Summation notation: A compact notation to indicate the sum of a set $\{a_1, a_2, \ldots, a_n\}$; the sum of a_1 through a_n is denoted in summation notation as $\sum_{i=1}^{n} a_i$.

Symmetric matrix: A square real matrix A such that $A = A^T$ is called symmetric.

Transition matrix: Let $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ and $T = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n\}$ be bases for an n-dimensional vector space V. The transition matrix from the T-basis to the S-basis is an $n \times n$ matrix, denoted $P_{S \leftarrow T}$, that converts the coordinates of a vector $\mathbf{v}$ relative to the T-basis into the coordinates of $\mathbf{v}$ relative to the S-basis; $\left[\mathbf{v}\right]_S = P_{S \leftarrow T}\left[\mathbf{v}\right]_T$.

Translation: Let $T\colon V \to V$ be defined by $T(\mathbf{v}) = \mathbf{v} + \mathbf{b}$ for all $\mathbf{v}$ in V and any fixed vector $\mathbf{b}$ in V. We call this the translation by the vector $\mathbf{b}$.

Transpose of a matrix: The transpose of an $m \times n$ matrix A is the $n \times m$ matrix obtained by forming columns from each row of A. The transpose of A is denoted A^T.

Trivial solution: The trivial solution of a homogeneous system $A\mathbf{x} = \mathbf{0}$ is the zero vector.

2-space: The set of all 2-vectors is called 2-space. For vectors whose entries are real numbers we denote 2-space as R^2.

Unit vector: A vector of length 1 is called a unit vector.

Unitary matrix: An $n \times n$ complex matrix A is called unitary provided $A^{-1} = \overline{A}^T$.

Upper triangular matrix: A square matrix with zero entries below its diagonal entries is called upper triangular.

Vector: The generic name for any member of a vector space. (See also 2-vector and n-vector.)

Vector addition: The sum of two vectors is also called vector addition. In R^n adding corresponding components of the vectors performs vector addition.

Zero matrix: A matrix with all zero entries is called the zero matrix.

Zero polynomial: A polynomial all of whose coefficients are zero is called the zero polynomial.

Zero subspace: The subspace consisting of exactly the zero vector of a vector space is called the zero subspace.

Zero vector: A vector with all zero entries is called the zero vector.

CHAPTER

1 Linear Equations and Matrices

1.1 Systems of Linear Equations

One of the most frequently recurring practical problems in many fields of study—such as mathematics, physics, biology, chemistry, economics, all phases of engineering, operations research, and the social sciences—is that of solving a system of linear equations. The equation

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b, \tag{1}$$

which expresses the real or complex quantity b in terms of the unknowns $x_1, x_2, \ldots, x_n$ and the real or complex constants $a_1, a_2, \ldots, a_n$, is called a **linear equation**. In many applications we are given b and must find numbers $x_1, x_2, \ldots, x_n$ satisfying (1).

A **solution** to linear Equation (1) is a sequence of n numbers $s_1, s_2, \ldots, s_n$, which has the property that (1) is satisfied when $x_1 = s_1, x_2 = s_2, \ldots, x_n = s_n$ are substituted in (1). Thus $x_1 = 2$, $x_2 = 3$, and $x_3 = -4$ is a solution to the linear equation

$$6x_1 - 3x_2 + 4x_3 = -13,$$

because

$$6(2) - 3(3) + 4(-4) = -13.$$

More generally, a **system of m linear equations in n unknowns**, $x_1, x_2, \ldots, x_n$, or a **linear system**, is a set of m linear equations each in n unknowns. A linear

Note: Appendix A reviews some very basic material dealing with sets and functions. It can be consulted at any time, as needed.

system can conveniently be written as

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \vdots \qquad \vdots \qquad\qquad \vdots \quad &\quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m. \end{aligned} \tag{2}$$

Thus the ith equation is

$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = b_i.$$

In (2) the a_{ij} are known constants. Given values of $b_1, b_2, \ldots, b_m$, we want to find values of $x_1, x_2, \ldots, x_n$ that will satisfy each equation in (2).

A **solution** to linear system (2) is a sequence of n numbers $s_1, s_2, \ldots, s_n$, which has the property that each equation in (2) is satisfied when $x_1 = s_1, x_2 = s_2, \ldots, x_n = s_n$ are substituted.

If the linear system (2) has no solution, it is said to be **inconsistent**; if it has a solution, it is called **consistent**. If $b_1 = b_2 = \cdots = b_m = 0$, then (2) is called a **homogeneous system**. Note that $x_1 = x_2 = \cdots = x_n = 0$ is always a solution to a homogeneous system; it is called the **trivial solution**. A solution to a homogeneous system in which not all of $x_1, x_2, \ldots, x_n$ are zero is called a **nontrivial solution**.

Consider another system of r linear equations in n unknowns:

$$\begin{aligned} c_{11}x_1 + c_{12}x_2 + \cdots + c_{1n}x_n &= d_1 \\ c_{21}x_1 + c_{22}x_2 + \cdots + c_{2n}x_n &= d_2 \\ \vdots \qquad \vdots \qquad\qquad \vdots \quad &\quad \vdots \\ c_{r1}x_1 + c_{r2}x_2 + \cdots + c_{rn}x_n &= d_r. \end{aligned} \tag{3}$$

We say that (2) and (3) are **equivalent** if they both have exactly the same solutions.

EXAMPLE 1 The linear system

$$\begin{aligned} x_1 - 3x_2 &= -7 \\ 2x_1 + x_2 &= 7 \end{aligned} \tag{4}$$

has only the solution $x_1 = 2$ and $x_2 = 3$. The linear system

$$\begin{aligned} 8x_1 - 3x_2 &= 7 \\ 3x_1 - 2x_2 &= 0 \\ 10x_1 - 2x_2 &= 14 \end{aligned} \tag{5}$$

also has only the solution $x_1 = 2$ and $x_2 = 3$. Thus (4) and (5) are equivalent. ■

To find a solution to a linear system, we shall use a technique called the **method of elimination**; that is, we eliminate some variables by adding a multiple of one equation to another equation. Elimination merely amounts to the development of a new linear system that is equivalent to the original system, but is much simpler to solve. Readers have probably confined their earlier work in this area to linear systems in which $m = n$, that is, linear systems having as many equations as unknowns. In this course we shall broaden our outlook by dealing with systems in which we have $m = n$, $m < n$, and $m > n$. Indeed, there are numerous applications in which $m \neq n$. If we deal with two, three, or four unknowns, we shall often write them as x, y, z, and w. In this section we use the method of elimination as it was studied in high school. In Section 2.2 we shall look at this method in a much more systematic manner.

EXAMPLE 2

The director of a trust fund has \$100,000 to invest. The rules of the trust state that both a certificate of deposit (CD) and a long-term bond must be used. The director's goal is to have the trust yield \$7800 on its investments for the year. The CD chosen returns 5% per annum, and the bond 9%. The director determines the amount x to invest in the CD and the amount y to invest in the bond as follows:

Since the total investment is \$100,000, we must have $x + y = 100{,}000$. Since the desired return is \$7800, we obtain the equation $0.05x + 0.09y = 7800$. Thus, we have the linear system

$$\begin{aligned} x + y &= 100{,}000 \\ 0.05x + 0.09y &= 7800. \end{aligned} \tag{6}$$

To eliminate x, we add (-0.05) times the first equation to the second, obtaining

$$0.04y = 2800,$$

an equation having no x term. We have eliminated the unknown x. Then solving for y, we have

$$y = 70{,}000,$$

and substituting into the first equation of (6), we obtain

$$x = 30{,}000.$$

To check that $x = 30{,}000$, $y = 70{,}000$ is a solution to (6), we verify that these values of x and y satisfy *each* of the equations in the given linear system. This solution is the only solution to (6); the system is consistent. The director of the trust should invest \$30,000 in the CD and \$70,000 in the long-term bond. ■

EXAMPLE 3

Consider the linear system

$$\begin{aligned} x - 3y &= -7 \\ 2x - 6y &= 7. \end{aligned} \tag{7}$$

Again, we decide to eliminate x. We add (-2) times the first equation to the second one, obtaining

$$0 = 21,$$

which makes no sense. This means that (7) has no solution; it is inconsistent. We could have come to the same conclusion from observing that in (7) the left side of the second equation is twice the left side of the first equation, but the right side of the second equation is not twice the right side of the first equation. ■

EXAMPLE 4 Consider the linear system

$$\begin{aligned} x + 2y + 3z &= 6 \\ 2x - 3y + 2z &= 14 \\ 3x + y - z &= -2. \end{aligned} \tag{8}$$

To eliminate x, we add (-2) times the first equation to the second one and (-3) times the first equation to the third one, obtaining

$$\begin{aligned} -7y - 4z &= 2 \\ -5y - 10z &= -20. \end{aligned} \tag{9}$$

This is a system of two equations in the unknowns y and z. We multiply the second equation of (9) by $\left(-\frac{1}{5}\right)$, yielding

$$\begin{aligned} -7y - 4z &= 2 \\ y + 2z &= 4, \end{aligned}$$

which we write, by interchanging equations, as

$$\begin{aligned} y + 2z &= 4 \\ -7y - 4z &= 2. \end{aligned} \tag{10}$$

We now eliminate y in (10) by adding 7 times the first equation to the second one, to obtain

$$10z = 30,$$

or

$$z = 3. \tag{11}$$

Substituting this value of z into the first equation of (10), we find that $y = -2$. Then substituting these values of y and z into the first equation of (8), we find that $x = 1$. We observe further that our elimination procedure has actually produced the linear system

$$\begin{aligned} x + 2y + 3z &= 6 \\ y + 2z &= 4 \\ z &= 3, \end{aligned} \tag{12}$$

obtained by using the first equations of (8) and (10) as well as (11). The importance of this procedure is that, although the linear systems (8) and (12) are equivalent, (12) has the advantage that it is easier to solve. ■

EXAMPLE 5 Consider the linear system

$$\begin{aligned} x + 2y - 3z &= -4 \\ 2x + y - 3z &= 4. \end{aligned} \tag{13}$$

Eliminating x, we add (-2) times the first equation to the second equation to get

$$-3y + 3z = 12. \tag{14}$$

We must now solve (14). A solution is

$$y = z - 4,$$

where z can be any real number. Then from the first equation of (13),

$$\begin{aligned} x &= -4 - 2y + 3z \\ &= -4 - 2(z - 4) + 3z \\ &= z + 4. \end{aligned}$$

Thus a solution to the linear system (13) is

$$\begin{aligned} x &= z + 4 \\ y &= z - 4 \\ z &= \text{any real number.} \end{aligned}$$

This means that the linear system (13) has infinitely many solutions. Every time we assign a value to z we obtain another solution to (13). Thus, if $z = 1$, then

$$x = 5, \quad y = -3, \quad \text{and} \quad z = 1$$

is a solution, while if $z = -2$, then

$$x = 2, \quad y = -6, \quad \text{and} \quad z = -2$$

is another solution. ■

These examples suggest that a linear system may have a unique solution, no solution, or infinitely many solutions.

Consider next a linear system of two equations in the unknowns x and y:

$$\begin{aligned} a_1x + a_2y &= c_1 \\ b_1x + b_2y &= c_2. \end{aligned} \tag{15}$$

The graph of each of these equations is a straight line, which we denote by ℓ_1 and ℓ_2, respectively. If $x = s_1$, $y = s_2$ is a solution to the linear system (15), then the point (s_1, s_2) lies on both lines ℓ_1 and ℓ_2. Conversely, if the point (s_1, s_2) lies on both lines ℓ_1 and ℓ_2, then $x = s_1$, $y = s_2$ is a solution to the linear system (15). Thus we are led geometrically to the same three possibilities mentioned previously. See Figure 1.1.

Next, consider a linear system of three equations in the unknowns x, y, and z:

$$\begin{aligned} a_1x + b_1y + c_1z &= d_1 \\ a_2x + b_2y + c_2z &= d_2 \\ a_3x + b_3y + c_3z &= d_3. \end{aligned} \tag{16}$$

The graph of each of these equations is a plane, denoted by P_1, P_2, and P_3, respectively. As in the case of a linear system of two equations in two unknowns,

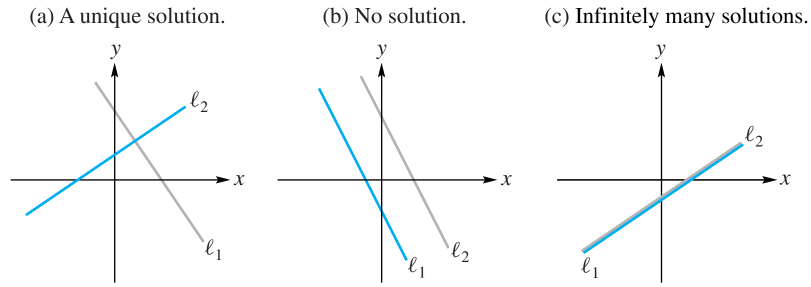


FIGURE 1.1

the linear system in (16) can have infinitely many solutions, a unique solution, or no solution. These situations are illustrated in Figure 1.2. For a more concrete illustration of some of the possible cases, consider that two intersecting walls and the ceiling (planes) of a room intersect in a unique point, a corner of the room, so the linear system has a unique solution. Next, think of the planes as pages of a book. Three pages of a book (held open) intersect in a straight line, the spine. Thus, the linear system has infinitely many solutions. On the other hand, when the book is closed, three pages of a book appear to be parallel and do not intersect, so the linear system has no solution.

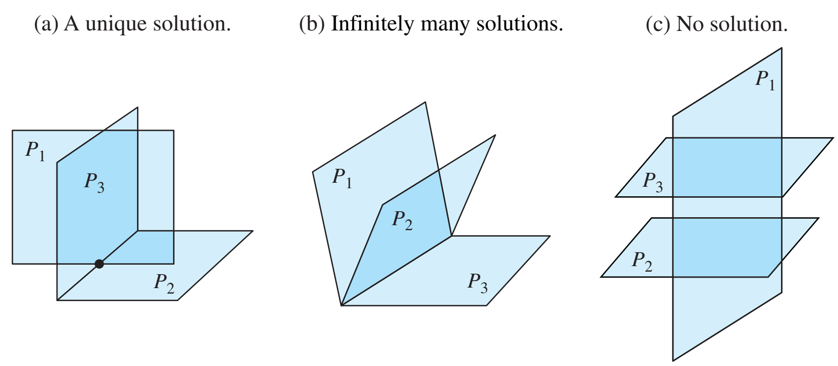


FIGURE 1.2

If we examine the method of elimination more closely, we find that it involves three manipulations that can be performed on a linear system to convert it into an equivalent system. These manipulations are as follows:

1. Interchange the ith and jth equations.
2. Multiply an equation by a nonzero constant.
3. Replace the ith equation by c times the jth equation plus the ith equation, $i \neq j$. That is, replace

$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = b_i$$

by

$$(ca_{j1} + a_{i1})x_1 + (ca_{j2} + a_{i2})x_2 + \cdots + (ca_{jn} + a_{in})x_n = cb_j + b_i.$$

It is not difficult to prove that performing these manipulations on a linear system leads to an equivalent system. The next example proves this for the second type of manipulation. Exercises 24 and 25 prove it for the first and third manipulations, respectively.

EXAMPLE 6

Suppose that the ith equation of the linear system (2) is multiplied by the nonzero constant c, producing the linear system

$$\begin{array}{ccccccc} a_{11}x_1 & + & a_{12}x_2 & + \cdots + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + \cdots + & a_{2n}x_n & = & b_2 \\ \vdots & & \vdots & & \vdots & & \vdots \\ ca_{i1}x_1 & + & ca_{i2}x_2 & + \cdots + & ca_{in}x_n & = & cb_i \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + \cdots + & a_{mn}x_n & = & b_m. \end{array} \tag{17}$$

If $x_1 = s_1, x_2 = s_2, \ldots, x_n = s_n$ is a solution to (2), then it is a solution to all the equations in (17), except possibly to the ith equation. For the ith equation we have

$$c(a_{i1}s_1 + a_{i2}s_2 + \cdots + a_{in}s_n) = cb_i$$

or

$$ca_{i1}s_1 + ca_{i2}s_2 + \cdots + ca_{in}s_n = cb_i.$$

Thus the ith equation of (17) is also satisfied. Hence every solution to (2) is also a solution to (17). Conversely, every solution to (17) also satisfies (2). Hence (2) and (17) are equivalent systems. ■

The following example gives an application leading to a linear system of two equations in three unknowns:

EXAMPLE 7

(Production Planning) A manufacturer makes three different types of chemical products: A, B, and C. Each product must go through two processing machines: X and Y. The products require the following times in machines X and Y:

1. One ton of A requires 2 hours in machine X and 2 hours in machine Y.
2. One ton of B requires 3 hours in machine X and 2 hours in machine Y.
3. One ton of C requires 4 hours in machine X and 3 hours in machine Y.

Machine X is available 80 hours per week, and machine Y is available 60 hours per week. Since management does not want to keep the expensive machines X and Y idle, it would like to know how many tons of each product to make so that the machines are fully utilized. It is assumed that the manufacturer can sell as much of the products as is made.

To solve this problem, we let x_1, x_2, and x_3 denote the number of tons of products A, B, and C, respectively, to be made. The number of hours that machine X will be used is

$$2x_1 + 3x_2 + 4x_3,$$

which must equal 80. Thus we have

$$2x_1 + 3x_2 + 4x_3 = 80.$$

Similarly, the number of hours that machine Y will be used is 60, so we have

$$2x_1 + 2x_2 + 3x_3 = 60.$$

Mathematically, our problem is to find nonnegative values of x_1, x_2, and x_3 so that

$$\begin{aligned} 2x_1 + 3x_2 + 4x_3 &= 80 \\ 2x_1 + 2x_2 + 3x_3 &= 60. \end{aligned}$$

This linear system has infinitely many solutions. Following the method of Example 4, we see that all solutions are given by

$$\begin{aligned} x_1 &= \frac{20 - x_3}{2} \\ x_2 &= 20 - x_3 \\ x_3 &= \text{any real number such that } 0 \leq x_3 \leq 20, \end{aligned}$$

since we must have $x_1 \geq 0$, $x_2 \geq 0$, and $x_3 \geq 0$. When $x_3 = 10$, we have

$$x_1 = 5, \qquad x_2 = 10, \qquad x_3 = 10$$

while

$$x_1 = \tfrac{13}{2}, \qquad x_2 = 13, \qquad x_3 = 7$$

when $x_3 = 7$. The reader should observe that one solution is just as good as the other. There is no best solution unless additional information or restrictions are given. ■

As you have probably already observed, the method of elimination has been described, so far, in general terms. Thus we have not indicated any rules for selecting the unknowns to be eliminated. Before providing a very systematic description of the method of elimination, we introduce in the next section the notion of a matrix. This will greatly simplify our notational problems and will enable us to develop tools to solve many important applied problems.

Key Terms

Linear equation
Solution of a linear equation
Linear system
Unknowns
Inconsistent system
Consistent system
Homogeneous system
Trivial solution
Nontrivial solution
Equivalent systems
Unique solution
No solution
Infinitely many solutions
Manipulations on linear systems
Method of elimination

1.1 Exercises

In Exercises 1 through 14, solve each given linear system by the method of elimination.

1. $\begin{aligned} x + 2y &= 8 \\ 3x - 4y &= 4 \end{aligned}$

2. $\begin{aligned} 2x - 3y + 4z &= -12 \\ x - 2y + z &= -5 \\ 3x + y + 2z &= 1 \end{aligned}$

3. $\begin{aligned} 3x + 2y + z &= 2 \\ 4x + 2y + 2z &= 8 \\ x - y + z &= 4 \end{aligned}$

4. $\begin{aligned} x + y &= 5 \\ 3x + 3y &= 10 \end{aligned}$

5. $\begin{aligned} 2x + 4y + 6z &= -12 \\ 2x - 3y - 4z &= 15 \\ 3x + 4y + 5z &= -8 \end{aligned}$

6. $\begin{aligned} x + y - 2z &= 5 \\ 2x + 3y + 4z &= 2 \end{aligned}$

7. $\begin{aligned} x + 4y - z &= 12 \\ 3x + 8y - 2z &= 4 \end{aligned}$

8. $\begin{aligned} 3x + 4y - z &= 8 \\ 6x + 8y - 2z &= 3 \end{aligned}$

9. $\begin{aligned} x + y + 3z &= 12 \\ 2x + 2y + 6z &= 6 \end{aligned}$

10. $\begin{aligned} x + y &= 1 \\ 2x - y &= 5 \\ 3x + 4y &= 2 \end{aligned}$

11. $\begin{aligned} 2x + 3y &= 13 \\ x - 2y &= 3 \\ 5x + 2y &= 27 \end{aligned}$

12. $\begin{aligned} x - 5y &= 6 \\ 3x + 2y &= 1 \\ 5x + 2y &= 1 \end{aligned}$

13. $\begin{aligned} x + 3y &= -4 \\ 2x + 5y &= -8 \\ x + 3y &= -5 \end{aligned}$

14. $\begin{aligned} 2x + 3y - z &= 6 \\ 2x - y + 2z &= -8 \\ 3x - y + z &= -7 \end{aligned}$

15. Given the linear system

$$\begin{aligned} 2x - y &= 5 \\ 4x - 2y &= t, \end{aligned}$$

(a) Determine a particular value of t so that the system is consistent.

(b) Determine a particular value of t so that the system is inconsistent.

(c) How many different values of t can be selected in part (b)?

16. Given the linear system

$$\begin{aligned} 3x + 4y &= s \\ 6x + 8y &= t, \end{aligned}$$

(a) Determine particular values for s and t so that the system is consistent.

(b) Determine particular values for s and t so that the system is inconsistent.

(c) What relationship between the values of s and t will guarantee that the system is consistent?

17. Given the linear system

$$\begin{aligned} x + 2y &= 10 \\ 3x + (6 + t)y &= 30, \end{aligned}$$

(a) Determine a particular value of t so that the system has infinitely many solutions.

(b) Determine a particular value of t so that the system has a unique solution.

(c) How many different values of t can be selected in part (b)?

18. Is every homogeneous linear system always consistent? Explain.

19. Given the linear system

$$\begin{aligned} 2x + 3y - z &= 0 \\ x - 4y + 5z &= 0, \end{aligned}$$

(a) Verify that $x_1 = 1$, $y_1 = -1$, $z_1 = -1$ is a solution.

(b) Verify that $x_2 = -2$, $y_2 = 2$, $z_2 = 2$ is a solution.

(c) Is $x = x_1 + x_2 = -1$, $y = y_1 + y_2 = 1$, and $z = z_1 + z_2 = 1$ a solution to the linear system?

(d) Is $3x$, $3y$, $3z$, where x, y, and z are as in part (c), a solution to the linear system?

20. Without using the method of elimination, solve the linear system

$$\begin{aligned} 2x + y - 2z &= -5 \\ 3y + z &= 7 \\ z &= 4. \end{aligned}$$

21. Without using the method of elimination, solve the linear system

$$\begin{aligned} 4x &= 8 \\ -2x + 3y &= -1 \\ 3x + 5y - 2z &= 11. \end{aligned}$$

22. Is there a value of r so that $x = 1$, $y = 2$, $z = r$ is a solution to the following linear system? If there is, find it.

$$\begin{aligned} 2x + 3y - z &= 11 \\ x - y + 2z &= -7 \\ 4x + y - 2z &= 12. \end{aligned}$$

23. Is there a value of r so that $x = r$, $y = 2$, $z = 1$ is a solution to the following linear system? If there is, find it.

$$\begin{aligned} 3x - 2z &= 4 \\ x - 4y + z &= -5 \\ -2x + 3y + 2z &= 9. \end{aligned}$$

24. Show that the linear system obtained by interchanging two equations in (2) is equivalent to (2).

25. Show that the linear system obtained by adding a multiple of an equation in (2) to another equation is equivalent to (2).

26. Describe the number of points that simultaneously lie in each of the three planes shown in each part of Figure 1.2.

27. Describe the number of points that simultaneously lie in each of the three planes shown in each part of Figure 1.3.

28. Let C_1 and C_2 be circles in the plane. Describe the number of possible points of intersection of C_1 and C_2. Illustrate each case with a figure.

29. Let S_1 and S_2 be spheres in space. Describe the number of possible points of intersection of S_1 and S_2. Illustrate each case with a figure.

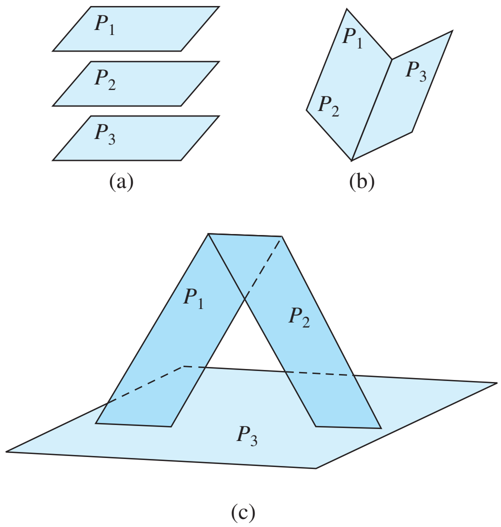


FIGURE 1.3

30. An oil refinery produces low-sulfur and high-sulfur fuel. Each ton of low-sulfur fuel requires 5 minutes in the blending plant and 4 minutes in the refining plant; each ton of high-sulfur fuel requires 4 minutes in the blending plant and 2 minutes in the refining plant. If the blending plant is available for 3 hours and the refining plant is available for 2 hours, how many tons of each type of fuel should be manufactured so that the plants are fully used?

31. A plastics manufacturer makes two types of plastic: regular and special. Each ton of regular plastic requires 2 hours in plant A and 5 hours in plant B; each ton of special plastic requires 2 hours in plant A and 3 hours in plant B. If plant A is available 8 hours per day and plant B is available 15 hours per day, how many tons of each type of plastic can be made daily so that the plants are fully used?

32. A dietician is preparing a meal consisting of foods A, B, and C. Each ounce of food A contains 2 units of protein, 3 units of fat, and 4 units of carbohydrate. Each ounce of food B contains 3 units of protein, 2 units of fat, and 1 unit of carbohydrate. Each ounce of food C contains 3 units of protein, 3 units of fat, and 2 units of carbohydrate. If the meal must provide exactly 25 units of protein, 24 units of fat, and 21 units of carbohydrate, how many ounces of each type of food should be used?

33. A manufacturer makes 2-minute, 6-minute, and 9-minute film developers. Each ton of 2-minute developer requires 6 minutes in plant A and 24 minutes in plant B. Each ton of 6-minute developer requires 12 minutes in plant A and 12 minutes in plant B. Each ton of 9-minute developer requires 12 minutes in plant A and 12 minutes in plant B. If plant A is available 10 hours per day and plant B is available 16 hours per day, how many tons of each type of developer can be produced so that the plants are fully used?

34. Suppose that the three points $(1, -5)$, $(-1, 1)$, and $(2, 7)$ lie on the parabola $p(x) = ax^2 + bx + c$.

(a) Determine a linear system of three equations in three unknowns that must be solved to find a, b, and c.

(b) Solve the linear system obtained in part (a) for a, b, and c.

35. An inheritance of \$24,000 is to be divided among three trusts, with the second trust receiving twice as much as the first trust. The three trusts pay interest annually at the rates of 9%, 10%, and 6%, respectively, and return a total in interest of \$2210 at the end of the first year. How much was invested in each trust?

36. For the software you are using, determine the command that “automatically” solves a linear system of equations.

37. Use the command from Exercise 36 to solve Exercises 3 and 4, and compare the output with the results you obtained by the method of elimination.

38. Solve the linear system

$$\begin{aligned} x + \tfrac{1}{2}y + \tfrac{1}{3}z &= 1 \\ \tfrac{1}{2}x + \tfrac{1}{3}y + \tfrac{1}{4}z &= \tfrac{11}{18} \\ \tfrac{1}{3}x + \tfrac{1}{4}y + \tfrac{1}{5}z &= \tfrac{9}{20} \end{aligned}$$

by using your software. Compare the computed solution with the exact solution $x = \frac{1}{2}$, $y = \frac{1}{3}$, $z = 1$.

39. If your software includes access to a computer algebra system (CAS), use it as follows:

(a) For the linear system in Exercise 38, replace the fraction $\frac{1}{2}$ with its decimal equivalent 0.5. Enter this system into your software and use the appropriate CAS commands to solve the system. Compare the solution with that obtained in Exercise 38.

(b) In some CAS environments you can select the number of digits to be used in the calculations. Perform part (a) with digit choices 2, 4, and 6 to see what influence such selections have on the computed solution.

40. If your software includes access to a CAS and you can select the number of digits used in calculations, do the following: Enter the linear system

$$\begin{aligned} 0.71x + 0.21y &= 0.92 \\ 0.23x + 0.58y &= 0.81 \end{aligned}$$

into the program. Have the software solve the system with digit choices 2, 5, 7, and 12. Briefly discuss any variations in the solutions generated.

1.2 Matrices

If we examine the method of elimination described in Section 1.1, we can make the following observation: Only the numbers in front of the unknowns $x_1, x_2, \ldots, x_n$ and the numbers $b_1, b_2, \ldots, b_m$ on the right side are being changed as we perform the steps in the method of elimination. Thus we might think of looking for a way of writing a linear system without having to carry along the unknowns. Matrices enable us to do this—that is, to write linear systems in a compact form that makes it easier to automate the elimination method by using computer software in order to obtain a fast and efficient procedure for finding solutions. The use of matrices, however, is not merely that of a convenient notation. We now develop operations on matrices and will work with matrices according to the rules they obey; this will enable us to solve systems of linear equations and to handle other computational problems in a fast and efficient manner. Of course, as any good definition should do, the notion of a matrix not only provides a new way of looking at old problems, but also gives rise to a great many new questions, some of which we study in this book.

DEFINITION 1.1

An $m \times n$ **matrix** A is a rectangular array of mn real or complex numbers arranged in m horizontal **rows** and n vertical **columns**:

$$A = \begin{array}{ll} \begin{bmatrix} a_{11} & a_{12} & \cdots & \cdots & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & \cdots & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \cdots & \cdots & \vdots \\ \vdots & \vdots & \cdots & \cdots & a_{ij} & \vdots \\ \vdots & \vdots & \cdots & \cdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & \cdots & \cdots & a_{mn} \end{bmatrix} . & \begin{array}{l} \\ \\ \\ \leftarrow i\text{th row} \\ \\ \\ \end{array} \\ \qquad\qquad\qquad\qquad \uparrow\ j\text{th column} & \end{array} \qquad (1)$$

The ***i*th row** of A is

$$\begin{bmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{bmatrix} \qquad (1 \leq i \leq m);$$

the ***j*th column** of A is

$$\begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix} \qquad (1 \leq j \leq n).$$

We shall say that A is ***m* by *n*** (written as $\boldsymbol{m \times n}$). If $m = n$, we say that A is a **square matrix of order *n***, and that the numbers $a_{11}, a_{22}, \ldots, a_{nn}$ form the **main diagonal** of A. We refer to the number a_{ij}, which is in the ith row and jth column of A, as the ***i*, *j*th element** of A, or the **(*i*, *j*) entry** of A, and we often write (1) as

$$A = \begin{bmatrix} a_{ij} \end{bmatrix}.$$

EXAMPLE 1 Let

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

$$D = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 0 & 1 \\ 3 & -1 & 2 \end{bmatrix}, \qquad E = \begin{bmatrix} 3 \end{bmatrix}, \qquad F = \begin{bmatrix} -1 & 0 & 2 \end{bmatrix}.$$

Then A is a 2×3 matrix with $a_{12} = 2$, $a_{13} = 3$, $a_{22} = 0$, and $a_{23} = 1$; B is a 2×2 matrix with $b_{11} = 1 + i$, $b_{12} = 4i$, $b_{21} = 2 - 3i$, and $b_{22} = -3$; C is a 3×1 matrix with $c_{11} = 1$, $c_{21} = -1$, and $c_{31} = 2$; D is a 3×3 matrix; E is a 1×1 matrix; and F is a 1×3 matrix. In D, the elements $d_{11} = 1$, $d_{22} = 0$, and $d_{33} = 2$ form the main diagonal. ■

For convenience, we focus much of our attention in the illustrative examples and exercises in Chapters 1–6 on matrices and expressions containing only real numbers. Complex numbers make a brief appearance in Chapter 7. An introduction to complex numbers, their properties, and examples and exercises showing how complex numbers are used in linear algebra may be found in Appendix B.

An $n \times 1$ matrix is also called an ***n*-vector** and is denoted by lowercase boldface letters. When n is understood, we refer to n-vectors merely as **vectors**. Vectors are discussed at length in Section 4.1.

EXAMPLE 2

$$\mathbf{u} = \begin{bmatrix} 1 \\ 2 \\ -1 \\ 0 \end{bmatrix} \text{ is a 4-vector and } \mathbf{v} = \begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix} \text{ is a 3-vector.}$$ ■

The n-vector all of whose entries are zero is denoted by $\mathbf{0}$.

Observe that if A is an $n \times n$ matrix, then the rows of A are $1 \times n$ matrices and the columns of A are $n \times 1$ matrices. The set of all n-vectors with real entries is denoted by R^n. Similarly, the set of all n-vectors with complex entries is denoted by C^n. As we have already pointed out, in the first six chapters of this book we work almost entirely with vectors in R^n.

EXAMPLE 3 **(Tabular Display of Data)** The following matrix gives the airline distances between the indicated cities (in statute miles):

$$\begin{array}{l} \\ \text{London} \\ \text{Madrid} \\ \text{New York} \\ \text{Tokyo} \end{array} \begin{array}{c} \begin{array}{cccc} \text{London} & \text{Madrid} & \text{New York} & \text{Tokyo} \end{array} \\ \begin{bmatrix} 0 & 785 & 3469 & 5959 \\ 785 & 0 & 3593 & 6706 \\ 3469 & 3593 & 0 & 6757 \\ 5959 & 6706 & 6757 & 0 \end{bmatrix} \end{array}$$ ■

EXAMPLE 4 **(Production)** Suppose that a manufacturer has four plants, each of which makes three products. If we let a_{ij} denote the number of units of product i made by plant

j in one week, then the 3×4 matrix

$$\begin{array}{c} \\ \text{Product 1} \\ \text{Product 2} \\ \text{Product 3} \end{array} \begin{array}{c} \begin{array}{cccc} \text{Plant 1} & \text{Plant 2} & \text{Plant 3} & \text{Plant 4} \end{array} \\ \begin{bmatrix} 560 & 360 & 380 & 0 \\ 340 & 450 & 420 & 80 \\ 280 & 270 & 210 & 380 \end{bmatrix} \end{array}$$

gives the manufacturer's production for the week. For example, plant 2 makes 270 units of product 3 in one week. ■

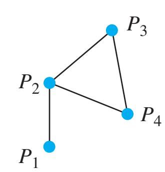
EXAMPLE 5 The windchill table that follows is a matrix.

	°F					
mph	15	10	5	0	−5	−10
5	12	7	0	−5	−10	−15
10	−3	−9	−15	−22	−27	−34
15	−11	−18	−25	−31	−38	−45
20	−17	−24	−31	−39	−46	−53

A combination of air temperature and wind speed makes a body feel colder than the actual temperature. For example, when the temperature is 10°F and the wind is 15 miles per hour, this causes a body heat loss equal to that when the temperature is −18°F with no wind. ■

EXAMPLE 6 By a **graph** we mean a set of points called **nodes** or **vertices**, some of which are connected by lines called **edges**. The nodes are usually labeled as $P_1, P_2, \ldots, P_k$, and for now we allow an edge to be traveled in either direction. One mathematical representation of a graph is constructed from a table. For example, the following table represents the graph shown:

	P_1	P_2	P_3	P_4
P_1	0	1	0	0
P_2	1	0	1	1
P_3	0	1	0	1
P_4	0	1	1	0



The (i, j) entry $= 1$ if there is an edge connecting vertex P_i to vertex P_j; otherwise, the (i, j) entry $= 0$. The **incidence matrix** $\boldsymbol{A}$ is the $k \times k$ matrix obtained by omitting the row and column labels from the preceding table. The incidence matrix for the corresponding graph is

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix}.$$

■

Internet search engines use matrices to keep track of the locations of information, the type of information at a location, keywords that appear in the information, and even the way websites link to one another. A large measure of the effectiveness of the search engine Google© is the manner in which matrices are used to determine which sites are referenced by other sites. That is, instead of directly keeping track of the information content of an actual web page or of an individual search topic, Google's matrix structure focuses on finding web pages that match the search topic, and then presents a list of such pages in the order of their "importance."

Suppose that there are n accessible web pages during a certain month. A simple way to view a matrix that is part of Google's scheme is to imagine an $n \times n$ matrix A, called the "connectivity matrix," that initially contains all zeros. To build the connections, proceed as follows. When you detect that website j links to website i, set entry a_{ij} equal to one. Since n is quite large, in the billions, most entries of the connectivity matrix A are zero. (Such a matrix is called sparse.) If row i of A contains many ones, then there are many sites linking to site i. Sites that are linked to by many other sites are considered more "important" (or to have a higher rank) by the software driving the Google search engine. Such sites would appear near the top of a list returned by a Google search on topics related to the information on site i. Since Google updates its connectivity matrix about every month, n increases over time and new links and sites are adjoined to the connectivity matrix.

In Chapter 8 we elaborate a bit on the fundamental technique used for ranking sites and give several examples related to the matrix concepts involved. Further information can be found in the following sources:

1. Berry, Michael W., and Murray Browne. *Understanding Search Engines—Mathematical Modeling and Text Retrieval*, 2d ed. Philadelphia: Siam, 2005.

2. www.google.com/technology/index.html

3. Moler, Cleve. "The World's Largest Matrix Computation: Google's PageRank Is an Eigenvector of a Matrix of Order 2.7 Billion," MATLAB *News and Notes*, October 2002, pp. 12–13.

Whenever a new object is introduced in mathematics, we must determine when two such objects are equal. For example, in the set of all rational numbers, the numbers $\frac{2}{3}$ and $\frac{4}{6}$ are called equal, although they have different representations. What we have in mind is the definition that a/b equals c/d when $ad = bc$. Accordingly, we now have the following definition:

DEFINITION 1.2 Two $m \times n$ matrices $A = \begin{bmatrix} a_{ij} \end{bmatrix}$ and $B = \begin{bmatrix} b_{ij} \end{bmatrix}$ are **equal** if they agree entry by entry, that is, if $a_{ij} = b_{ij}$ for $i = 1, 2, \dots, m$ and $j = 1, 2, \dots, n$.

EXAMPLE 7 The matrices

$$A = \begin{bmatrix} 1 & 2 & -1 \\ 2 & -3 & 4 \\ 0 & -4 & 5 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 2 & w \\ 2 & x & 4 \\ y & -4 & z \end{bmatrix}$$

are equal if and only if $w = -1$, $x = -3$, $y = 0$, and $z = 5$. ■

Matrix Operations

We next define a number of operations that will produce new matrices out of given matrices. When we are dealing with linear systems, for example, this will enable us to manipulate the matrices that arise and to avoid writing down systems over and over again. These operations and manipulations are also useful in other applications of matrices.

Matrix Addition

DEFINITION 1.3 If $A = \begin{bmatrix} a_{ij} \end{bmatrix}$ and $B = \begin{bmatrix} b_{ij} \end{bmatrix}$ are both $m \times n$ matrices, then the **sum** $A+B$ is an $m \times n$ matrix $C = \begin{bmatrix} c_{ij} \end{bmatrix}$ defined by $c_{ij} = a_{ij} + b_{ij}$, $i = 1, 2, \dots, m$; $j = 1, 2, \dots, n$. Thus, to obtain the sum of A and B, we merely add corresponding entries.

EXAMPLE 8 Let

$$A = \begin{bmatrix} 1 & -2 & 3 \\ 2 & -1 & 4 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 0 & 2 & 1 \\ 1 & 3 & -4 \end{bmatrix}.$$

Then

$$A + B = \begin{bmatrix} 1+0 & -2+2 & 3+1 \\ 2+1 & -1+3 & 4+(-4) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 4 \\ 3 & 2 & 0 \end{bmatrix}.$$

■

EXAMPLE 9 **(Production)** A manufacturer of a certain product makes three models, A, B, and C. Each model is partially made in factory F_1 in Taiwan and then finished in factory F_2 in the United States. The total cost of each product consists of the manufacturing cost and the shipping cost. Then the costs at each factory (in dollars) can be described by the 3×2 matrices F_1 and F_2:

$$\begin{array}{cc} & \begin{array}{cc} \text{Manufacturing cost} & \text{Shipping cost} \end{array} \\ F_1 = & \begin{bmatrix} 32 & 40 \\ 50 & 80 \\ 70 & 20 \end{bmatrix} \begin{array}{l} \text{Model A} \\ \text{Model B} \\ \text{Model C} \end{array} \end{array}$$

$$\begin{array}{cc} & \begin{array}{cc} \text{Manufacturing cost} & \text{Shipping cost} \end{array} \\ F_2 = & \begin{bmatrix} 40 & 60 \\ 50 & 50 \\ 130 & 20 \end{bmatrix} \begin{array}{l} \text{Model A} \\ \text{Model B} \\ \text{Model C} \end{array} \end{array}.$$

The matrix $F_1 + F_2$ gives the total manufacturing and shipping costs for each product. Thus the total manufacturing and shipping costs of a model C product are \$200 and \$40, respectively. ■

If $\mathbf{x}$ is an n-vector, then it is easy to show that $\mathbf{x} + \mathbf{0} = \mathbf{x}$, where $\mathbf{0}$ is the n-vector all of whose entries are zero. (See Exercise 16.)

It should be noted that the sum of the matrices A and B is defined only when A and B have the same number of rows and the same number of columns, that is, only when A and B are of the same size.

We now make the convention that when $A + B$ is written, both A and B are of the same size.

The basic properties of matrix addition are considered in the next section and are similar to those satisfied by the real numbers.

Scalar Multiplication

DEFINITION 1.4

If $A = \begin{bmatrix} a_{ij} \end{bmatrix}$ is an $m \times n$ matrix and r is a real number, then the **scalar multiple** of A by r, rA, is the $m \times n$ matrix $C = \begin{bmatrix} c_{ij} \end{bmatrix}$, where $c_{ij} = ra_{ij}$, $i = 1, 2, \dots, m$ and $j = 1, 2, \dots, n$; that is, the matrix C is obtained by multiplying each entry of A by r.

EXAMPLE 10

We have

$$-2\begin{bmatrix} 4 & -2 & -3 \\ 7 & -3 & 2 \end{bmatrix} = \begin{bmatrix} (-2)(4) & (-2)(-2) & (-2)(-3) \\ (-2)(7) & (-2)(-3) & (-2)(2) \end{bmatrix}$$
$$= \begin{bmatrix} -8 & 4 & 6 \\ -14 & 6 & -4 \end{bmatrix}.$$ ■

Thus far, addition of matrices has been defined for only two matrices. Our work with matrices will call for adding more than two matrices. Theorem 1.1 in Section 1.4 shows that addition of matrices satisfies the associative property: $A + (B + C) = (A + B) + C$.

If A and B are $m \times n$ matrices, we write $A + (-1)B$ as $A - B$ and call this the **difference between A and B**.

EXAMPLE 11

Let

$$A = \begin{bmatrix} 2 & 3 & -5 \\ 4 & 2 & 1 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 2 & -1 & 3 \\ 3 & 5 & -2 \end{bmatrix}.$$

Then

$$A - B = \begin{bmatrix} 2-2 & 3+1 & -5-3 \\ 4-3 & 2-5 & 1+2 \end{bmatrix} = \begin{bmatrix} 0 & 4 & -8 \\ 1 & -3 & 3 \end{bmatrix}.$$ ■

Application

Vectors in R^n can be used to handle large amounts of data. Indeed, a number of computer software products, notably, MATLAB®, make extensive use of vectors. The following example illustrates these ideas:

EXAMPLE 12

(Inventory Control) Suppose that a store handles 100 different items. The inventory on hand at the beginning of the week can be described by the inventory vector $\mathbf{u}$ in R^{100}. The number of items sold at the end of the week can be described by the 100-vector $\mathbf{v}$, and the vector

$$\mathbf{u} - \mathbf{v}$$

represents the inventory at the end of the week. If the store receives a new shipment of goods, represented by the 100-vector $\mathbf{w}$, then its new inventory would be

$$\mathbf{u} - \mathbf{v} + \mathbf{w}.$$ ■

We shall sometimes use the **summation notation**, and we now review this useful and compact notation.

By $\sum_{i=1}^{n} a_i$ we mean $a_1 + a_2 + \cdots + a_n$. The letter i is called the **index of summation**; it is a dummy variable that can be replaced by another letter. Hence we can write

$$\sum_{i=1}^{n} a_i = \sum_{j=1}^{n} a_j = \sum_{k=1}^{n} a_k.$$

Thus

$$\sum_{i=1}^{4} a_i = a_1 + a_2 + a_3 + a_4.$$

The summation notation satisfies the following properties:

1. $\sum_{i=1}^{n}(r_i + s_i)a_i = \sum_{i=1}^{n} r_i a_i + \sum_{i=1}^{n} s_i a_i$

2. $\sum_{i=1}^{n} c(r_i a_i) = c\sum_{i=1}^{n} r_i a_i$

3. $\sum_{j=1}^{n}\left(\sum_{i=1}^{m} a_{ij}\right) = \sum_{i=1}^{m}\left(\sum_{j=1}^{n} a_{ij}\right)$

Property 3 can be interpreted as follows: The left side is obtained by adding all the entries in each column and then adding all the resulting numbers. The right side is obtained by adding all the entries in each row and then adding all the resulting numbers.

If $A_1, A_2, \ldots, A_k$ are $m \times n$ matrices and $c_1, c_2, \ldots, c_k$ are real numbers, then an expression of the form

$$c_1 A_1 + c_2 A_2 + \cdots + c_k A_k \tag{2}$$

is called a **linear combination** of $A_1, A_2, \ldots, A_k$, and $c_1, c_2, \ldots, c_k$ are called **coefficients**.

The linear combination in Equation (2) can also be expressed in summation notation as

$$\sum_{i=1}^{k} c_i A_i = c_1 A_1 + c_2 A_2 + \cdots + c_k A_k.$$

EXAMPLE 13 The following are linear combinations of matrices:

$$3\begin{bmatrix} 0 & -3 & 5 \\ 2 & 3 & 4 \\ 1 & -2 & -3 \end{bmatrix} - \frac{1}{2}\begin{bmatrix} 5 & 2 & 3 \\ 6 & 2 & 3 \\ -1 & -2 & 3 \end{bmatrix},$$

$$2\begin{bmatrix} 3 & -2 \end{bmatrix} - 3\begin{bmatrix} 5 & 0 \end{bmatrix} + 4\begin{bmatrix} -2 & 5 \end{bmatrix},$$

$$-0.5\begin{bmatrix} 1 \\ -4 \\ -6 \end{bmatrix} + 0.4\begin{bmatrix} 0.1 \\ -4 \\ 0.2 \end{bmatrix}.$$

Using scalar multiplication and matrix addition, we can compute each of these linear combinations. Verify that the results of such computations are, respectively,

$$\begin{bmatrix} -\frac{5}{2} & -10 & \frac{27}{2} \\ 3 & 8 & \frac{21}{2} \\ \frac{7}{2} & -5 & -\frac{21}{2} \end{bmatrix}, \qquad \begin{bmatrix} -17 & 16 \end{bmatrix}, \qquad \text{and} \qquad \begin{bmatrix} -0.46 \\ 0.4 \\ 3.08 \end{bmatrix}.$$

■

EXAMPLE 14

Let

$$\mathbf{p} = \begin{bmatrix} 18.95 \\ 14.75 \\ 8.60 \end{bmatrix}$$

be a 3-vector that represents the current prices of three items at a store. Suppose that the store announces a sale so that the price of each item is reduced by 20%.

(a) Determine a 3-vector that gives the price changes for the three items.

(b) Determine a 3-vector that gives the new prices of the items.

Solution

(a) Since each item is reduced by 20%, the 3-vector

$$-0.20\mathbf{p} = \begin{bmatrix} (-0.20)18.95 \\ (-0.20)14.75 \\ (-0.20)8.60 \end{bmatrix} = \begin{bmatrix} -3.79 \\ -2.95 \\ -1.72 \end{bmatrix} = -\begin{bmatrix} 3.79 \\ 2.95 \\ 1.72 \end{bmatrix}$$

gives the price changes for the three items.

(b) The new prices of the items are given by the expression

$$\mathbf{p} - 0.20\mathbf{p} = \begin{bmatrix} 18.95 \\ 14.75 \\ 8.60 \end{bmatrix} - \begin{bmatrix} 3.79 \\ 2.95 \\ 1.72 \end{bmatrix} = \begin{bmatrix} 15.16 \\ 11.80 \\ 6.88 \end{bmatrix}.$$

Observe that this expression can also be written as

$$\mathbf{p} - 0.20\mathbf{p} = 0.80\mathbf{p}.$$

■

The next operation on matrices is useful in a number of situations.

DEFINITION 1.5

If $A = \begin{bmatrix} a_{ij} \end{bmatrix}$ is an $m \times n$ matrix, then the **transpose** of A, $A^T = \begin{bmatrix} a_{ij}^T \end{bmatrix}$, is the $n \times m$ matrix defined by $a_{ij}^T = a_{ji}$. Thus the transpose of A is obtained from A by interchanging the rows and columns of A.

EXAMPLE 15

Let

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

$$D = \begin{bmatrix} 3 & -5 & 1 \end{bmatrix}, \qquad E = \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix}.$$

Then

$$A^T = \begin{bmatrix} 4 & 0 \\ -2 & 5 \\ 3 & -2 \end{bmatrix}, \qquad B^T = \begin{bmatrix} 6 & 3 & 0 \\ 2 & -1 & 4 \\ -4 & 2 & 3 \end{bmatrix},$$

$$C^T = \begin{bmatrix} 5 & -3 & 2 \\ 4 & 2 & -3 \end{bmatrix}, \quad D^T = \begin{bmatrix} 3 \\ -5 \\ 1 \end{bmatrix}, \quad \text{and} \quad E^T = \begin{bmatrix} 2 & -1 & 3 \end{bmatrix}.$$

■

Key Terms

Matrix
Rows
Columns
Size of a matrix
Square matrix
Main diagonal
Element or entry of a matrix
Equal matrices
n-vector (or vector)
R^n, C^n
0, zero vector
Google
Matrix addition
Scalar multiple
Difference of matrices
Summation notation
Index of summation
Linear combination
Coefficients
Transpose

1.2 Exercises

1. Let

$$A = \begin{bmatrix} 2 & -3 & 5 \\ 6 & -5 & 4 \end{bmatrix}, \quad B = \begin{bmatrix} 4 \\ -3 \\ 5 \end{bmatrix},$$

and

$$C = \begin{bmatrix} 7 & 3 & 2 \\ -4 & 3 & 5 \\ 6 & 1 & -1 \end{bmatrix}.$$

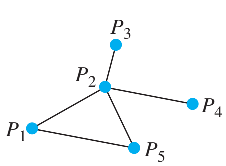
(a) What is a_{12}, a_{22}, a_{23}?

(b) What is b_{11}, b_{31}?

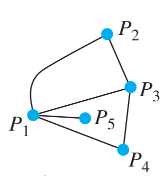
(c) What is c_{13}, c_{31}, c_{33}?

2. Determine the incidence matrix associated with each of the following graphs:

(a)



(b)



3. For each of the following incidence matrices, construct a graph. Label the vertices $P_1, P_2, \ldots, P_5$.

(a) $A = \begin{bmatrix} 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 \end{bmatrix}$

(b) $A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix}$

4. If

$$\begin{bmatrix} a+b & c+d \\ c-d & a-b \end{bmatrix} = \begin{bmatrix} 4 & 6 \\ 10 & 2 \end{bmatrix},$$

find a, b, c, and d.

5. If

$$\begin{bmatrix} a+2b & 2a-b \\ 2c+d & c-2d \end{bmatrix} = \begin{bmatrix} 4 & -2 \\ 4 & -3 \end{bmatrix},$$

find a, b, c, and d.

In Exercises 6 through 9, let

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 4 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 \\ 2 & 1 \\ 3 & 2 \end{bmatrix},$$

$$C = \begin{bmatrix} 3 & -1 & 3 \\ 4 & 1 & 5 \\ 2 & 1 & 3 \end{bmatrix}, \quad D = \begin{bmatrix} 3 & -2 \\ 2 & 4 \end{bmatrix},$$

$$E = \begin{bmatrix} 2 & -4 & 5 \\ 0 & 1 & 4 \\ 3 & 2 & 1 \end{bmatrix}, \quad F = \begin{bmatrix} -4 & 5 \\ 2 & 3 \end{bmatrix},$$

$$\text{and} \quad O = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

6. If possible, compute the indicated linear combination:
(a) $C + E$ and $E + C$ **(b)** $A + B$
(c) $D - F$ **(d)** $-3C + 5O$
(e) $2C - 3E$ **(f)** $2B + F$

7. If possible, compute the indicated linear combination:
(a) $3D + 2F$ **(b)** $3(2A)$ and $6A$
(c) $3A + 2A$ and $5A$
(d) $2(D + F)$ and $2D + 2F$
(e) $(2 + 3)D$ and $2D + 3D$
(f) $3(B + D)$

8. If possible, compute the following:
(a) A^T and $(A^T)^T$
(b) $(C + E)^T$ and $C^T + E^T$
(c) $(2D + 3F)^T$ **(d)** $D - D^T$
(e) $2A^T + B$ **(f)** $(3D - 2F)^T$

9. If possible, compute the following:
(a) $(2A)^T$ **(b)** $(A - B)^T$
(c) $(3B^T - 2A)^T$ **(d)** $(3A^T - 5B^T)^T$
(e) $(-A)^T$ and $-(A^T)$ **(f)** $(C + E + F^T)^T$

10. Is the matrix $\begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}$ a linear combination of the matrices $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$? Justify your answer.

11. Is the matrix $\begin{bmatrix} 4 & 1 \\ 0 & -3 \end{bmatrix}$ a linear combination of the matrices $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$? Justify your answer.

12. Let

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 6 & -2 & 3 \\ 5 & 2 & 4 \end{bmatrix} \quad \text{and} \quad I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

If λ is a real number, compute $\lambda I_3 - A$.

13. If A is an $n \times n$ matrix, what are the entries on the main diagonal of $A - A^T$? Justify your answer.

14. Explain why every incidence matrix A associated with a graph is the same as A^T.

15. Let the $n \times n$ matrix A be equal to A^T. Briefly describe the pattern of the entries in A.

16. If $\mathbf{x}$ is an n-vector, show that $\mathbf{x} + \mathbf{0} = \mathbf{x}$.

17. Show that the summation notation satisfies the following properties:

(a) $\displaystyle\sum_{i=1}^{n} (r_i + s_i)a_i = \sum_{i=1}^{n} r_i a_i + \sum_{i=1}^{n} s_i a_i$

(b) $\displaystyle\sum_{i=1}^{n} c(r_i a_i) = c\left(\sum_{i=1}^{n} r_i a_i\right)$

18. Show that $\displaystyle\sum_{i=1}^{n}\left(\sum_{j=1}^{m} a_{ij}\right) = \sum_{j=1}^{m}\left(\sum_{i=1}^{n} a_{ij}\right)$.

19. Identify the following expressions as true or false. If true, prove the result; if false, give a counterexample.

(a) $\displaystyle\sum_{i=1}^{n} (a_i + 1) = \left(\sum_{i=1}^{n} a_i\right) + n$

(b) $\displaystyle\sum_{i=1}^{n}\left(\sum_{j=1}^{m} 1\right) = mn$

(c) $\displaystyle\sum_{j=1}^{m}\left(\sum_{i=1}^{n} a_i b_j\right) = \left[\sum_{i=1}^{n} a_i\right]\left[\sum_{j=1}^{m} b_j\right]$

20. A large steel manufacturer, who has 2000 employees, lists each employee's salary as a component of a vector $\mathbf{u}$ in R^{2000}. If an 8% across-the-board salary increase has been approved, find an expression involving $\mathbf{u}$ that gives all the new salaries.

21. A brokerage firm records the high and low values of the price of IBM stock each day. The information for a given week is presented in two vectors, $\mathbf{t}$ and $\mathbf{b}$, in R^5, showing the high and low values, respectively. What expression gives the average daily values of the price of IBM stock for the entire 5-day week?

22. For the software you are using, determine the commands to enter a matrix, add matrices, multiply a scalar times a matrix, and obtain the transpose of a matrix for matrices with numerical entries. Practice the commands, using the linear combinations in Example 13.

23. Determine whether the software you are using includes a computer algebra system (CAS), and if it does, do the following:

(a) Find the command for entering a symbolic matrix. (This command may be different than that for entering a numeric matrix.)

(b) Enter several symbolic matrices like

$$A = \begin{bmatrix} r & s & t \\ u & v & w \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix}.$$

Compute expressions like $A + B$, $2A$, $3A + B$, $A - 2B$, $A^T + B^T$, etc. (In some systems you must explicitly indicate scalar multiplication with an asterisk.)

24. For the software you are using, determine whether there is a command that will display a graph for an incidence matrix. If there is, display the graphs for the incidence matrices in Exercise 3 and compare them with those that you drew by hand. Explain why the computer-generated graphs need not be identical to those you drew by hand.

1.3 Matrix Multiplication

In this section we introduce the operation of matrix multiplication. Unlike matrix addition, matrix multiplication has some properties that distinguish it from multiplication of real numbers.

DEFINITION 1.6

The **dot product**, or **inner product**, of the n-vectors in R^n

$$\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

is defined as

$$\mathbf{a} \cdot \mathbf{b} = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n = \sum_{i=1}^{n} a_i b_i .^*$$

The dot product is an important operation that will be used here and in later sections.

EXAMPLE 1

The dot product of

$$\mathbf{u} = \begin{bmatrix} 1 \\ -2 \\ 3 \\ 4 \end{bmatrix} \quad \text{and} \quad \mathbf{v} = \begin{bmatrix} 2 \\ 3 \\ -2 \\ 1 \end{bmatrix}$$

is

$$\mathbf{u} \cdot \mathbf{v} = (1)(2) + (-2)(3) + (3)(-2) + (4)(1) = -6. \qquad \blacksquare$$

EXAMPLE 2

Let $\mathbf{a} = \begin{bmatrix} x \\ 2 \\ 3 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} 4 \\ 1 \\ 2 \end{bmatrix}$. If $\mathbf{a} \cdot \mathbf{b} = -4$, find x.

Solution

We have

$$\begin{aligned} \mathbf{a} \cdot \mathbf{b} = 4x + 2 + 6 &= -4 \\ 4x + 8 &= -4 \\ x &= -3. \end{aligned} \qquad \blacksquare$$

*The dot product of vectors in C^n is defined in Appendix B.2.

EXAMPLE 3 **(Computing a Course Average)** Suppose that an instructor uses four grades to determine a student's course average: quizzes, two hourly exams, and a final exam. These are weighted as 10%, 30%, 30%, and 30%, respectively. If a student's scores are 78, 84, 62, and 85, respectively, we can compute the course average by letting

$$\mathbf{w} = \begin{bmatrix} 0.10 \\ 0.30 \\ 0.30 \\ 0.30 \end{bmatrix} \quad \text{and} \quad \mathbf{g} = \begin{bmatrix} 78 \\ 84 \\ 62 \\ 85 \end{bmatrix}$$

and computing

$$\mathbf{w} \cdot \mathbf{g} = (0.10)(78) + (0.30)(84) + (0.30)(62) + (0.30)(85) = 77.1.$$

Thus the student's course average is 77.1. ■

■ Matrix Multiplication

DEFINITION 1.7 If $A = \begin{bmatrix} a_{ij} \end{bmatrix}$ is an $m \times p$ matrix and $B = \begin{bmatrix} b_{ij} \end{bmatrix}$ is a $p \times n$ matrix, then the **product** of A and B, denoted AB, is the $m \times n$ matrix $C = \begin{bmatrix} c_{ij} \end{bmatrix}$, defined by

$$\begin{aligned} c_{ij} &= a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{ip}b_{pj} \\ &= \sum_{k=1}^{p} a_{ik}b_{kj} \quad (1 \le i \le m, 1 \le j \le n). \end{aligned} \tag{1}$$

Equation (1) says that the i, jth element in the product matrix is the dot product of the transpose of the ith row, $\text{row}_i(A)$—that is, $(\text{row}_i(A))^T$—and the jth column, $\text{col}_j(B)$, of B; this is shown in Figure 1.4.

$$\begin{array}{r} \\ \\ \\ \text{row}_i(A) \\ \\ \\ \end{array} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{ip} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mp} \end{bmatrix} \begin{array}{c} \text{col}_j(B) \\ \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1j} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2j} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots & & \vdots \\ b_{p1} & b_{p2} & \cdots & b_{pj} & \cdots & b_{pn} \end{bmatrix} \end{array}$$

$$= \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & c_{ij} & \vdots \\ c_{m1} & c_{m2} & \cdots & c_{mn} \end{bmatrix}$$

$$(\text{row}_i(A))^T \cdot \text{col}_j(B) = \sum_{k=1}^{p} a_{ik}b_{kj} = c_{ij}$$

FIGURE 1.4

Observe that the product of A and B is defined only when the number of rows of B is exactly the same as the number of columns of A, as indicated in Figure 1.5.

EXAMPLE 4 Let

$$A = \begin{bmatrix} 1 & 2 & -1 \\ 3 & 1 & 4 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} -2 & 5 \\ 4 & -3 \\ 2 & 1 \end{bmatrix}.$$

$$\underset{m \times p}{A} \qquad \underset{p \times n}{B} = \underset{m \times n}{AB}$$

the same

size of AB

FIGURE 1.5

Then

$$AB = \begin{bmatrix} (1)(-2) + (2)(4) + (-1)(2) & (1)(5) + (2)(-3) + (-1)(1) \\ (3)(-2) + (1)(4) + (4)(2) & (3)(5) + (1)(-3) + (4)(1) \end{bmatrix}$$
$$= \begin{bmatrix} 4 & -2 \\ 6 & 16 \end{bmatrix}.$$
■

EXAMPLE 5 Let

$$A = \begin{bmatrix} 1 & -2 & 3 \\ 4 & 2 & 1 \\ 0 & 1 & -2 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 4 \\ 3 & -1 \\ -2 & 2 \end{bmatrix}.$$

Compute the (3, 2) entry of AB.

Solution

If $AB = C$, then the (3, 2) entry of AB is c_{32}, which is $(\text{row}_3(A))^T \cdot \text{col}_2(B)$. We now have

$$(\text{row}_3(A))^T \cdot \text{col}_2(B) = \begin{bmatrix} 0 \\ 1 \\ -2 \end{bmatrix} \cdot \begin{bmatrix} 4 \\ -1 \\ 2 \end{bmatrix} = -5.$$
■

EXAMPLE 6 Let

$$A = \begin{bmatrix} 1 & x & 3 \\ 2 & -1 & 1 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 2 \\ 4 \\ y \end{bmatrix}.$$

If $AB = \begin{bmatrix} 12 \\ 6 \end{bmatrix}$, find x and y.

Solution

We have

$$AB = \begin{bmatrix} 1 & x & 3 \\ 2 & -1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 4 \\ y \end{bmatrix} = \begin{bmatrix} 2 + 4x + 3y \\ 4 - 4 + y \end{bmatrix} = \begin{bmatrix} 12 \\ 6 \end{bmatrix}.$$

Then

$$\begin{aligned} 2 + 4x + 3y &= 12 \\ y &= 6, \end{aligned}$$

so $x = -2$ and $y = 6$. ■

The basic properties of matrix multiplication will be considered in the next section. However, multiplication of matrices requires much more care than their addition, since the algebraic properties of matrix multiplication differ from those satisfied by the real numbers. Part of the problem is due to the fact that AB is defined only when the number of columns of A is the same as the number of rows of B. Thus, if A is an $m \times p$ matrix and B is a $p \times n$ matrix, then AB is an $m \times n$ matrix. What about BA? Four different situations may occur:

1. BA may not be defined; this will take place if $n \neq m$.
2. If BA is defined, which means that $m = n$, then BA is $p \times p$ while AB is $m \times m$; thus, if $m \neq p$, AB and BA are of different sizes.
3. If AB and BA are both of the same size, they may be equal.
4. If AB and BA are both of the same size, they may be unequal.

EXAMPLE 7 If A is a 2×3 matrix and B is a 3×4 matrix, then AB is a 2×4 matrix while BA is undefined. ■

EXAMPLE 8 Let A be 2×3 and let B be 3×2. Then AB is 2×2 while BA is 3×3. ■

EXAMPLE 9 Let

$$A = \begin{bmatrix} 1 & 2 \\ -1 & 3 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}.$$

Then

$$AB = \begin{bmatrix} 2 & 3 \\ -2 & 2 \end{bmatrix} \quad \text{while} \quad BA = \begin{bmatrix} 1 & 7 \\ -1 & 3 \end{bmatrix}.$$

Thus $AB \neq BA$. ■

One might ask why matrix equality and matrix addition are defined in such a natural way, while matrix multiplication appears to be much more complicated. Only a thorough understanding of the composition of functions and the relationship that exists between matrices and what are called linear transformations would show that the definition of multiplication given previously is the natural one. These topics are covered later in the book. For now, Example 10 provides a motivation for the definition of matrix multiplication.

EXAMPLE 10 **(Ecology)** Pesticides are sprayed on plants to eliminate harmful insects. However, some of the pesticide is absorbed by the plant. The pesticides are absorbed by herbivores when they eat the plants that have been sprayed. To determine the amount of pesticide absorbed by a herbivore, we proceed as follows. Suppose that we have three pesticides and four plants. Let a_{ij} denote the amount of pesticide i (in milligrams) that has been absorbed by plant j. This information can be represented by the matrix

$$A = \begin{array}{c} \begin{array}{cccc} \text{Plant 1} & \text{Plant 2} & \text{Plant 3} & \text{Plant 4} \end{array} \\ \begin{bmatrix} 2 & 3 & 4 & 3 \\ 3 & 2 & 2 & 5 \\ 4 & 1 & 6 & 4 \end{bmatrix} \end{array} \begin{array}{l} \\ \text{Pesticide 1} \\ \text{Pesticide 2} \\ \text{Pesticide 3} \end{array}.$$

Now suppose that we have three herbivores, and let b_{ij} denote the number of plants of type i that a herbivore of type j eats per month. This information can be represented by the matrix

$$B = \begin{array}{c} \begin{array}{ccc} \text{Herbivore 1} & \text{Herbivore 2} & \text{Herbivore 3} \end{array} \\ \left[\begin{array}{ccc} 20 & 12 & 8 \\ 28 & 15 & 15 \\ 30 & 12 & 10 \\ 40 & 16 & 20 \end{array}\right] \end{array} \begin{array}{l} \\ \text{Plant 1} \\ \text{Plant 2} \\ \text{Plant 3} \\ \text{Plant 4} \end{array}.$$

The (i, j) entry in AB gives the amount of pesticide of type i that animal j has absorbed. Thus, if $i = 2$ and $j = 3$, the (2, 3) entry in AB is

$$\begin{aligned}(\mathrm{row}_2(A))^T \cdot \mathrm{col}_3(B) &= 3(8) + 2(15) + 2(10) + 5(20) \\ &= 174 \text{ mg of pesticide 2 absorbed by herbivore 3.}\end{aligned}$$

If we now have p carnivores (such as a human) who eat the herbivores, we can repeat the analysis to find out how much of each pesticide has been absorbed by each carnivore. ■

It is sometimes useful to be able to find a column in the matrix product AB without having to multiply the two matrices. It is not difficult to show (Exercise 46) that the jth column of the matrix product AB is equal to the matrix product $A\mathrm{col}_j(B)$.

EXAMPLE 11 Let

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ -1 & 5 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} -2 & 3 & 4 \\ 3 & 2 & 1 \end{bmatrix}.$$

Then the second column of AB is

$$A\mathrm{col}_2(B) = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ -1 & 5 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 7 \\ 17 \\ 7 \end{bmatrix}.$$

■

Remark If $\mathbf{u}$ and $\mathbf{v}$ are n-vectors ($n \times 1$ matrices), then it is easy to show by matrix multiplication (Exercise 41) that

$$\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T \mathbf{v}.$$

This observation is applied in Chapter 5.

■ The Matrix–Vector Product Written in Terms of Columns

Let

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

be an $m \times n$ matrix and let

$$\mathbf{c} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}$$

be an n-vector, that is, an $n \times 1$ matrix. Since A is $m \times n$ and $\mathbf{c}$ is $n \times 1$, the matrix product $A\mathbf{c}$ is the $m \times 1$ matrix

$$A\mathbf{c} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} (\mathrm{row}_1(A))^T \cdot \mathbf{c} \\ (\mathrm{row}_2(A))^T \cdot \mathbf{c} \\ \vdots \\ (\mathrm{row}_m(A))^T \cdot \mathbf{c} \end{bmatrix}$$
$$= \begin{bmatrix} a_{11}c_1 + a_{12}c_2 + \cdots + a_{1n}c_n \\ a_{21}c_1 + a_{22}c_2 + \cdots + a_{2n}c_n \\ \vdots \\ a_{m1}c_1 + a_{m2}c_2 + \cdots + a_{mn}c_n \end{bmatrix}. \tag{2}$$

This last expression can be written as

$$c_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + c_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + c_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} \tag{3}$$
$$= c_1\mathrm{col}_1(A) + c_2\mathrm{col}_2(A) + \cdots + c_n\mathrm{col}_n(A).$$

Thus the product $A\mathbf{c}$ of an $m \times n$ matrix A and an $n \times 1$ matrix $\mathbf{c}$ can be written as a linear combination of the columns of A, where the coefficients are the entries in the matrix $\mathbf{c}$.

In our study of linear systems of equations we shall see that these systems can be expressed in terms of a matrix–vector product. This point of view provides us with an important way to think about solutions of linear systems.

EXAMPLE 12 Let

$$A = \begin{bmatrix} 2 & -1 & -3 \\ 4 & 2 & -2 \end{bmatrix} \quad \text{and} \quad \mathbf{c} = \begin{bmatrix} 2 \\ -3 \\ 4 \end{bmatrix}.$$

Then the product $A\mathbf{c}$, written as a linear combination of the columns of A, is

$$A\mathbf{c} = \begin{bmatrix} 2 & -1 & -3 \\ 4 & 2 & -2 \end{bmatrix} \begin{bmatrix} 2 \\ -3 \\ 4 \end{bmatrix} = 2\begin{bmatrix} 2 \\ 4 \end{bmatrix} - 3\begin{bmatrix} -1 \\ 2 \end{bmatrix} + 4\begin{bmatrix} -3 \\ -2 \end{bmatrix} = \begin{bmatrix} -5 \\ -6 \end{bmatrix}.$$

■

If A is an $m \times p$ matrix and B is a $p \times n$ matrix, we can then conclude that the jth column of the product AB can be written as a linear combination of the columns of matrix A, where the coefficients are the entries in the jth column of matrix B:

$$\operatorname{col}_j(AB) = A\operatorname{col}_j(B) = b_{1j}\operatorname{col}_1(A) + b_{2j}\operatorname{col}_2(A) + \cdots + b_{pj}\operatorname{col}_p(A).$$

EXAMPLE 13 If A and B are the matrices defined in Example 11, then

$$AB = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ -1 & 5 \end{bmatrix} \begin{bmatrix} -2 & 3 & 4 \\ 3 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 4 & 7 & 6 \\ 6 & 17 & 16 \\ 17 & 7 & 1 \end{bmatrix}.$$

The columns of AB as linear combinations of the columns of A are given by

$$\operatorname{col}_1(AB) = \begin{bmatrix} 4 \\ 6 \\ 17 \end{bmatrix} = A\operatorname{col}_1(B) = -2\begin{bmatrix} 1 \\ 3 \\ -1 \end{bmatrix} + 3\begin{bmatrix} 2 \\ 4 \\ 5 \end{bmatrix}$$

$$\operatorname{col}_2(AB) = \begin{bmatrix} 7 \\ 17 \\ 7 \end{bmatrix} = A\operatorname{col}_2(B) = 3\begin{bmatrix} 1 \\ 3 \\ -1 \end{bmatrix} + 2\begin{bmatrix} 2 \\ 4 \\ 5 \end{bmatrix}$$

$$\operatorname{col}_3(AB) = \begin{bmatrix} 6 \\ 16 \\ 1 \end{bmatrix} = A\operatorname{col}_3(B) = 4\begin{bmatrix} 1 \\ 3 \\ -1 \end{bmatrix} + 1\begin{bmatrix} 2 \\ 4 \\ 5 \end{bmatrix}.$$

■

■ Linear Systems

Consider the linear system of m equations in n unknowns,

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \vdots \qquad \vdots \qquad\qquad \vdots \quad &\quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m. \end{aligned} \tag{4}$$

Now define the following matrices:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \qquad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \qquad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}.$$

Then

$$
\begin{aligned}
A\mathbf{x} &= \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \\
&= \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \qquad\qquad \vdots \qquad\qquad\qquad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{bmatrix}.
\end{aligned} \tag{5}
$$

The entries in the product $A\mathbf{x}$ at the end of (5) are merely the left sides of the equations in (4). Hence the linear system (4) can be written in matrix form as

$$A\mathbf{x} = \mathbf{b}.$$

The matrix A is called the **coefficient matrix** of the linear system (4), and the matrix

$$\left[\begin{array}{cccc:c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right],$$

obtained by adjoining column $\mathbf{b}$ to A, is called the **augmented matrix** of the linear system (4). The augmented matrix of (4) is written as $\left[\begin{array}{c:c} A & \mathbf{b} \end{array}\right]$. Conversely, any matrix with more than one column can be thought of as the augmented matrix of a linear system. The coefficient and augmented matrices play key roles in our method for solving linear systems.

Recall from Section 1.1 that if

$$b_1 = b_2 = \cdots = b_m = 0$$

in (4), the linear system is called a **homogeneous system**. A homogeneous system can be written as

$$A\mathbf{x} = \mathbf{0},$$

where A is the coefficient matrix.

EXAMPLE 14

Consider the linear system

$$
\begin{aligned}
-2x \phantom{{}+3y} + z &= 5 \\
2x + 3y - 4z &= 7 \\
3x + 2y + 2z &= 3.
\end{aligned}
$$

Letting

$$A = \begin{bmatrix} -2 & 0 & 1 \\ 2 & 3 & -4 \\ 3 & 2 & 2 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}, \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} 5 \\ 7 \\ 3 \end{bmatrix},$$

we can write the given linear system in matrix form as

$$A\mathbf{x} = \mathbf{b}.$$

The coefficient matrix is A, and the augmented matrix is

$$\left[\begin{array}{rrr|r} -2 & 0 & 1 & 5 \\ 2 & 3 & -4 & 7 \\ 3 & 2 & 2 & 3 \end{array}\right].$$

■

EXAMPLE 15 The matrix

$$\left[\begin{array}{rrr|r} 2 & -1 & 3 & 4 \\ 3 & 0 & 2 & 5 \end{array}\right]$$

is the augmented matrix of the linear system

$$\begin{aligned} 2x - y + 3z &= 4 \\ 3x + 2z &= 5. \end{aligned}$$

■

We can express (5) in another form, as follows, using (2) and (3):

$$\begin{aligned} A\mathbf{x} &= \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{bmatrix} \\ &= \begin{bmatrix} a_{11}x_1 \\ a_{21}x_1 \\ \vdots \\ a_{m1}x_1 \end{bmatrix} + \begin{bmatrix} a_{12}x_2 \\ a_{22}x_2 \\ \vdots \\ a_{m2}x_2 \end{bmatrix} + \cdots + \begin{bmatrix} a_{1n}x_n \\ a_{2n}x_n \\ \vdots \\ a_{mn}x_n \end{bmatrix} \\ &= x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} \\ &= x_1 \,\mathbf{col}_1(A) + x_2 \,\mathbf{col}_2(A) + \cdots + x_n \,\mathbf{col}_n(A). \end{aligned}$$

Thus $A\mathbf{x}$ is a linear combination of the columns of A with coefficients that are the entries of $\mathbf{x}$. It follows that the matrix form of a linear system, $A\mathbf{x} = \mathbf{b}$, can be expressed as

$$x_1 \,\mathbf{col}_1(A) + x_2 \,\mathbf{col}_2(A) + \cdots + x_n \,\mathbf{col}_n(A) = \mathbf{b}. \tag{6}$$

Conversely, an equation of the form in (6) always describes a linear system of the form in (4).

EXAMPLE 16 Consider the linear system $A\mathbf{x} = \mathbf{b}$, where the coefficient matrix

$$A = \begin{bmatrix} 3 & 1 & 2 \\ 4 & -5 & 6 \\ 0 & 7 & -3 \\ -1 & 2 & 0 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} 4 \\ 1 \\ 0 \\ 2 \end{bmatrix}.$$

Writing $A\mathbf{x} = \mathbf{b}$ as a linear combination of the columns of A as in (6), we have

$$x_1 \begin{bmatrix} 3 \\ 4 \\ 0 \\ -1 \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ -5 \\ 7 \\ 2 \end{bmatrix} + x_3 \begin{bmatrix} 2 \\ 6 \\ -3 \\ 0 \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \\ 0 \\ 2 \end{bmatrix}.$$

■

The expression for the linear system $A\mathbf{x} = \mathbf{b}$ as shown in (6), provides an important way to think about solutions of linear systems.

> $A\mathbf{x} = \mathbf{b}$ is consistent if and only if $\mathbf{b}$ can be expressed as a linear combination of the columns of the matrix A.

We encounter this approach in Chapter 2.

Key Terms

Dot product (inner product)
Matrix–vector product
Coefficient matrix
Augmented matrix

1.3 Exercises

In Exercises 1 and 2, compute $\mathbf{a} \cdot \mathbf{b}$.

1. **(a)** $\mathbf{a} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 4 \\ -1 \end{bmatrix}$

(b) $\mathbf{a} = \begin{bmatrix} -3 \\ -2 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$

(c) $\mathbf{a} = \begin{bmatrix} 4 \\ 2 \\ -1 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 1 \\ 3 \\ 6 \end{bmatrix}$

(d) $\mathbf{a} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$

2. **(a)** $\mathbf{a} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$

(b) $\mathbf{a} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

(c) $\mathbf{a} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}$

(d) $\mathbf{a} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$

3. Let $\mathbf{a} = \mathbf{b} = \begin{bmatrix} -3 \\ 2 \\ x \end{bmatrix}$. If $\mathbf{a} \cdot \mathbf{b} = 17$, find x.

4. Determine the value of x so that $\mathbf{v} \cdot \mathbf{w} = 0$, where

$$\mathbf{v} = \begin{bmatrix} 1 \\ -3 \\ 4 \\ x \end{bmatrix} \quad \text{and} \quad \mathbf{w} = \begin{bmatrix} x \\ 2 \\ -1 \\ 1 \end{bmatrix}.$$

5. Determine values of x and y so that $\mathbf{v} \cdot \mathbf{w} = 0$ and $\mathbf{v} \cdot \mathbf{u} = 0$, where $\mathbf{v} = \begin{bmatrix} x \\ 1 \\ y \end{bmatrix}$, $\mathbf{w} = \begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix}$, and $\mathbf{u} = \begin{bmatrix} 1 \\ 8 \\ 2 \end{bmatrix}$.

6. Determine values of x and y so that $\mathbf{v} \cdot \mathbf{w} = 0$ and $\mathbf{v} \cdot \mathbf{u} = 0$, where $\mathbf{v} = \begin{bmatrix} x \\ 1 \\ y \end{bmatrix}$, $\mathbf{w} = \begin{bmatrix} x \\ -2 \\ 0 \end{bmatrix}$, and $\mathbf{u} = \begin{bmatrix} 0 \\ -9 \\ y \end{bmatrix}$.

7. Let $\mathbf{w} = \begin{bmatrix} \sin\theta \\ \cos\theta \end{bmatrix}$. Compute $\mathbf{w} \cdot \mathbf{w}$.

8. Find all values of x so that $\mathbf{u} \cdot \mathbf{u} = 50$, where $\mathbf{u} = \begin{bmatrix} x \\ 3 \\ 4 \end{bmatrix}$.

9. Find all values of x so that $\mathbf{v} \cdot \mathbf{v} = 1$, where $\mathbf{v} = \begin{bmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ x \end{bmatrix}$.

10. Let $A = \begin{bmatrix} 1 & 2 & x \\ 3 & -1 & 2 \end{bmatrix}$ and $B = \begin{bmatrix} y \\ x \\ 1 \end{bmatrix}$.

If $AB = \begin{bmatrix} 6 \\ 8 \end{bmatrix}$, find x and y.

Consider the following matrices for Exercises 11 through 15:

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 4 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 \\ 2 & 1 \\ 3 & 2 \end{bmatrix},$$

$$C = \begin{bmatrix} 3 & -1 & 3 \\ 4 & 1 & 5 \\ 2 & 1 & 3 \end{bmatrix}, \quad D = \begin{bmatrix} 3 & -2 \\ 2 & 5 \end{bmatrix},$$

$$E = \begin{bmatrix} 2 & -4 & 5 \\ 0 & 1 & 4 \\ 3 & 2 & 1 \end{bmatrix}, \quad \textit{and} \quad F = \begin{bmatrix} -1 & 2 \\ 0 & 4 \\ 3 & 5 \end{bmatrix}.$$

11. If possible, compute the following:

(a) AB **(b)** BA **(c)** $F^T E$

(d) $CB + D$ **(e)** $AB + D^2$, where $D^2 = DD$

12. If possible, compute the following:

(a) $DA + B$ **(b)** EC **(c)** CE

(d) $EB + F$ **(e)** $FC + D$

13. If possible, compute the following:

(a) $FD - 3B$ **(b)** $AB - 2D$

(c) $F^T B + D$ **(d)** $2F - 3(AE)$

(e) $BD + AE$

14. If possible, compute the following:

(a) $A(BD)$ **(b)** $(AB)D$

(c) $A(C + E)$ **(d)** $AC + AE$

(e) $(2AB)^T$ and $2(AB)^T$ **(f)** $A(C - 3E)$

15. If possible, compute the following:

(a) A^T **(b)** $(A^T)^T$

(c) $(AB)^T$ **(d)** $B^T A^T$

(e) $(C + E)^T B$ and $C^T B + E^T B$

(f) $A(2B)$ and $2(AB)$

16. Let $A = \begin{bmatrix} 1 & 2 & -3 \end{bmatrix}$, $B = \begin{bmatrix} -1 & 4 & 2 \end{bmatrix}$, and $C = \begin{bmatrix} -3 & 0 & 1 \end{bmatrix}$. If possible, compute the following:

(a) AB^T **(b)** CA^T **(c)** $(BA^T)C$

(d) $A^T B$ **(e)** CC^T **(f)** $C^T C$

(g) $B^T C A A^T$

17. Let $A = \begin{bmatrix} 2 & 3 \\ -1 & 4 \\ 0 & 3 \end{bmatrix}$ and $B = \begin{bmatrix} 3 & -1 & 3 \\ 1 & 2 & 4 \end{bmatrix}$.

Compute the following entries of AB:

(a) the (1, 2) entry **(b)** the (2, 3) entry

(c) the (3, 1) entry **(d)** the (3, 3) entry

18. If $I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $D = \begin{bmatrix} 2 & 3 \\ -1 & -2 \end{bmatrix}$, compute DI_2 and $I_2 D$.

19. Let

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 2 & -1 \\ -3 & 4 \end{bmatrix}.$$

Show that $AB \neq BA$.

20. If A is the matrix in Example 4 and O is the 3×2 matrix every one of whose entries is zero, compute AO.

In Exercises 21 and 22, let

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

and

$$B = \begin{bmatrix} 1 & 0 & -1 & 2 \\ 3 & 3 & -3 & 4 \\ 4 & 2 & 5 & 1 \end{bmatrix}.$$

21. Using the method in Example 11, compute the following columns of AB:

(a) the first column **(b)** the third column

22. Using the method in Example 11, compute the following columns of AB:

(a) the second column **(b)** the fourth column

23. Let

$$A = \begin{bmatrix} 2 & -3 & 4 \\ -1 & 2 & 3 \\ 5 & -1 & -2 \end{bmatrix} \quad \text{and} \quad \mathbf{c} = \begin{bmatrix} 2 \\ 1 \\ 4 \end{bmatrix}.$$

Express $A\mathbf{c}$ as a linear combination of the columns of A.

24. Let

$$A = \begin{bmatrix} 1 & -2 & -1 \\ 2 & 4 & 3 \\ 3 & 0 & -2 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & -1 \\ 3 & 2 \\ 2 & 4 \end{bmatrix}.$$

Express the columns of AB as linear combinations of the columns of A.

25. Let $A = \begin{bmatrix} 2 & -3 & 1 \\ 1 & 2 & 4 \end{bmatrix}$ and $B = \begin{bmatrix} 3 \\ 5 \\ 2 \end{bmatrix}$.

(a) Verify that $AB = 3\mathbf{a}_1 + 5\mathbf{a}_2 + 2\mathbf{a}_3$, where $\mathbf{a}_j$ is the jth column of A for $j = 1, 2, 3$.

(b) Verify that $AB = \begin{bmatrix} (\text{row}_1(A))B \\ (\text{row}_2(A))B \end{bmatrix}$.

26. **(a)** Find a value of r so that $AB^T = 0$, where $A = \begin{bmatrix} r & 1 & -2 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 3 & -1 \end{bmatrix}$.

(b) Give an alternative way to write this product.

27. Find a value of r and a value of s so that $AB^T = 0$, where $A = \begin{bmatrix} 1 & r & 1 \end{bmatrix}$ and $B = \begin{bmatrix} -2 & 2 & s \end{bmatrix}$.

28. **(a)** Let A be an $m \times n$ matrix with a row consisting entirely of zeros. Show that if B is an $n \times p$ matrix, then AB has a row of zeros.

(b) Let A be an $m \times n$ matrix with a column consisting entirely of zeros and let B be $p \times m$. Show that BA has a column of zeros.

29. Let $A = \begin{bmatrix} -3 & 2 & 1 \\ 4 & 5 & 0 \end{bmatrix}$ with $\mathbf{a}_j =$ the jth column of A, $j = 1, 2, 3$. Verify that

$$A^TA = \begin{bmatrix} \mathbf{a}_1^T\mathbf{a}_1 & \mathbf{a}_1^T\mathbf{a}_2 & \mathbf{a}_1^T\mathbf{a}_3 \\ \mathbf{a}_2^T\mathbf{a}_1 & \mathbf{a}_2^T\mathbf{a}_2 & \mathbf{a}_2^T\mathbf{a}_3 \\ \mathbf{a}_3^T\mathbf{a}_1 & \mathbf{a}_3^T\mathbf{a}_2 & \mathbf{a}_3^T\mathbf{a}_3 \end{bmatrix}.$$

30. Consider the following linear system:

$$\begin{aligned} 2x_1 + 3x_2 - 3x_3 + x_4 + x_5 &= 7 \\ 3x_1 + 2x_3 + 3x_5 &= -2 \\ 2x_1 + 3x_2 - 4x_4 &= 3 \\ x_3 + x_4 + x_5 &= 5. \end{aligned}$$

(a) Find the coefficient matrix.

(b) Write the linear system in matrix form.

(c) Find the augmented matrix.

31. Write the linear system whose augmented matrix is

$$\left[\begin{array}{rrrr|r} -2 & -1 & 0 & 4 & 5 \\ -3 & 2 & 7 & 8 & 3 \\ 1 & 0 & 0 & 2 & 4 \\ 3 & 0 & 1 & 3 & 6 \end{array}\right].$$

32. Write the following linear system in matrix form:

$$\begin{aligned} -2x_1 + 3x_2 &= 5 \\ x_1 - 5x_2 &= 4 \end{aligned}$$

33. Write the following linear system in matrix form:

$$\begin{aligned} 2x_1 + 3x_2 &= 0 \\ 3x_2 + x_3 &= 0 \\ 2x_1 - x_2 &= 0 \end{aligned}$$

34. Write the linear system whose augmented matrix is

(a) $\left[\begin{array}{rrrr|r} 2 & 1 & 3 & 4 & 0 \\ 3 & -1 & 2 & 0 & 3 \\ -2 & 1 & -4 & 3 & 2 \end{array}\right]$.

(b) $\left[\begin{array}{rrrr|r} 2 & 1 & 3 & 4 & 0 \\ 3 & -1 & 2 & 0 & 3 \\ -2 & 1 & -4 & 3 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$.

35. How are the linear systems obtained in Exercise 34 related?

36. Write each of the following linear systems as a linear combination of the columns of the coefficient matrix:

(a) $\begin{aligned} 3x_1 + 2x_2 + x_3 &= 4 \\ x_1 - x_2 + 4x_3 &= -2 \end{aligned}$

(b) $\begin{aligned} -x_1 + x_2 &= 3 \\ 2x_1 - x_2 &= -2 \\ 3x_1 + x_2 &= 1 \end{aligned}$

37. Write each of the following linear combinations of columns as a linear system of the form in (4):

(a) $x_1 \begin{bmatrix} 2 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$

(b) $x_1 \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} + x_3 \begin{bmatrix} 3 \\ 4 \\ 5 \end{bmatrix} + x_4 \begin{bmatrix} 1 \\ 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 2 \\ 5 \\ 8 \end{bmatrix}$

38. Write each of the following as a linear system in matrix form:

(a) $x_1 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ 5 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

(b) $x_1 \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

39. Determine a solution to each of the following linear systems, using the fact that $A\mathbf{x} = \mathbf{b}$ is consistent if and only if $\mathbf{b}$ is a linear combination of the columns of A:

(a) $\begin{bmatrix} 1 & 2 & 1 \\ -3 & 6 & -3 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

(b) $\begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \\ 3 & 4 & 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 20 \\ 20 \\ 20 \end{bmatrix}$

40. Construct a coefficient matrix A so that $\mathbf{x} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$ is a solution to the system $A\mathbf{x} = \mathbf{b}$, where $\mathbf{b} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$. Can there be more than one such coefficient matrix? Explain.

41. Show that if $\mathbf{u}$ and $\mathbf{v}$ are n-vectors, then $\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T \mathbf{v}$.

42. Let A be an $m \times n$ matrix and B an $n \times p$ matrix. What, if anything, can you say about the matrix product AB when

(a) A has a column consisting entirely of zeros?

(b) B has a row consisting entirely of zeros?

43. If $A = \left[a_{ij} \right]$ is an $n \times n$ matrix, then the **trace** of A, $\text{Tr}(A)$, is defined as the sum of all elements on the main diagonal of A, $\text{Tr}(A) = \sum_{i=1}^{n} a_{ii}$. Show each of the following:

(a) $\text{Tr}(cA) = c\,\text{Tr}(A)$, where c is a real number

(b) $\text{Tr}(A + B) = \text{Tr}(A) + \text{Tr}(B)$

(c) $\text{Tr}(AB) = \text{Tr}(BA)$

(d) $\text{Tr}(A^T) = \text{Tr}(A)$

(e) $\text{Tr}(A^T A) \geq 0$

44. Compute the trace (see Exercise 43) of each of the following matrices:

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

(c) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

45. Show that there are no 2×2 matrices A and B such that $AB - BA = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

46. **(a)** Show that the jth column of the matrix product AB is equal to the matrix product $A\mathbf{b}_j$, where $\mathbf{b}_j$ is the jth column of B. It follows that the product AB can be written in terms of columns as

$$AB = \left[A\mathbf{b}_1 \quad A\mathbf{b}_2 \quad \cdots \quad A\mathbf{b}_n \right].$$

(b) Show that the ith row of the matrix product AB is equal to the matrix product $\mathbf{a}_i B$, where $\mathbf{a}_i$ is the ith row of A. It follows that the product AB can be written in terms of rows as

$$AB = \begin{bmatrix} \mathbf{a}_1 B \\ \mathbf{a}_2 B \\ \vdots \\ \mathbf{a}_m B \end{bmatrix}.$$

47. Show that the jth column of the matrix product AB is a linear combination of the columns of A with coefficients the entries in $\mathbf{b}_j$, the jth column of B.

48. The vector

$$\mathbf{u} = \begin{bmatrix} 20 \\ 30 \\ 80 \\ 10 \end{bmatrix}$$

gives the number of receivers, CD players, speakers, and DVD recorders that are on hand in an audio shop. The vector

$$\mathbf{v} = \begin{bmatrix} 200 \\ 120 \\ 80 \\ 70 \end{bmatrix}$$

gives the price (in dollars) of each receiver, CD player, speaker, and DVD recorder, respectively. What does the dot product $\mathbf{u} \cdot \mathbf{v}$ tell the shop owner?

49. **(*Manufacturing Costs*)** A furniture manufacturer makes chairs and tables, each of which must go through an assembly process and a finishing process. The times required for these processes are given (in hours) by the matrix

$$A = \begin{array}{c} \begin{array}{cc} \text{Assembly process} & \text{Finishing process} \end{array} \\ \left[\begin{array}{cc} 2 & 2 \\ 3 & 4 \end{array} \right] \end{array} \begin{array}{l} \\ \text{Chair} \\ \text{Table} \end{array}.$$

The manufacturer has a plant in Salt Lake City and another in Chicago. The hourly rates for each of the processes are given (in dollars) by the matrix

$$B = \begin{array}{c} \begin{array}{cc} \text{Salt Lake City} & \text{Chicago} \end{array} \\ \left[\begin{array}{cc} 9 & 10 \\ 10 & 12 \end{array} \right] \end{array} \begin{array}{l} \\ \text{Assembly process} \\ \text{Finishing process} \end{array}.$$

What do the entries in the matrix product AB tell the manufacturer?

50. **(*Medicine*)** A diet research project includes adults and children of both sexes. The composition of the participants in the project is given by the matrix

$$A = \begin{array}{c} \begin{array}{cc} \text{Adults} & \text{Children} \end{array} \\ \left[\begin{array}{cc} 80 & 120 \\ 100 & 200 \end{array} \right] \end{array} \begin{array}{l} \\ \text{Male} \\ \text{Female} \end{array}.$$

The number of daily grams of protein, fat, and carbohydrate consumed by each child and adult is given by the matrix

$$B = \begin{array}{c} \begin{array}{ccc} \text{Protein} & \text{Fat} & \text{Carbohydrate} \end{array} \\ \left[\begin{array}{ccc} 20 & 20 & 20 \\ 10 & 20 & 30 \end{array} \right] \end{array} \begin{array}{l} \\ \text{Adult} \\ \text{Child} \end{array}.$$

(a) How many grams of protein are consumed daily by the males in the project?

(b) How many grams of fat are consumed daily by the females in the project?

51. Let $\mathbf{x}$ be an n-vector.

(a) Is it possible for $\mathbf{x} \cdot \mathbf{x}$ to be negative? Explain.

(b) If $\mathbf{x} \cdot \mathbf{x} = 0$, what is $\mathbf{x}$?

52. Let $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$ be n-vectors and let k be a real number.

(a) Show that $\mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a}$.

(b) Show that $(\mathbf{a} + \mathbf{b}) \cdot \mathbf{c} = \mathbf{a} \cdot \mathbf{c} + \mathbf{b} \cdot \mathbf{c}$.

(c) Show that $(k\mathbf{a}) \cdot \mathbf{b} = \mathbf{a} \cdot (k\mathbf{b}) = k(\mathbf{a} \cdot \mathbf{b})$.

53. Let A be an $m \times n$ matrix whose entries are real numbers. Show that if $AA^T = O$ (the $m \times m$ matrix all of whose entries are zero), then $A = O$.

54. Use the matrices A and C in Exercise 11 and the matrix multiplication command in your software to compute AC and CA. Discuss the results.

55. Using your software, compute $B^T B$ and BB^T for

$$B = \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \frac{1}{6} \end{bmatrix}.$$

Discuss the nature of the results.

1.4 Algebraic Properties of Matrix Operations

In this section we consider the algebraic properties of the matrix operations just defined. Many of these properties are similar to the familiar properties that hold for real numbers. However, there will be striking differences between the set of real numbers and the set of matrices in their algebraic behavior under certain operations—for example, under multiplication (as seen in Section 1.3). The proofs of most of the properties will be left as exercises.

Theorem 1.1 **Properties of Matrix Addition**

Let A, B, and C be $m \times n$ matrices.

(a) $A + B = B + A$.

(b) $A + (B + C) = (A + B) + C$.

(c) There is a unique $m \times n$ matrix O such that

$$A + O = A \tag{1}$$

for any $m \times n$ matrix A. The matrix O is called the $m \times n$ **zero matrix**.

(d) For each $m \times n$ matrix A, there is a unique $m \times n$ matrix D such that

$$A + D = O. \tag{2}$$

We shall write D as $-A$, so (2) can be written as

$$A + (-A) = O.$$

The matrix $-A$ is called the **negative** of A. We also note that $-A$ is $(-1)A$.

Proof

(a) Let

$$A = \begin{bmatrix} a_{ij} \end{bmatrix}, \quad B = \begin{bmatrix} b_{ij} \end{bmatrix},$$
$$A + B = C = \begin{bmatrix} c_{ij} \end{bmatrix}, \quad \text{and} \quad B + A = D = \begin{bmatrix} d_{ij} \end{bmatrix}.$$

We must show that $c_{ij} = d_{ij}$ for all i, j. Now $c_{ij} = a_{ij} + b_{ij}$ and $d_{ij} = b_{ij} + a_{ij}$ for all i, j. Since a_{ij} and b_{ij} are real numbers, we have $a_{ij} + b_{ij} = b_{ij} + a_{ij}$, which implies that $c_{ij} = d_{ij}$ for all i, j.

(c) Let $U = \begin{bmatrix} u_{ij} \end{bmatrix}$. Then $A + U = A$ if and only if† $a_{ij} + u_{ij} = a_{ij}$, which holds if and only if $u_{ij} = 0$. Thus U is the $m \times n$ matrix all of whose entries are zero: U is denoted by O. ■

The 2×2 zero matrix is

$$O = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

EXAMPLE 1 If

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

then

$$\begin{bmatrix} 4 & -1 \\ 2 & 3 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 4+0 & -1+0 \\ 2+0 & 3+0 \end{bmatrix} = \begin{bmatrix} 4 & -1 \\ 2 & 3 \end{bmatrix}.$$ ■

The 2×3 zero matrix is

$$O = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

EXAMPLE 2 If $A = \begin{bmatrix} 1 & 3 & -2 \\ -2 & 4 & 3 \end{bmatrix}$, then $-A = \begin{bmatrix} -1 & -3 & 2 \\ 2 & -4 & -3 \end{bmatrix}$. ■

Theorem 1.2 **Properties of Matrix Multiplication**

(a) If A, B, and C are matrices of the appropriate sizes, then

$$A(BC) = (AB)C.$$

(b) If A, B, and C are matrices of the appropriate sizes, then

$$(A + B)C = AC + BC.$$

(c) If A, B, and C are matrices of the appropriate sizes, then

$$C(A + B) = CA + CB. \tag{3}$$

Proof

(a) Suppose that A is $m \times n$, B is $n \times p$, and C is $p \times q$. We shall prove the result for the special case $m = 2$, $n = 3$, $p = 4$, and $q = 3$. The general proof is completely analogous.

Let $A = \begin{bmatrix} a_{ij} \end{bmatrix}$, $B = \begin{bmatrix} b_{ij} \end{bmatrix}$, $C = \begin{bmatrix} c_{ij} \end{bmatrix}$, $AB = D = \begin{bmatrix} d_{ij} \end{bmatrix}$, $BC = E = \begin{bmatrix} e_{ij} \end{bmatrix}$, $(AB)C = F = \begin{bmatrix} f_{ij} \end{bmatrix}$, and $A(BC) = G = \begin{bmatrix} g_{ij} \end{bmatrix}$. We must show that $f_{ij} = g_{ij}$ for all i, j. Now

$$f_{ij} = \sum_{k=1}^{4} d_{ik} c_{kj} = \sum_{k=1}^{4} \left(\sum_{r=1}^{3} a_{ir} b_{rk} \right) c_{kj}$$

†The connector "if and only if" means that both statements are true or both statements are false. Thus (i) if $A + U = A$, then $a_{ij} + u_{ij} = a_{ij}$; and (ii) if $a_{ij} + u_{ij} = a_{ij}$, then $A + U = A$. See Appendix C, "Introduction to Proofs."

and

$$g_{ij} = \sum_{r=1}^{3} a_{ir} e_{rj} = \sum_{r=1}^{3} a_{ir} \left(\sum_{k=1}^{4} b_{rk} c_{kj} \right).$$

Then, by the properties satisfied by the summation notation,

$$\begin{aligned} f_{ij} &= \sum_{k=1}^{4} (a_{i1} b_{1k} + a_{i2} b_{2k} + a_{i3} b_{3k}) c_{kj} \\ &= a_{i1} \sum_{k=1}^{4} b_{1k} c_{kj} + a_{i2} \sum_{k=1}^{4} b_{2k} c_{kj} + a_{i3} \sum_{k=1}^{4} b_{3k} c_{kj} \\ &= \sum_{r=1}^{3} a_{ir} \left(\sum_{k=1}^{4} b_{rk} c_{kj} \right) = g_{ij}. \end{aligned}$$

The proofs of (b) and (c) are left as Exercise 4. ■

EXAMPLE 3 Let

$$A = \begin{bmatrix} 5 & 2 & 3 \\ 2 & -3 & 4 \end{bmatrix}, \qquad B = \begin{bmatrix} 2 & -1 & 1 & 0 \\ 0 & 2 & 2 & 2 \\ 3 & 0 & -1 & 3 \end{bmatrix},$$

and

$$C = \begin{bmatrix} 1 & 0 & 2 \\ 2 & -3 & 0 \\ 0 & 0 & 3 \\ 2 & 1 & 0 \end{bmatrix}.$$

Then

$$A(BC) = \begin{bmatrix} 5 & 2 & 3 \\ 2 & -3 & 4 \end{bmatrix} \begin{bmatrix} 0 & 3 & 7 \\ 8 & -4 & 6 \\ 9 & 3 & 3 \end{bmatrix} = \begin{bmatrix} 43 & 16 & 56 \\ 12 & 30 & 8 \end{bmatrix}$$

and

$$(AB)C = \begin{bmatrix} 19 & -1 & 6 & 13 \\ 16 & -8 & -8 & 6 \end{bmatrix} \begin{bmatrix} 1 & 0 & 2 \\ 2 & -3 & 0 \\ 0 & 0 & 3 \\ 2 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 43 & 16 & 56 \\ 12 & 30 & 8 \end{bmatrix}.$$ ■

EXAMPLE 4 Let

$$A = \begin{bmatrix} 2 & 2 & 3 \\ 3 & -1 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 0 & 1 \\ 2 & 3 & -1 \end{bmatrix}, \quad \text{and} \quad C = \begin{bmatrix} 1 & 0 \\ 2 & 2 \\ 3 & -1 \end{bmatrix}.$$

Then

$$(A + B)C = \begin{bmatrix} 2 & 2 & 4 \\ 5 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 2 & 2 \\ 3 & -1 \end{bmatrix} = \begin{bmatrix} 18 & 0 \\ 12 & 3 \end{bmatrix}$$

and (verify)

$$AC + BC = \begin{bmatrix} 15 & 1 \\ 7 & -4 \end{bmatrix} + \begin{bmatrix} 3 & -1 \\ 5 & 7 \end{bmatrix} = \begin{bmatrix} 18 & 0 \\ 12 & 3 \end{bmatrix}.$$

■

Recall Example 9 in Section 1.3, which shows that AB need not always equal BA. This is the first significant difference between multiplication of matrices and multiplication of real numbers.

Theorem 1.3 **Properties of Scalar Multiplication**

If r and s are real numbers and A and B are matrices of the appropriate sizes, then

(a) $r(sA) = (rs)A$

(b) $(r + s)A = rA + sA$

(c) $r(A + B) = rA + rB$

(d) $A(rB) = r(AB) = (rA)B$

Proof

Exercises 13, 14, 16, and 18. ■

EXAMPLE 5 Let

$$A = \begin{bmatrix} 4 & 2 & 3 \\ 2 & -3 & 4 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 3 & -2 & 1 \\ 2 & 0 & -1 \\ 0 & 1 & 2 \end{bmatrix}.$$

Then

$$2(3A) = 2\begin{bmatrix} 12 & 6 & 9 \\ 6 & -9 & 12 \end{bmatrix} = \begin{bmatrix} 24 & 12 & 18 \\ 12 & -18 & 24 \end{bmatrix} = 6A.$$

We also have

$$A(2B) = \begin{bmatrix} 4 & 2 & 3 \\ 2 & -3 & 4 \end{bmatrix}\begin{bmatrix} 6 & -4 & 2 \\ 4 & 0 & -2 \\ 0 & 2 & 4 \end{bmatrix} = \begin{bmatrix} 32 & -10 & 16 \\ 0 & 0 & 26 \end{bmatrix} = 2(AB).$$

■

EXAMPLE 6 Scalar multiplication can be used to change the size of entries in a matrix to meet prescribed properties. Let

$$A = \begin{bmatrix} 3 \\ 7 \\ 2 \\ 1 \end{bmatrix}.$$

Then for $k = \frac{1}{7}$, the largest entry of kA is 1. Also if the entries of A represent the volume of products in gallons, for $k = 4$, kA gives the volume in quarts. ■

So far we have seen that multiplication and addition of matrices have much in common with multiplication and addition of real numbers. We now look at some properties of the transpose.

Theorem 1.4 **Properties of Transpose**

If r is a scalar and A and B are matrices of the appropriate sizes, then

(a) $(A^T)^T = A$

(b) $(A + B)^T = A^T + B^T$

(c) $(AB)^T = B^T A^T$

(d) $(rA)^T = rA^T$

Proof

We leave the proofs of (a), (b), and (d) as Exercises 26 and 27.

(c) Let $A = \begin{bmatrix} a_{ij} \end{bmatrix}$ and $B = \begin{bmatrix} b_{ij} \end{bmatrix}$; let $AB = C = \begin{bmatrix} c_{ij} \end{bmatrix}$. We must prove that c_{ij}^T is the (i, j) entry in $B^T A^T$. Now

$$c_{ij}^T = c_{ji} = \sum_{k=1}^{n} a_{jk} b_{ki} = \sum_{k=1}^{n} a_{kj}^T b_{ik}^T$$

$$= \sum_{k=1}^{n} b_{ik}^T a_{kj}^T = \text{the } (i, j) \text{ entry in } B^T A^T. \quad \blacksquare$$

EXAMPLE 7 Let

$$A = \begin{bmatrix} 1 & 2 & 3 \\ -2 & 0 & 1 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 3 & -1 & 2 \\ 3 & 2 & -1 \end{bmatrix}.$$

Then

$$A^T = \begin{bmatrix} 1 & -2 \\ 2 & 0 \\ 3 & 1 \end{bmatrix} \quad \text{and} \quad B^T = \begin{bmatrix} 3 & 3 \\ -1 & 2 \\ 2 & -1 \end{bmatrix}.$$

Also,

$$A + B = \begin{bmatrix} 4 & 1 & 5 \\ 1 & 2 & 0 \end{bmatrix} \quad \text{and} \quad (A + B)^T = \begin{bmatrix} 4 & 1 \\ 1 & 2 \\ 5 & 0 \end{bmatrix}.$$

Now

$$A^T + B^T = \begin{bmatrix} 4 & 1 \\ 1 & 2 \\ 5 & 0 \end{bmatrix} = (A + B)^T. \quad \blacksquare$$

EXAMPLE 8 Let

$$A = \begin{bmatrix} 1 & 3 & 2 \\ 2 & -1 & 3 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 0 & 1 \\ 2 & 2 \\ 3 & -1 \end{bmatrix}.$$

Then

$$AB = \begin{bmatrix} 12 & 5 \\ 7 & -3 \end{bmatrix} \quad \text{and} \quad (AB)^T = \begin{bmatrix} 12 & 7 \\ 5 & -3 \end{bmatrix}.$$

On the other hand,

$$A^T = \begin{bmatrix} 1 & 2 \\ 3 & -1 \\ 2 & 3 \end{bmatrix} \quad \text{and} \quad B^T = \begin{bmatrix} 0 & 2 & 3 \\ 1 & 2 & -1 \end{bmatrix}.$$

Then

$$B^T A^T = \begin{bmatrix} 12 & 7 \\ 5 & -3 \end{bmatrix} = (AB)^T. \quad \blacksquare$$

We also note two other peculiarities of matrix multiplication. If a and b are real numbers, then $ab = 0$ can hold only if a or b is zero. However, this is not true for matrices.

EXAMPLE 9 If

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 4 & -6 \\ -2 & 3 \end{bmatrix},$$

then neither A nor B is the zero matrix, but $AB = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$. ■

If a, b, and c are real numbers for which $ab = ac$ and $a \neq 0$, it follows that $b = c$. That is, we can cancel out the nonzero factor a. However, the cancellation law does not hold for matrices, as the following example shows.

EXAMPLE 10 If

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 1 \\ 3 & 2 \end{bmatrix}, \quad \text{and} \quad C = \begin{bmatrix} -2 & 7 \\ 5 & -1 \end{bmatrix},$$

then

$$AB = AC = \begin{bmatrix} 8 & 5 \\ 16 & 10 \end{bmatrix},$$

but $B \neq C$. ■

We summarize some of the differences between matrix multiplication and the multiplication of real numbers as follows: For matrices A, B, and C of the appropriate sizes,

1. AB need not equal BA.
2. AB may be the zero matrix with $A \neq O$ and $B \neq O$.
3. AB may equal AC with $B \neq C$.

In this section we have developed a number of properties about matrices and their transposes. If a future problem involves these concepts, refer to these properties to help solve the problem. These results can be used to develop many more results.

Key Terms

Properties of matrix addition
Zero matrix
Properties of matrix multiplication
Properties of scalar multiplication
Properties of transpose

1.4 Exercises

1. Prove Theorem 1.1(b).

2. Prove Theorem 1.1(d).

3. Verify Theorem 1.2(a) for the following matrices:

$$A = \begin{bmatrix} 1 & 3 \\ 2 & -1 \end{bmatrix}, \quad B = \begin{bmatrix} -1 & 3 & 2 \\ 1 & -3 & 4 \end{bmatrix},$$
$$\text{and} \quad C = \begin{bmatrix} 1 & 0 \\ 3 & -1 \\ 1 & 2 \end{bmatrix}.$$

4. Prove Theorem 1.2(b) and (c).

5. Verify Theorem 1.2(c) for the following matrices:

$$A = \begin{bmatrix} 2 & -3 & 2 \\ 3 & -1 & -2 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 3 & -2 \end{bmatrix},$$
$$\text{and} \quad C = \begin{bmatrix} 1 & -3 \\ -3 & 4 \end{bmatrix}.$$

6. Let $A = \begin{bmatrix} a_{ij} \end{bmatrix}$ be the $n \times n$ matrix defined by $a_{ii} = k$ and $a_{ij} = 0$ if $i \neq j$. Show that if B is any $n \times n$ matrix, then $AB = kB$.

7. Let A be an $m \times n$ matrix and $C = \begin{bmatrix} c_1 & c_2 & \cdots & c_m \end{bmatrix}$ a $1 \times m$ matrix. Prove that

$$CA = \sum_{j=1}^{m} c_j A_j,$$

where A_j is the jth row of A.

8. Let $A = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$.

   (a) Determine a simple expression for A^2.

   (b) Determine a simple expression for A^3.

   (c) Conjecture the form of a simple expression for A^k, k a positive integer.

   (d) Prove or disprove your conjecture in part (c).

9. Find a pair of unequal 2×2 matrices A and B, other than those given in Example 9, such that $AB = O$.

10. Find two different 2×2 matrices A such that $A^2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

11. Find two unequal 2×2 matrices A and B such that $AB = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

12. Find two different 2×2 matrices A such that $A^2 = O$.

13. Prove Theorem 1.3(a).

14. Prove Theorem 1.3(b).

15. Verify Theorem 1.3(b) for $r = 4$, $s = -2$, and $A = \begin{bmatrix} 2 & -3 \\ 4 & 2 \end{bmatrix}$.

16. Prove Theorem 1.3(c).

17. Verify Theorem 1.3(c) for $r = -3$,

$$A = \begin{bmatrix} 4 & 2 \\ 1 & -3 \\ 3 & 2 \end{bmatrix}, \quad \text{and} \quad B = \begin{bmatrix} 0 & 2 \\ 4 & 3 \\ -2 & 1 \end{bmatrix}.$$

18. Prove Theorem 1.3(d).

19. Verify Theorem 1.3(d) for the following matrices:

$$A = \begin{bmatrix} 1 & 3 \\ 2 & -1 \end{bmatrix}, \quad B = \begin{bmatrix} -1 & 3 & 2 \\ 1 & -3 & 4 \end{bmatrix},$$
$$\text{and} \quad r = -3.$$

20. The matrix A contains the weight (in pounds) of objects packed on board a spacecraft on earth. The objects are to be used on the moon where things weigh about $\frac{1}{6}$ as much. Write an expression kA that calculates the weight of the objects on the moon.

21. **(a)** A is a 360×2 matrix. The first column of A is $\cos 0^\circ, \cos 1^\circ, \dots, \cos 359^\circ$; and the second column is $\sin 0^\circ, \sin 1^\circ, \dots, \sin 359^\circ$. The graph of the ordered pairs in A is a circle of radius 1 centered at the origin. Write an expression kA for ordered pairs whose graph is a circle of radius 3 centered at the origin.

    (b) Explain how to prove the claims about the circles in part (a).

22. Determine a scalar r such that $A\mathbf{x} = r\mathbf{x}$, where

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \quad \text{and} \quad \mathbf{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

23. Determine a scalar r such that $A\mathbf{x} = r\mathbf{x}$, where

$$A = \begin{bmatrix} 1 & 2 & -1 \\ 1 & 0 & 1 \\ 4 & -4 & 5 \end{bmatrix} \quad \text{and} \quad \mathbf{x} = \begin{bmatrix} -\frac{1}{2} \\ \frac{1}{4} \\ 1 \end{bmatrix}.$$

24. Prove that if $A\mathbf{x} = r\mathbf{x}$ for $n \times n$ matrix A, $n \times 1$ matrix $\mathbf{x}$, and scalar r, then $A\mathbf{y} = r\mathbf{y}$, where $\mathbf{y} = s\mathbf{x}$ for any scalar s.

25. Determine a scalar s such that $A^2\mathbf{x} = s\mathbf{x}$ when $A\mathbf{x} = r\mathbf{x}$.

26. Prove Theorem 1.4(a).

27. Prove Theorem 1.4(b) and (d).

28. Verify Theorem 1.4(a), (b), and (d) for

$$A = \begin{bmatrix} 1 & 3 & 2 \\ 2 & 1 & -3 \end{bmatrix}, \quad B = \begin{bmatrix} 4 & 2 & -1 \\ -2 & 1 & 5 \end{bmatrix},$$
$$\text{and} \quad r = -4.$$

29. Verify Theorem 1.4(c) for

$$A = \begin{bmatrix} 1 & 3 & 2 \\ 2 & 1 & -3 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 3 & -1 \\ 2 & 4 \\ 1 & 2 \end{bmatrix}.$$

30. Let

$$A = \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix}, \quad B = \begin{bmatrix} 3 \\ -2 \\ -4 \end{bmatrix}, \quad \text{and} \quad C = \begin{bmatrix} -1 \\ 5 \\ 1 \end{bmatrix}.$$

(a) Compute $(AB^T)C$.

(b) Compute B^TC and multiply the result by A on the right. (*Hint*: B^TC is 1×1).

(c) Explain why $(AB^T)C = (B^TC)A$.

31. Determine a constant k such that $(kA)^T(kA) = 1$, where $A = \begin{bmatrix} -2 \\ 1 \\ -1 \end{bmatrix}$. Is there more than one value of k that could be used?

32. Find three 2×2 matrices, A, B, and C such that $AB = AC$ with $B \neq C$ and $A \neq O$.

33. Let A be an $n \times n$ matrix and c a real number. Show that if $cA = O$, then $c = 0$ or $A = O$.

34. Determine all 2×2 matrices A such that $AB = BA$ for any 2×2 matrix B.

35. Show that $(A - B)^T = A^T - B^T$.

36. Let $\mathbf{x}_1$ and $\mathbf{x}_2$ be solutions to the homogeneous linear system $A\mathbf{x} = \mathbf{0}$.

(a) Show that $\mathbf{x}_1 + \mathbf{x}_2$ is a solution.

(b) Show that $\mathbf{x}_1 - \mathbf{x}_2$ is a solution.

(c) For any scalar r, show that $r\mathbf{x}_1$ is a solution.

(d) For any scalars r and s, show that $r\mathbf{x}_1 + s\mathbf{x}_2$ is a solution.

37. Show that if $A\mathbf{x} = \mathbf{b}$ has more than one solution, then it has infinitely many solutions. (*Hint*: If $\mathbf{x}_1$ and $\mathbf{x}_2$ are solutions, consider $\mathbf{x}_3 = r\mathbf{x}_1 + s\mathbf{x}_2$, where $r + s = 1$.)

38. Show that if $\mathbf{x}_1$ and $\mathbf{x}_2$ are solutions to the linear system $A\mathbf{x} = \mathbf{b}$, then $\mathbf{x}_1 - \mathbf{x}_2$ is a solution to the associated homogeneous system $A\mathbf{x} = \mathbf{0}$.

39. Let

$$A = \begin{bmatrix} 6 & -1 & 1 \\ 0 & 13 & -16 \\ 0 & 8 & -11 \end{bmatrix} \quad \text{and} \quad \mathbf{x} = \begin{bmatrix} 10.5 \\ 21.0 \\ 10.5 \end{bmatrix}.$$

(a) Determine a scalar r such that $A\mathbf{x} = r\mathbf{x}$.

(b) Is it true that $A^T\mathbf{x} = r\mathbf{x}$ for the value r determined in part (a)?

40. Repeat Exercise 39 with

$$A = \begin{bmatrix} -3.35 & -3.00 & 3.60 \\ 1.20 & 2.05 & -6.20 \\ -3.60 & -2.40 & 3.85 \end{bmatrix}$$
$$\text{and} \quad \mathbf{x} = \begin{bmatrix} 12.5 \\ -12.5 \\ 6.25 \end{bmatrix}.$$

41. Let $A = \begin{bmatrix} 0.1 & 0.01 \\ 0.001 & 0.0001 \end{bmatrix}$. In your software, set the display format to show as many decimal places as possible, then compute

$$B = 10 * A,$$
$$C = \underbrace{A + A + A + A + A + A + A + A + A + A}_{10 \text{ summands}},$$

and

$$D = B - C.$$

If D is not O, then you have verified that scalar multiplication by a positive integer and successive addition are not the same in your computing environment. (It is not unusual that $D \neq O$, since many computing environments use only a “model” of exact arithmetic, called floating-point arithmetic.)

42. Let $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$. In your software, set the display to show as many decimal places as possible. Experiment to find a positive integer k such that $A + 10^{-k} * A$ is equal to A. If you find such an integer k, you have verified that there is more than one matrix in your computational environment that plays the role of O.

1.5 Special Types of Matrices and Partitioned Matrices

We have already introduced one special type of matrix O, the matrix all of whose entries are zero. We now consider several other types of matrices whose structures are rather specialized and for which it will be convenient to have special names.

An $n \times n$ matrix $A = \begin{bmatrix} a_{ij} \end{bmatrix}$ is called a **diagonal matrix** if $a_{ij} = 0$ for $i \neq j$. Thus, for a diagonal matrix, the terms *off* the main diagonal are all zero. Note that O is a diagonal matrix. A **scalar matrix** is a diagonal matrix whose diagonal elements are equal. The scalar matrix $I_n = \begin{bmatrix} d_{ij} \end{bmatrix}$, where $d_{ii} = 1$ and $d_{ij} = 0$ for $i \neq j$, is called the $n \times n$ **identity matrix**.

EXAMPLE 1 Let

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}, \quad \text{and} \quad I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Then A, B, and I_3 are diagonal matrices; B and I_3 are scalar matrices; and I_3 is the 3×3 identity matrix. ■

It is easy to show (Exercise 1) that if A is any $m \times n$ matrix, then

$$AI_n = A \quad \text{and} \quad I_m A = A.$$

Also, if A is a scalar matrix, then $A = rI_n$ for some scalar r.

Suppose that A is a square matrix. We now define the powers of a matrix, for p a positive integer, by

$$A^p = \underbrace{A \cdot A \cdot \cdots \cdot A}_{p \text{ factors}}.$$

If A is $n \times n$, we also define

$$A^0 = I_n.$$

For nonnegative integers p and q, the familiar laws of exponents for the real numbers can also be proved for matrix multiplication of a square matrix A (Exercise 8):

$$A^p A^q = A^{p+q} \quad \text{and} \quad (A^p)^q = A^{pq}.$$

It should also be noted that the rule

$$(AB)^p = A^p B^p$$

does not hold for square matrices unless $AB = BA$ (Exercise 9).

An $n \times n$ matrix $A = \left[a_{ij} \right]$ is called **upper triangular** if $a_{ij} = 0$ for $i > j$. It is called **lower triangular** if $a_{ij} = 0$ for $i < j$. A diagonal matrix is both upper triangular and lower triangular.

EXAMPLE 2

The matrix

$$A = \begin{bmatrix} 1 & 3 & 3 \\ 0 & 3 & 5 \\ 0 & 0 & 2 \end{bmatrix}$$

is upper triangular, and

$$B = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 3 & 0 \\ 3 & 5 & 2 \end{bmatrix}$$

is lower triangular. ■

DEFINITION 1.8

A matrix A with real entries is called **symmetric** if $A^T = A$.

DEFINITION 1.9

A matrix A with real entries is called **skew symmetric** if $A^T = -A$.

EXAMPLE 3

$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{bmatrix}$ is a symmetric matrix. ■

EXAMPLE 4

$B = \begin{bmatrix} 0 & 2 & 3 \\ -2 & 0 & -4 \\ -3 & 4 & 0 \end{bmatrix}$ is a skew symmetric matrix. ■

We can make a few observations about symmetric and skew symmetric matrices; the proofs of most of these statements will be left as exercises.

It follows from the preceding definitions that if A is symmetric or skew symmetric, then A is a square matrix. If A is a symmetric matrix, then the entries of A are symmetric with respect to the main diagonal of A. Also, A is symmetric if and only if $a_{ij} = a_{ji}$, and A is skew symmetric if and only if $a_{ij} = -a_{ji}$. Moreover, if A is skew symmetric, then the entries on the main diagonal of A are all zero. An important property of symmetric and skew symmetric matrices is the following: If A is an $n \times n$ matrix, then we can show that $A = S + K$, where S is symmetric and K is skew symmetric. Moreover, this decomposition is unique (Exercise 29).

■ Partitioned Matrices

If we start out with an $m \times n$ matrix $A = \left[a_{ij} \right]$ and then cross out some, but not all, of its rows or columns, we obtain a **submatrix** of A.

EXAMPLE 5

Let

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ -2 & 4 & -3 & 5 \\ 3 & 0 & 5 & -3 \end{bmatrix}.$$

If we cross out the second row and third column, we get the submatrix

$$\begin{bmatrix} 1 & 2 & 4 \\ 3 & 0 & -3 \end{bmatrix}. \qquad \blacksquare$$

A matrix can be partitioned into submatrices by drawing horizontal lines between rows and vertical lines between columns. Of course, the partitioning can be carried out in many different ways.

EXAMPLE 6 The matrix

$$A = \left[\begin{array}{ccc|cc} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ \hline a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} \end{array}\right] = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

can be partitioned as indicated previously. We could also write

$$A = \left[\begin{array}{cc|cc|c} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ \hline a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} \end{array}\right] = \begin{bmatrix} \widehat{A}_{11} & \widehat{A}_{12} & \widehat{A}_{13} \\ \widehat{A}_{21} & \widehat{A}_{22} & \widehat{A}_{23} \end{bmatrix}, \qquad (1)$$

which gives another partitioning of A. We thus speak of **partitioned matrices**. $\blacksquare$

EXAMPLE 7 The augmented matrix (defined in Section 1.3) of a linear system is a partitioned matrix. Thus, if $A\mathbf{x} = \mathbf{b}$, we can write the augmented matrix of this system as $\left[\begin{array}{c|c} A & \mathbf{b} \end{array}\right]$. $\blacksquare$

If A and B are both $m \times n$ matrices that are partitioned in the same way, then $A + B$ is produced simply by adding the corresponding submatrices of A and B. Similarly, if A is a partitioned matrix, then the scalar multiple cA is obtained by forming the scalar multiple of each submatrix.

If A is partitioned as shown in (1) and

$$B = \left[\begin{array}{cc|cc} b_{11} & b_{12} & b_{13} & b_{14} \\ b_{21} & b_{22} & b_{23} & b_{24} \\ \hline b_{31} & b_{32} & b_{33} & b_{34} \\ b_{41} & b_{42} & b_{43} & b_{44} \\ \hline b_{51} & b_{52} & b_{53} & b_{54} \end{array}\right] = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \\ B_{31} & B_{32} \end{bmatrix},$$

then by straightforward computations we can show that

$$AB = \left[\begin{array}{c|c} (\widehat{A}_{11}B_{11} + \widehat{A}_{12}B_{21} + \widehat{A}_{13}B_{31}) & (\widehat{A}_{11}B_{12} + \widehat{A}_{12}B_{22} + \widehat{A}_{13}B_{32}) \\ \hline (\widehat{A}_{21}B_{11} + \widehat{A}_{22}B_{21} + \widehat{A}_{23}B_{31}) & (\widehat{A}_{21}B_{12} + \widehat{A}_{22}B_{22} + \widehat{A}_{23}B_{32}) \end{array}\right].$$

EXAMPLE 8

Let

$$A = \left[\begin{array}{cc|cc} 1 & 0 & 1 & 0 \\ 0 & 2 & 3 & -1 \\ \hline 2 & 0 & -4 & 0 \\ 0 & 1 & 0 & 3 \end{array}\right] = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

and let

$$B = \left[\begin{array}{ccc|ccc} 2 & 0 & 0 & 1 & 1 & -1 \\ 0 & 1 & 1 & -1 & 2 & 2 \\ \hline 1 & 3 & 0 & 0 & 1 & 0 \\ -3 & -1 & 2 & 1 & 0 & -1 \end{array}\right] = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}.$$

Then

$$AB = C = \left[\begin{array}{ccc|ccc} 3 & 3 & 0 & 1 & 2 & -1 \\ 6 & 12 & 0 & -3 & 7 & 5 \\ \hline 0 & -12 & 0 & 2 & -2 & -2 \\ -9 & -2 & 7 & 2 & 2 & -1 \end{array}\right] = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix},$$

where C_{11} should be $A_{11}B_{11} + A_{12}B_{21}$. We verify that C_{11} is this expression as follows:

$$\begin{aligned} A_{11}B_{11} + A_{12}B_{21} &= \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 3 & -1 \end{bmatrix}\begin{bmatrix} 1 & 3 & 0 \\ -3 & -1 & 2 \end{bmatrix} \\ &= \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 2 \end{bmatrix} + \begin{bmatrix} 1 & 3 & 0 \\ 6 & 10 & -2 \end{bmatrix} \\ &= \begin{bmatrix} 3 & 3 & 0 \\ 6 & 12 & 0 \end{bmatrix} = C_{11}. \end{aligned}$$

■

This method of multiplying partitioned matrices is also known as **block multiplication**. Partitioned matrices can be used to great advantage when matrices exceed the memory capacity of a computer. Thus, in multiplying two partitioned matrices, one can keep the matrices on disk and bring into memory only the submatrices required to form the submatrix products. The products, of course, can be downloaded as they are formed. The partitioning must be done in such a way that the products of corresponding submatrices are defined.

Partitioning of a matrix implies a subdivision of the information into blocks, or units. The reverse process is to consider individual matrices as blocks and adjoin them to form a partitioned matrix. The only requirement is that after the blocks have been joined, all rows have the same number of entries and all columns have the same number of entries.

EXAMPLE 9

Let

$$B = \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & -1 & 0 \end{bmatrix}, \quad \text{and} \quad D = \begin{bmatrix} 9 & 8 & -4 \\ 6 & 7 & 5 \end{bmatrix}.$$

Then we have

$$\begin{bmatrix} B & D \end{bmatrix} = \left[\begin{array}{c:ccc} 2 & 9 & 8 & -4 \\ 3 & 6 & 7 & 5 \end{array}\right], \quad \begin{bmatrix} D \\ C \end{bmatrix} = \left[\begin{array}{ccc} 9 & 8 & -4 \\ 6 & 7 & 5 \\ \hdashline 1 & -1 & 0 \end{array}\right],$$

and

$$\begin{bmatrix} \begin{bmatrix} D \\ C \end{bmatrix} & C^T \end{bmatrix} = \left[\begin{array}{ccc:c} 9 & 8 & -4 & 1 \\ 6 & 7 & 5 & -1 \\ \hdashline 1 & -1 & 0 & 0 \end{array}\right].$$

■

Adjoining matrix blocks to expand information structures is done regularly in a variety of applications. It is common for a business to keep monthly sales data for a year in a 1×12 matrix and then adjoin such matrices to build a sales history matrix for a period of years. Similarly, results of new laboratory experiments are adjoined to existing data to update a database in a research facility.

We have already noted in Example 7 that the augmented matrix of the linear system $A\mathbf{x} = \mathbf{b}$ is a partitioned matrix. At times we shall need to solve several linear systems in which the coefficient matrix A is the same, but the right sides of the systems are different, say, $\mathbf{b}$, $\mathbf{c}$, and $\mathbf{d}$. In these cases we shall find it convenient to consider the partitioned matrix $\left[\begin{array}{c:c:c:c} A & \mathbf{b} & \mathbf{c} & \mathbf{d} \end{array}\right]$. (See Section 4.8.)

■ Nonsingular Matrices

We now come to a special type of square matrix and formulate the notion corresponding to the reciprocal of a nonzero real number.

DEFINITION 1.10 An $n \times n$ matrix A is called **nonsingular**, or **invertible**, if there exists an $n \times n$ matrix B such that $AB = BA = I_n$; such a B is called an **inverse** of A. Otherwise, A is called **singular**, or **noninvertible**.

Remark In Theorem 2.11, Section 2.3, we show that if $AB = I_n$, then $BA = I_n$. Thus, to verify that B is an inverse of A, we need verify only that $AB = I_n$.

EXAMPLE 10 Let $A = \begin{bmatrix} 2 & 3 \\ 2 & 2 \end{bmatrix}$ and $B = \begin{bmatrix} -1 & \frac{3}{2} \\ 1 & -1 \end{bmatrix}$. Since $AB = BA = I_2$, we conclude that B is an inverse of A. ■

Theorem 1.5 The inverse of a matrix, if it exists, is unique.

Proof

Let B and C be inverses of A. Then

$$AB = BA = I_n \quad \text{and} \quad AC = CA = I_n.$$

We then have $B = BI_n = B(AC) = (BA)C = I_nC = C$, which proves that the inverse of a matrix, if it exists, is unique. ■

Because of this uniqueness, we write the inverse of a nonsingular matrix A as A^{-1}. Thus

$$AA^{-1} = A^{-1}A = I_n.$$

EXAMPLE 11 Let

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

If A^{-1} exists, let

$$A^{-1} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

Then we must have

$$AA^{-1} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix} = I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

so that

$$\begin{bmatrix} a + 2c & b + 2d \\ 3a + 4c & 3b + 4d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Equating corresponding entries of these two matrices, we obtain the linear systems

$$\begin{aligned} a + 2c &= 1 \\ 3a + 4c &= 0 \end{aligned} \quad \text{and} \quad \begin{aligned} b + 2d &= 0 \\ 3b + 4d &= 1. \end{aligned}$$

The solutions are (verify) $a = -2$, $c = \frac{3}{2}$, $b = 1$, and $d = -\frac{1}{2}$. Moreover, since the matrix

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{bmatrix}$$

also satisfies the property that

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

we conclude that A is nonsingular and that

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

■

EXAMPLE 12 Let

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}.$$

If A^{-1} exists, let

$$A^{-1} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

Then we must have

$$AA^{-1} = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix} = I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

so that

$$\begin{bmatrix} a+2c & b+2d \\ 2a+4c & 2b+4d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Equating corresponding entries of these two matrices, we obtain the linear systems

$$\begin{aligned} a + 2c &= 1 \\ 2a + 4c &= 0 \end{aligned} \quad \text{and} \quad \begin{aligned} b + 2d &= 0 \\ 2b + 4d &= 1. \end{aligned}$$

These linear systems have no solutions, so our assumption that A^{-1} exists is incorrect. Thus A is singular. ■

We next establish several properties of inverses of matrices.

Theorem 1.6 If A and B are both nonsingular $n \times n$ matrices, then AB is nonsingular and $(AB)^{-1} = B^{-1}A^{-1}$.

Proof

We have $(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = (AI_n)A^{-1} = AA^{-1} = I_n$. Similarly, $(B^{-1}A^{-1})(AB) = I_n$. Therefore AB is nonsingular. Since the inverse of a matrix is unique, we conclude that $(AB)^{-1} = B^{-1}A^{-1}$. ■

Corollary 1.1 If $A_1, A_2, \dots, A_r$ are $n \times n$ nonsingular matrices, then $A_1A_2 \cdots A_r$ is nonsingular and $(A_1A_2 \cdots A_r)^{-1} = A_r^{-1}A_{r-1}^{-1} \cdots A_1^{-1}$.

Proof

Exercise 44. ■

Theorem 1.7 If A is a nonsingular matrix, then A^{-1} is nonsingular and $(A^{-1})^{-1} = A$.

Proof

Exercise 45. ■

Theorem 1.8 If A is a nonsingular matrix, then A^T is nonsingular and $(A^{-1})^T = (A^T)^{-1}$.

Proof

We have $AA^{-1} = I_n$. Taking transposes of both sides, we get

$$(A^{-1})^T A^T = I_n^T = I_n.$$

Taking transposes of both sides of the equation $A^{-1}A = I_n$, we find, similarly, that

$$(A^T)(A^{-1})^T = I_n.$$

These equations imply that $(A^{-1})^T = (A^T)^{-1}$. ■

EXAMPLE 13 If

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

then from Example 11

$$A^{-1} = \begin{bmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{bmatrix} \quad \text{and} \quad (A^{-1})^T = \begin{bmatrix} -2 & \frac{3}{2} \\ 1 & -\frac{1}{2} \end{bmatrix}.$$

Also (verify),

$$A^T = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} \quad \text{and} \quad (A^T)^{-1} = \begin{bmatrix} -2 & \frac{3}{2} \\ 1 & -\frac{1}{2} \end{bmatrix}.$$

■

Suppose that A is nonsingular. Then $AB = AC$ implies that $B = C$ (Exercise 50), and $AB = O$ implies that $B = O$ (Exercise 51).

It follows from Theorem 1.8 that if A is a symmetric nonsingular matrix, then A^{-1} is symmetric. (See Exercise 54.)

■ Linear Systems and Inverses

If A is an $n \times n$ matrix, then the linear system $A\mathbf{x} = \mathbf{b}$ is a system of n equations in n unknowns. Suppose that A is nonsingular. Then A^{-1} exists, and we can multiply $A\mathbf{x} = \mathbf{b}$ by A^{-1} on the left on both sides, yielding

$$\begin{aligned} A^{-1}(A\mathbf{x}) &= A^{-1}\mathbf{b} \\ (A^{-1}A)\mathbf{x} &= A^{-1}\mathbf{b} \\ I_n\mathbf{x} &= A^{-1}\mathbf{b} \\ \mathbf{x} &= A^{-1}\mathbf{b}. \end{aligned} \tag{2}$$

Moreover, $\mathbf{x} = A^{-1}\mathbf{b}$ is clearly a solution to the given linear system. Thus, if A is nonsingular, we have a unique solution. We restate this result for emphasis:

> If A is an $n \times n$ matrix, then the linear system $A\mathbf{x} = \mathbf{b}$ has the unique solution $\mathbf{x} = A^{-1}\mathbf{b}$. Moreover, if $\mathbf{b} = \mathbf{0}$, then the unique solution to the homogeneous system $A\mathbf{x} = \mathbf{0}$ is $\mathbf{x} = \mathbf{0}$.

If A is a nonsingular $n \times n$ matrix, Equation (2) implies that if the linear system $A\mathbf{x} = \mathbf{b}$ needs to be solved repeatedly for different $\mathbf{b}$'s, we need compute A^{-1} only once; then whenever we change $\mathbf{b}$, we find the corresponding solution $\mathbf{x}$ by forming $A^{-1}\mathbf{b}$. Although this is certainly a valid approach, its value is of a more theoretical rather than practical nature, since a more efficient procedure for solving such problems is presented in Section 2.5.

EXAMPLE 14 Suppose that A is the matrix of Example 11 so that

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

If

$$\mathbf{b} = \begin{bmatrix} 8 \\ 6 \end{bmatrix},$$

then the solution to the linear system $A\mathbf{x} = \mathbf{b}$ is

$$\mathbf{x} = A^{-1}\mathbf{b} = \begin{bmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} 8 \\ 6 \end{bmatrix} = \begin{bmatrix} -10 \\ 9 \end{bmatrix}.$$

On the other hand, if

$$\mathbf{b} = \begin{bmatrix} 10 \\ 20 \end{bmatrix},$$

then

$$\mathbf{x} = A^{-1}\begin{bmatrix} 10 \\ 20 \end{bmatrix} = \begin{bmatrix} 0 \\ 5 \end{bmatrix}.$$

■

■ Application A: Recursion Relation; the Fibonacci Sequence

In 1202, Leonardo of Pisa, also called Fibonacci,* wrote a book on mathematics in which he posed the following problem: A pair of newborn rabbits begins to breed at the age of 1 month, and thereafter produces one pair of offspring per month. Suppose that we start with a pair of newly born rabbits and that none of the rabbits produced from this pair dies. How many pairs of rabbits will there be at the beginning of each month?

At the beginning of month 0, we have the newly born pair of rabbits P_1. At the beginning of month 1 we still have only the original pair of rabbits P_1, which have not yet produced any offspring. At the beginning of month 2 we have the original pair P_1 and its first pair of offspring, P_2. At the beginning of month 3 we have the original pair P_1, its first pair of offspring P_2 born at the beginning of month 2, and its second pair of offspring, P_3. At the beginning of month 4 we have P_1, P_2, and P_3; P_4, the offspring of P_1; and P_5, the offspring of P_2. Let u_n denote the number of pairs of rabbits at the beginning of month n. We see that

$$u_0 = 1, \quad u_1 = 1, \quad u_2 = 2, \quad u_3 = 3, \quad u_4 = 5, \quad u_5 = 8.$$

The sequence expands rapidly, and we get

$$1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, \ldots$$

To obtain a formula for u_n, we proceed as follows. The number of pairs of rabbits that are alive at the beginning of month n is u_{n-1}, the number of pairs who were alive the previous month, plus the number of pairs newly born at the beginning of month n. The latter number is u_{n-2}, since a pair of rabbits produces a pair of offspring, starting with its second month of life. Thus

$$u_n = u_{n-1} + u_{n-2}. \tag{3}$$

*Leonardo Fibonacci of Pisa (about 1170–1250) was born and lived most of his life in Pisa, Italy. When he was about 20, his father was appointed director of Pisan commercial interests in northern Africa, now a part of Algeria. Leonardo accompanied his father to Africa and for several years traveled extensively throughout the Mediterranean area on behalf of his father. During these travels he learned the Hindu–Arabic method of numeration and calculation and decided to promote its use in Italy. This was one purpose of his most famous book, *Liber Abaci*, which appeared in 1202 and contained the rabbit problem stated here.

That is, each number is the sum of its two predecessors. The resulting sequence of numbers, called a **Fibonacci sequence**, occurs in a remarkable variety of applications, such as the distribution of leaves on certain trees, the arrangements of seeds on sunflowers, search techniques in numerical analysis, the generation of random numbers in statistics, and others.

To compute u_n by the **recursion relation** (or difference equation) (3), we have to compute $u_0, u_1, \ldots, u_{n-2}, u_{n-1}$. This can be rather tedious for large n. We now develop a formula that will enable us to calculate u_n directly.

In addition to Equation (3), we write

$$u_{n-1} = u_{n-1},$$

so we now have

$$\begin{aligned} u_n &= u_{n-1} + u_{n-2} \\ u_{n-1} &= u_{n-1}, \end{aligned}$$

which can be written in matrix form as

$$\begin{bmatrix} u_n \\ u_{n-1} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} u_{n-1} \\ u_{n-2} \end{bmatrix}. \tag{4}$$

We now define, in general,

$$\mathbf{w}_k = \begin{bmatrix} u_{k+1} \\ u_k \end{bmatrix} \quad \text{and} \quad A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \qquad (0 \le k \le n-1)$$

so that

$$\begin{aligned} \mathbf{w}_0 &= \begin{bmatrix} u_1 \\ u_0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \\ \mathbf{w}_1 &= \begin{bmatrix} u_2 \\ u_1 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \ldots, \quad \mathbf{w}_{n-2} = \begin{bmatrix} u_{n-1} \\ u_{n-2} \end{bmatrix}, \quad \text{and} \quad \mathbf{w}_{n-1} = \begin{bmatrix} u_n \\ u_{n-1} \end{bmatrix}. \end{aligned}$$

Then (4) can be written as

$$\mathbf{w}_{n-1} = A\mathbf{w}_{n-2}.$$

Thus

$$\begin{aligned} \mathbf{w}_1 &= A\mathbf{w}_0 \\ \mathbf{w}_2 &= A\mathbf{w}_1 = A(A\mathbf{w}_0) = A^2\mathbf{w}_0 \\ \mathbf{w}_3 &= A\mathbf{w}_2 = A(A^2\mathbf{w}_0) = A^3\mathbf{w}_0 \\ &\vdots \\ \mathbf{w}_{n-1} &= A^{n-1}\mathbf{w}_0. \end{aligned}$$

Hence, to find u_n, we merely have to calculate A^{n-1}, which is still rather tedious if n is large. In Chapter 7 we develop a more efficient way to compute the Fibonacci numbers that involves powers of a *diagonal* matrix. (See the discussion exercises in Chapter 7.)

Key Terms

Diagonal matrix
Identity matrix
Powers of a matrix
Upper triangular matrix
Lower triangular matrix
Symmetric matrix
Skew symmetric matrix
Submatrix
Partitioning
Partitioned matrix
Nonsingular (invertible) matrix
Inverse
Singular (noninvertible) matrix
Properties of nonsingular matrices
Linear system with nonsingular coefficient matrix
Fibonacci sequence

1.5 Exercises

1. **(a)** Show that if A is any $m \times n$ matrix, then $I_m A = A$ and $A I_n = A$.

(b) Show that if A is an $n \times n$ scalar matrix, then $A = r I_n$ for some real number r.

2. Prove that the sum, product, and scalar multiple of diagonal, scalar, and upper (lower) triangular matrices is diagonal, scalar, and upper (lower) triangular, respectively.

3. Prove: If A and B are $n \times n$ diagonal matrices, then $AB = BA$.

4. Let

$$A = \begin{bmatrix} 3 & 2 & -1 \\ 0 & -4 & 3 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 6 & -3 & 2 \\ 0 & 2 & 4 \\ 0 & 0 & 3 \end{bmatrix}.$$

Verify that $A + B$ and AB are upper triangular.

5. Describe all matrices that are both upper and lower triangular.

6. Let $A = \begin{bmatrix} 1 & 2 \\ 3 & -2 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & -1 \\ 2 & 3 \end{bmatrix}$. Compute each of the following:

(a) A^2 **(b)** B^3 **(c)** $(AB)^2$

7. Let $A = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 1 & 1 \\ 3 & 1 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & 0 & 1 \\ -1 & 1 & 1 \\ 2 & 0 & 1 \end{bmatrix}$.
Compute each of the following:

(a) A^3 **(b)** B^2 **(c)** $(AB)^3$

8. Let p and q be nonnegative integers and let A be a square matrix. Show that

$$A^p A^q = A^{p+q} \quad \text{and} \quad (A^p)^q = A^{pq}.$$

9. If $AB = BA$ and p is a nonnegative integer, show that $(AB)^p = A^p B^p$.

10. If p is a nonnegative integer and c is a scalar, show that $(cA)^p = c^p A^p$.

11. For a square matrix A and a nonnegative integer p, show that $(A^T)^p = (A^p)^T$.

12. For a nonsingular matrix A and a nonnegative integer p, show that $(A^p)^{-1} = (A^{-1})^p$.

13. For a nonsingular matrix A and nonzero scalar k, show that $(kA)^{-1} = \frac{1}{k}A^{-1}$.

14. **(a)** Show that every scalar matrix is symmetric.

(b) Is every scalar matrix nonsingular? Explain.

(c) Is every diagonal matrix a scalar matrix? Explain.

15. Find a 2×2 matrix $B \neq O$ and $B \neq I_2$ such that $AB = BA$, where $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$. How many such matrices B are there?

16. Find a 2×2 matrix $B \neq O$ and $B \neq I_2$ such that $AB = BA$, where $A = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$. How many such matrices B are there?

17. Prove or disprove: For any $n \times n$ matrix A, $A^T A = A A^T$.

18. **(a)** Show that A is symmetric if and only if $a_{ij} = a_{ji}$ for all i, j.

(b) Show that A is skew symmetric if and only if $a_{ij} = -a_{ji}$ for all i, j.

(c) Show that if A is skew symmetric, then the elements on the main diagonal of A are all zero.

19. Show that if A is a symmetric matrix, then A^T is symmetric.

20. Describe all skew symmetric scalar matrices.

21. Show that if A is any $m \times n$ matrix, then AA^T and $A^T A$ are symmetric.

22. Show that if A is any $n \times n$ matrix, then

(a) $A + A^T$ is symmetric.

(b) $A - A^T$ is skew symmetric.

23. Show that if A is a symmetric matrix, then A^k, $k = 2, 3, \ldots$, is symmetric.

24. Let A and B be symmetric matrices.

(a) Show that $A + B$ is symmetric.

(b) Show that AB is symmetric if and only if $AB = BA$.

25. **(a)** Show that if A is an upper triangular matrix, then A^T is lower triangular.

(b) Show that if A is a lower triangular matrix, then A^T is upper triangular.

26. If A is a skew symmetric matrix, what type of matrix is A^T? Justify your answer.

27. Show that if A is skew symmetric, then the elements on the main diagonal of A are all zero.

28. Show that if A is skew symmetric, then A^k is skew symmetric for any positive odd integer k.

29. Show that if A is an $n \times n$ matrix, then $A = S + K$, where S is symmetric and K is skew symmetric. Also show that this decomposition is unique. (*Hint*: Use Exercise 22.)

30. Let

$$A = \begin{bmatrix} 1 & 3 & -2 \\ 4 & 6 & 2 \\ 5 & 1 & 3 \end{bmatrix}.$$

Find the matrices S and K described in Exercise 29.

31. Show that the matrix $A = \begin{bmatrix} 2 & 3 \\ 4 & 6 \end{bmatrix}$ is singular.

32. If $D = \begin{bmatrix} 4 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$, find D^{-1}.

33. Find the inverse of each of the following matrices:

(a) $A = \begin{bmatrix} 1 & 3 \\ 5 & 2 \end{bmatrix}$ **(b)** $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$

34. If A is a nonsingular matrix whose inverse is $\begin{bmatrix} 2 & 1 \\ 4 & 1 \end{bmatrix}$, find A.

35. If

$$A^{-1} = \begin{bmatrix} 3 & 2 \\ 1 & 3 \end{bmatrix} \quad \text{and} \quad B^{-1} = \begin{bmatrix} 2 & 5 \\ 3 & -2 \end{bmatrix},$$

find $(AB)^{-1}$.

36. Suppose that

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

Solve the linear system $A\mathbf{x} = \mathbf{b}$ for each of the following matrices $\mathbf{b}$:

(a) $\begin{bmatrix} 4 \\ 6 \end{bmatrix}$ **(b)** $\begin{bmatrix} 8 \\ 15 \end{bmatrix}$

37. The linear system $AC\mathbf{x} = \mathbf{b}$ is such that A and C are nonsingular with

$$A^{-1} = \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix}, \quad C^{-1} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}, \quad \text{and } \mathbf{b} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}.$$

Find the solution $\mathbf{x}$.

38. The linear system $A^2\mathbf{x} = \mathbf{b}$ is such that A is nonsingular with

$$A^{-1} = \begin{bmatrix} 3 & 0 \\ 2 & 1 \end{bmatrix} \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}.$$

Find the solution $\mathbf{x}$.

39. The linear system $A^T\mathbf{x} = \mathbf{b}$ is such that A is nonsingular with

$$A^{-1} = \begin{bmatrix} 4 & 1 \\ 1 & 0 \end{bmatrix} \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} 1 \\ -2 \end{bmatrix}.$$

Find the solution $\mathbf{x}$.

40. The linear system $C^T A\mathbf{x} = \mathbf{b}$ is such that A and C are nonsingular, with

$$A^{-1} = \begin{bmatrix} 2 & 1 \\ 3 & 1 \end{bmatrix}, \quad C^{-1} = \begin{bmatrix} 0 & 2 \\ 1 & 4 \end{bmatrix}, \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Find the solution $\mathbf{x}$.

41. Consider the linear system $A\mathbf{x} = \mathbf{b}$, where A is the matrix defined in Exercise 33(a).

(a) Find a solution if $\mathbf{b} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$.

(b) Find a solution if $\mathbf{b} = \begin{bmatrix} 5 \\ 6 \end{bmatrix}$.

42. Find two 2×2 singular matrices whose sum is nonsingular.

43. Find two 2×2 nonsingular matrices whose sum is singular.

44. Prove Corollary 1.1.

45. Prove Theorem 1.7.

46. Prove that if one row (column) of the $n \times n$ matrix A consists entirely of zeros, then A is singular. (*Hint*: Assume that A is nonsingular; that is, there exists an $n \times n$ matrix B such that $AB = BA = I_n$. Establish a contradiction.)

47. Prove: If A is a diagonal matrix with nonzero diagonal entries $a_{11}, a_{22}, \ldots, a_{nn}$, then A is nonsingular and A^{-1} is a diagonal matrix with diagonal entries $1/a_{11}, 1/a_{22}, \ldots, 1/a_{nn}$.

48. Let $A = \begin{bmatrix} 2 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 5 \end{bmatrix}$. Compute A^4.

49. For an $n \times n$ diagonal matrix A whose diagonal entries are $a_{11}, a_{22}, \ldots, a_{nn}$, compute A^p for a nonnegative integer p.

50. Show that if $AB = AC$ and A is nonsingular, then $B = C$.