Chapter 1

Matrices

You should be familiar with

- Two- and three-dimensional geometry
- Elementary functions

Linear algebra is a branch of mathematics that is used by engineers and applied scientists to design and analyze complex systems. Civil engineers use linear algebra to design and analyze load-bearing structures such as bridges. Mechanical engineers use linear algebra to design and analyze suspension systems, and electrical engineers use it to design and analyze electrical circuits. Electrical, biomedical, and aerospace engineers use linear algebra to enhance X-rays, tomographs, and images from space. This introduction is intended to serve as a basis for the study of numerical linear algebra, the study of procedures used on a computer to perform linear algebra computations, most notably matrix operations. As you will see, there is a big difference between theoretical linear algebra and applying linear algebra on a computer and obtaining reliable results. It is assumed only that the reader has completed one or more calculus courses and has had some exposure to vectors and matrices, although the text provides a review of the basic concepts. It will be helpful but not necessary if the reader has taken a course in discrete mathematics that provided some exposure to mathematical proofs.

Section 1.1 discusses matrix operations, including matrix multiplication and that matrix multiplication obeys many of the familiar laws of arithmetic apart from the commutative law. While matrix multiplication is most often performed on a computer, it is necessary to understand its definition, fundamental properties, and applications. For instance, a linear system of equations is elegantly expressed in matrix form. This section also introduces the matrix trace operator and the very useful fact that trace $(AB) = \text{trace}\,(BA)$ for square matrices A and B. This section concludes with a presentation of basic MATLAB operations for executing these fundamental matrix operations.

A linear transformation is an absolutely critical concept in linear algebra, and Section 1.2 presents the concept and shows how a linear transformation performs a rotation of a figure in the xy-plane or in three-dimensional space. This application of linear transformations is fundamental to computer graphics.

Section 1.3 discusses powers of matrices and shows the connection between matrix powers and the number of possible paths between two vertices of a graph. This section also presents the interesting Fibonacci matrix.

Section 1.4 introduces the matrix inverse and a number of its properties. It is shown that a linear system has a unique solution when its coefficient matrix has an inverse.

Section 1.5 discusses the matrix transpose and this motivates the definition of a symmetric matrix. As we will see in later chapters, symmetric matrices have many applications in engineering and science.

1.1 MATRIX ARITHMETIC

A matrix is a rectangular array of numbers with m rows and n columns. The symbol $\mathbb{R}^{m\times n}$ denotes the collection of all $m\times n$ matrices whose entries are real numbers. Matrices will usually be denoted by capital letters, and the notation $A = [a_{ij}]$ specifies that the matrix is composed of entries a_{ij} located in the ith row and jth column of A.

A vector is a matrix with either one row or one column; for instance,

$$x = \begin{bmatrix} 1 \\ -2 \\ 6 \end{bmatrix}$$

is a column vector, and

$$y = \begin{bmatrix} 6 & -1 & 3 \end{bmatrix}$$

is a row vector. The elements of a vector require only one subscript. For the vector x, $x_2 = -4$.

Numerical Linear Algebra with Applications. http://dx.doi.org/10.1016/B978-0-12-394435-1.00001-6

Example 1.1. The formula $a_{ij} = 1/(i+j)$ for $1 \le i \le 3, 1 \le j \le 4$ defines a 3×4 matrix $A = [a_{ij}]$, namely,

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

The first column of A is the column vector $\begin{bmatrix} \frac{1}{2} \\ \frac{1}{3} \\ \frac{1}{4} \end{bmatrix}$. ■

Definition 1.1 (Equality of matrices). Matrices A and B are said to be equal if they have the same size and their corresponding elements are equal; i.e., A and B have dimension $m \times n$, and $A = [a_{ij}], B = [b_{ij}]$, with $a_{ij} = b_{ij}$ for $1 \le i \le m, 1 \le j \le n$.

Definition 1.2 (Addition of matrices). Let $A = [a_{ij}]$ and $B = [b_{ij}]$ be of the same size. Then $A+B$ is the matrix obtained by adding corresponding elements of A and B; that is,

$$A + B = [a_{ij}] + [b_{ij}] = [a_{ij} + b_{ij}].$$

Definition 1.3 (Scalar multiple of a matrix). Let $A = [a_{ij}]$ and t be a number (*scalar*). Then tA is the matrix obtained by multiplying all elements of A by t; that is,

$$tA = t[a_{ij}] = [ta_{ij}].$$

Definition 1.4 (Negative of a matrix). Let $A = [a_{ij}]$. Then $-A$ is the matrix obtained by replacing the elements of A by their negatives; that is,

$$-A = -[a_{ij}] = [-a_{ij}].$$

Definition 1.5 (Subtraction of matrices). Matrix subtraction is defined for two matrices $A = [a_{ij}]$ and $B = [b_{ij}]$ of the same size, in the usual way; that is,

$$A - B = [a_{ij}] - [b_{ij}] = [a_{ij} - b_{ij}].$$

Definition 1.6 (The zero matrix). Each $m \times n$ matrix, all of whose elements are zero, is called the *zero* matrix (of size $m \times n$) and is denoted by the symbol 0.

The matrix operations of addition, scalar multiplication, negation and subtraction satisfy the usual laws of arithmetic. (In what follows, s and t are arbitrary scalars and A, B, C are matrices of the same size.)

1. $(A + B) + C = A + (B + C)$;
2. $A + B = B + A$;
3. $0 + A = A$;
4. $A + (-A) = 0$;
5. $(s + t)A = sA + tA$, $(s - t)A = sA - tA$;
6. $t(A + B) = tA + tB$, $t(A - B) = tA - tB$;
7. $s(tA) = (st)A$;
8. $1A = A$, $0A = 0$, $(-1)A = -A$;
9. $tA = 0 \Rightarrow t = 0$ or $A = 0$.

Other similar properties will be used when needed.

1.1.1 Matrix Product

Definition 1.7 (Matrix product). Let $A = [a_{ij}]$ be a matrix of size $m \times p$ and $B = [b_{jk}]$ be a matrix of size $p \times n$ (i.e., the number of columns of A equals the number of rows of B). Then AB is the $m \times n$ matrix $C = [c_{ik}]$ whose (i,j)th element is defined by the formula

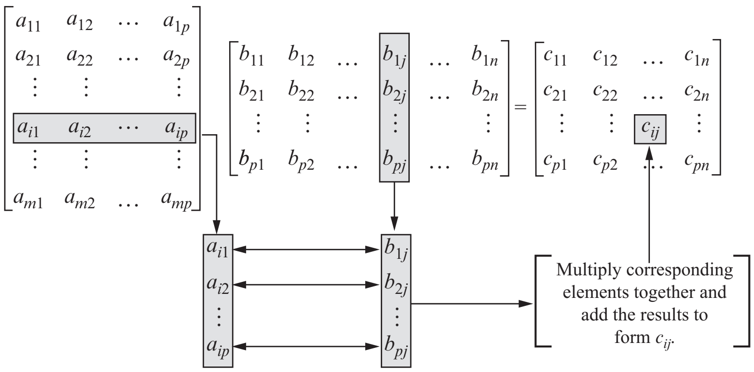

FIGURE 1.1 Matrix multiplication.

$$c_{ij} = \sum_{k=1}^{p} a_{ik}b_{kj} = a_{i1}b_{1j} + \cdots + a_{ip}b_{pj}.$$

A way to look at this is that c_{ij} is the sum of the products of corresponding elements from row i of A and column j of B. For hand computation, fix on row 1 of A. Form the sum of products of corresponding elements from row 1 of A and column 1 of B, then the sum of products of corresponding elements from row 1 of A and column 2 of B, and so forth, until forming the sum of the products of corresponding elements of row 1 of A and column n of B. This computes the first row of the product matrix C. Now use row 2 of A in the same fashion to compute the second row of C. Continue until you have all m rows of C (Figure 1.1).

Example 1.2.

$$\begin{aligned}
\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} &= \begin{bmatrix} 1\times 5+2\times 7 & 1\times 6+2\times 8 \\ 3\times 5+4\times 7 & 3\times 6+4\times 8 \end{bmatrix} = \begin{bmatrix} 19 & 22 \\ 43 & 50 \end{bmatrix}, \\
\begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} &= \begin{bmatrix} 23 & 34 \\ 31 & 46 \end{bmatrix} \neq \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}, \\
\begin{bmatrix} 1 \\ 2 \end{bmatrix}\begin{bmatrix} 3 & 4 \end{bmatrix} &= \begin{bmatrix} 3 & 4 \\ 6 & 8 \end{bmatrix}, \\
\begin{bmatrix} 3 & 4 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \end{bmatrix} &= \begin{bmatrix} 11 \end{bmatrix}, \\
\begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix} &= \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.
\end{aligned}$$

■

Remark 1.1. Matrix multiplication is a computationally expensive operation. On a computer, multiplication is a much more time-consuming operation than addition. Consider computing the product of an $m \times k$ matrix A and a $k \times n$ matrix B. The computation of $(AB)_{ij}$ requires calculating k products. This must be done n times to form each row of AB, so the computation of a row of AB requires kn multiplications. There are m rows in AB, so the total number of multiplications is $m(kn) = mkn$. If A and B are both $n \times n$ matrices, n^3 multiplications must be performed. For example, if the matrices have dimension 10×10, the computation of their product requires 1000 multiplications. To multiply two 100×100 matrices involves computing 1,000,000 products. A matrix most of whose entries are zero is called *sparse*. There are faster ways to multiply *sparse* matrices, and we will deal with these matrices in Chapters 21 and 22.

Theorem 1.1. *Matrix multiplication obeys many of the familiar laws of arithmetic apart from the commutative law.*

1. $(AB)C = A(BC)$ *if* A, B, C *are* $m\times p, p\times k, k\times n$*, respectively;*
2. $t(AB) = (tA)B = A(tB)$, $A(-B) = (-A)B = -(AB)$*;*
3. $(A+B)C = AC + BC$ *if* A *and* B *are* $m\times n$ *and* C *is* $n\times p$*;*
4. $D(A+B) = DA + DB$ *if* A *and* B *are* $m\times n$ *and* D *is* $p\times m$*.*

We prove the associative law only:

Proof. Assume that A is an $m \times p$ matrix, B is a $p \times k$ matrix, and C is a $k \times n$ matrix. Observe that $(AB)C$ and $A(BC)$ are both of size $m \times n$.

Let $A = [a_{iq}], B = [b_{ql}], C = [c_{lj}]$. Then

$$\begin{aligned} ((AB)C)_{ij} &= \sum_{q=1}^{k} (AB)_{iq} c_{qj} = \sum_{q=1}^{k} \left(\sum_{l=1}^{p} a_{il} b_{lq} \right) c_{qj} \\ &= \sum_{q=1}^{k} \sum_{l=1}^{p} a_{il} b_{lq} c_{qj}. \end{aligned}$$

Similarly,

$$(A(BC))_{ij} = \sum_{l=1}^{p} \sum_{q=1}^{k} a_{il} b_{lq} c_{qj}.$$

However, the double summations are equal. Sums of the form

$$\sum_{q=1}^{k} \sum_{l=1}^{p} d_{lq} \quad \text{and} \quad \sum_{l=1}^{p} \sum_{q=1}^{k} d_{lq}$$

represent the sum of the kp elements of the rectangular array $[d_{lq}]$, by rows and by columns, respectively. Consequently, $((AB)C)_{ij} = (A(BC))_{ij}$ for $1 \le i \le m, 1 \le j \le n$. Hence, $(AB)C = A(BC)$. $\square$

One of the primary uses of matrix multiplication is formulating a system of equations as a matrix problem. The system of m linear equations in n unknowns

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

is equivalent to a single-matrix equation

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix},$$

that is, $Ax = b$, where $A = [a_{ij}]$ is the *coefficient matrix* of the system, $x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ is the *vector of unknowns* and $b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$ is the *vector of constants.*

Another useful matrix equation equivalent to the above system of linear equations is

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

We will begin a study of $n \times n$ linear systems in Chapter 2 and continue the study throughout the book. In Chapter 16, most of the systems we deal with will have dimension $m \times n$, where $m \neq n$.

Example 1.3. The system

$$\begin{aligned} x + y + z &= 1, \\ x - y + z &= 0, \\ 3x + 5y - z &= 2 \end{aligned}$$

is equivalent to the matrix equation

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 3 & 5 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}$$

and to the equation

$$x \begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix} + y \begin{bmatrix} 1 \\ -1 \\ 5 \end{bmatrix} + z \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}.$$

The solution to the system is $\begin{bmatrix} 0.0000 \\ 0.5000 \\ 0.5000 \end{bmatrix}$:

$$\begin{aligned} (1)\,0.0000 + (1)\,0.5000 + (1)\,0.5000 &= 1, \\ (1)\,0.0000 - (1)\,0.5000 + (1)\,0.5000 &= 0, \\ 3\,(0.0000) + 5\,(0.5000) - (1)\,0.5000 &= 2. \end{aligned}$$

■

1.1.2 The Trace

The trace is a matrix operation that is frequently used in matrix formulas, and it is very simple to compute.

Definition 1.8. If A is an $n \times n$ matrix, the trace of A, written trace (A), is the sum of the diagonal elements; that is,

$$\operatorname{trace}(A) = a_{11} + a_{22} + \cdots + a_{nn} = \sum_{i=1}^{n} a_{ii}.$$

Example 1.4. If $A = \begin{bmatrix} 5 & 8 & 12 & -1 \\ 7 & 4 & -8 & 7 \\ 0 & 3 & -6 & 5 \\ -1 & -9 & 4 & 3 \end{bmatrix}$, then $\operatorname{trace}(A) = 5 + 4 + (-6) + 3 = 6$. ■

There are a number of relationships satisfied by trace. For instance, $\operatorname{trace}(A + B) = \operatorname{trace}(A) + \operatorname{trace}(B)$ (Problem 1.22(a)), and $\operatorname{trace}(cA) = c\operatorname{trace}(A)$, where c is a scalar (Problem 1.22(b)). A more complex relationship is the trace of product of two matrices.

Theorem 1.2. *If A is an $n \times n$ matrix and B is an $n \times n$ matrix, then* $\operatorname{trace}(AB) = \operatorname{trace}(BA)$.

Proof. By the definition of matrix multiplication,

$$\operatorname{trace}(AB) = \sum_{i=1}^{n} (AB)_{ii} = \sum_{i=1}^{n} \left(\sum_{k=1}^{n} a_{ik} b_{ki} \right) = \sum_{k=1}^{n} \left(\sum_{i=1}^{n} b_{ik} a_{ki} \right) = \operatorname{trace}(BA).$$

□

1.1.3 MATLAB Examples

There are numerous examples throughout this book that involve the use of MATLAB. It is fundamental to our use of MATLAB that you are familiar with "*vectorization*." When performing vector or matrix operations using the operators "*", "/", and "^", it may be necessary to use the dot operator ("."). As stated in the preface, the reader is expected to be familiar with MATLAB, but it is not necessary to be an expert.

Example 1.5. *Matrix Operations*
The operators $+$, $-$, * work as expected in MATLAB, and the command `trace` computes the trace of a matrix.

```
>> A = [1 5 1;2 -1 6;1 0 3]

A =
     1     5     1
     2    -1     6
     1     0     3

>> B = [2 3 0;3 -1 7;4 8 9]

B =
     2     3     0
     3    -1     7
     4     8     9

>> 5*A -10*B + 3*A*B

ans =
    48    13   137
    55   170   101
     7     1     6

>> trace(A + B)

ans =
    13

>> 7*trace(A + B)

ans =
    91

>> trace(A*B)

ans =
   103

>> trace(B*A)

ans =
   103

>> A.*B

ans =
     2    15     0
     6     1    42
     4     0    27
```

■

1.2 LINEAR TRANSFORMATIONS

Throughout this book we will assume that matrices have elements that are real numbers. The real numbers include the integers $(\dots, -2, -1, 0, 1, 2, \dots)$, which are a subset of the *rational numbers* (p/q), where p and q are positive integers, $q \neq 0$. The remaining numbers are called *irrational numbers*; for instance, π and e are irrational numbers. We will use the symbol $\mathbb{R}$ to denote the collection of real numbers. An *n-dimensional column vector* is an $n \times 1$ matrix. The collection of all n-dimensional column vectors is denoted by $\mathbb{R}^n$.

Every matrix is associated with a type of function called a *linear transformation*.

Definition 1.9 (Linear transformation). We can associate an $m \times n$ matrix A with the function $T_A : \mathbb{R}^n \to \mathbb{R}^m$, defined by $T_A(x) = Ax$ for all $x \in \mathbb{R}^n$. More explicitly, using components, the above function takes the form

$$\begin{aligned} y_1 &= a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ y_2 &= a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ &\vdots \\ y_m &= a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n, \end{aligned}$$

where $y_1, y_2, \dots, y_m$ are the components of the column vector $T_A(x)$, in other words $y = Ax$.

A linear transformation has the property that

$$T_A(sx + ty) = sT_A(x) + tT_A(y)$$

for all $s, t \in \mathbb{R}$ and all n-dimensional column vectors x, y. This is true because

$$T_A(sx + ty) = A(sx + ty) = s(Ax) + t(Ay) = sT_A(x) + tT_A(y).$$

1.2.1 Rotations

One well-known example of a linear transformation arises from rotating the (x, y)-plane in two-dimensional Euclidean space counterclockwise about the origin (0, 0) through θ radians. A point (x, y) will be transformed into the point $(\bar{x}, \bar{y})$. By referring to Figure 1.2, the coordinates of the rotated point can be found using a little trigonometry.

$$\begin{aligned} \bar{x} &= d\cos(\theta + \alpha) = d\cos(\theta)\cos(\alpha) - d\sin(\theta)\sin(\alpha) = x\cos(\theta) - y\sin(\theta), \\ \bar{y} &= d\sin(\theta + \alpha) = d\sin(\theta)\cos(\alpha) + d\cos(\theta)\sin(\alpha) = x\sin(\theta) + y\cos(\theta). \end{aligned}$$

The equations in matrix form are

$$R = \begin{bmatrix} \bar{x} \\ \bar{y} \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

Example 1.6. Rotate the line $y = 5x + 1$ an angle of 30° counterclockwise. Graph the original and the rotated line.

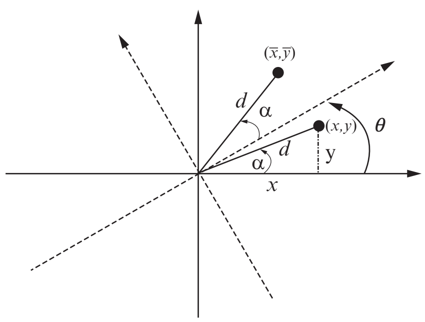

FIGURE 1.2 Rotating the xy-plane.

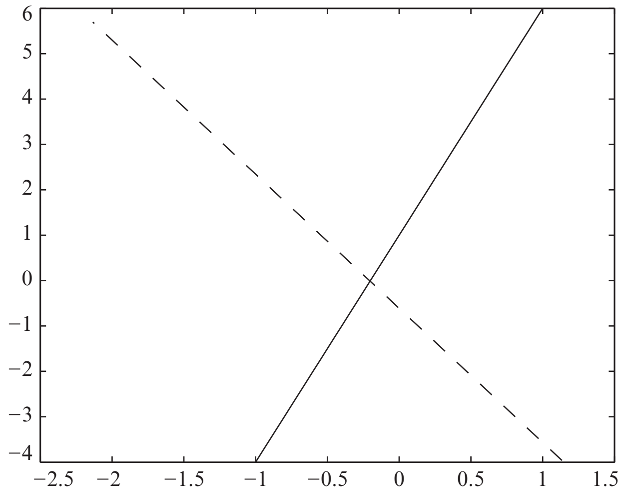

FIGURE 1.3 Rotated line

Since 30° is $\pi/6$ radians, the rotation matrix is

$$\begin{bmatrix} \cos\frac{\pi}{6} & -\sin\frac{\pi}{6} \\ \sin\frac{\pi}{6} & \cos\frac{\pi}{6} \end{bmatrix} = \begin{bmatrix} 0.866 & -0.5 \\ 0.5 & 0.866 \end{bmatrix}.$$

Now compute the rotation.

$$\begin{bmatrix} \overline{x} \\ \overline{y} \end{bmatrix} = \begin{bmatrix} 0.866 & -0.5 \\ 0.5 & 0.866 \end{bmatrix}\begin{bmatrix} x \\ 5x+1 \end{bmatrix} = \begin{bmatrix} -1.634x - 0.5 \\ 4.83x + 0.866 \end{bmatrix}.$$

Choose two points on the line $y = 5x + 1$, say (0, 1) and (1, 6), apply the transformation to these points, and determine two points on the line.

$$\begin{bmatrix} \overline{x_1} \\ \overline{y_1} \end{bmatrix} = \begin{bmatrix} -0.5 \\ 0.866 \end{bmatrix},$$
$$\begin{bmatrix} \overline{x_2} \\ \overline{y_2} \end{bmatrix} = \begin{bmatrix} -2.134 \\ 5.696 \end{bmatrix}.$$

Figure 1.3 is the graph of the original and the rotated line. ■

In three-dimensional Euclidean space, the equations

$$\overline{x} = x\cos\theta - y\sin\theta, \quad \overline{y} = x\sin\theta + y\cos\theta, \quad \overline{z} = z,$$
$$\overline{x} = x, \quad \overline{y} = y\cos\phi - z\sin\phi, \quad \overline{z} = y\sin\phi + z\cos\phi,$$
$$\overline{x} = x\cos\psi - z\sin\psi, \quad \overline{y} = y, \quad \overline{z} = x\sin\psi + z\cos\psi$$

correspond to rotations about the positive z-, x-, and y-axes, counterclockwise through θ, ϕ, ψ radians, respectively.

The product of two matrices is related to the product of the corresponding linear transformations:

If A is $m\times k$ and B is $k\times n$, the linear transformation $T_A T_B$ first performs transformation T_B, and then T_A. For instance, we might rotate about the x-axis, followed by a rotation about the z-axis. This transformation is in fact equal to the linear transformation T_{AB}, since

$$T_A T_B(x) = A(Bx) = (AB)x = T_{AB}(x).$$

The following example is useful for producing rotations in three-dimensional animated design (see Ref. [7, pp. 97-112]).

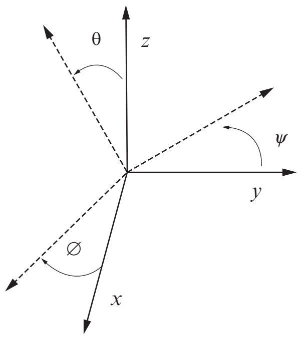

FIGURE 1.4 Rotate three-dimensional coordinate system.

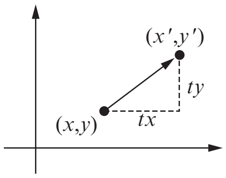

FIGURE 1.5 Translate a point in two dimensions.

Example 1.7. The linear transformation resulting from successively rotating three-dimensional space about the positive z, x, and y-axes, counterclockwise through θ, ϕ, ψ radians, respectively (Figure 1.4), is equal to T_{ABC}, where

$$C = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & -\sin\phi \\ 0 & \sin\phi & \cos\phi \end{bmatrix}, \quad A = \begin{bmatrix} \cos\psi & 0 & -\sin\psi \\ 0 & 1 & 0 \\ \sin\psi & 0 & \cos\psi \end{bmatrix}.$$

The matrix ABC is somewhat complex:

$$\begin{aligned} A(BC) &= \begin{bmatrix} \cos\psi & 0 & -\sin\psi \\ 0 & 1 & 0 \\ \sin\psi & 0 & \cos\psi \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \cos\phi\sin\theta & \cos\phi\cos\theta & -\sin\phi \\ \sin\phi\sin\theta & \sin\phi\cos\theta & \cos\phi \end{bmatrix} \\ &= \begin{bmatrix} \cos\psi\cos\theta - \sin\psi\sin\phi\sin\theta & -\cos\psi\sin\theta - \sin\psi\sin\phi\sin\theta & -\sin\psi\cos\phi \\ \cos\phi\sin\theta & \cos\phi\cos\theta & -\sin\phi \\ \sin\psi\cos\theta + \cos\psi\sin\phi\sin\theta & -\sin\psi\sin\theta + \cos\psi\sin\phi\cos\theta & \cos\psi\cos\phi \end{bmatrix}. \end{aligned}$$

■

Now consider a new problem. Reposition a point (x, y) along a straight line a distance of (tx, ty), where t is a scalar. The new location of the point is $(x + tx, y + ty)$ (Figure 1.5).

To determine a linear transformation, we use a 3×3 matrix

$$T = \begin{bmatrix} 1 & 0 & tx \\ 0 & 1 & ty \\ 0 & 0 & 1 \end{bmatrix}$$

and multiply T by the column vector $\begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$. Then

$$\begin{bmatrix} 1 & 0 & tx \\ 0 & 1 & ty \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} x + tx \\ y + ty \\ 1 \end{bmatrix}.$$

In order to combine translation and rotation using matrix multiplication, we need to create a 3×3 matrix that performs a two-dimensional rotation. Define

$$R = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Now,

$$\begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} x\cos\theta - y\sin\theta \\ x\sin\theta + y\cos\theta \\ 1 \end{bmatrix}.$$

In each case, we can ignore the z component of 1.

Example 1.8. Now we can perform an interesting and practical matrix calculation. Take the line $y = 5x + 1$ and rotate it 30° counterclockwise about the point (2, 11).

To solve this problem, first translate the point (2, 11) to (0, 0), rotate 30° counterclockwise, and then translate the point from the origin back to (2, 11) (Figure 1.6).

Here are the matrices involved.

$T_1 = \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & -11 \\ 0 & 0 & 1 \end{bmatrix}$. Translate (2, 11) to the origin using $tx = -2$ and $ty = -11$.

$R = \begin{bmatrix} \cos\frac{\pi}{6} & -\sin\frac{\pi}{6} & 0 \\ \sin\frac{\pi}{6} & \cos\frac{\pi}{6} & 0 \\ 0 & 0 & 1 \end{bmatrix}$. Rotate 30°

$T_2 = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 11 \\ 0 & 0 & 1 \end{bmatrix}$. Translate back to (2, 11) using $tx = 2$ and $ty = 11$.

Compute the product $F = T_2 R T_1$.

$$F = \begin{bmatrix} 0.8660 & -0.5000 & 5.7679 \\ 0.5000 & 0.8660 & 0.4737 \\ 0 & 0 & 1.0000 \end{bmatrix}.$$

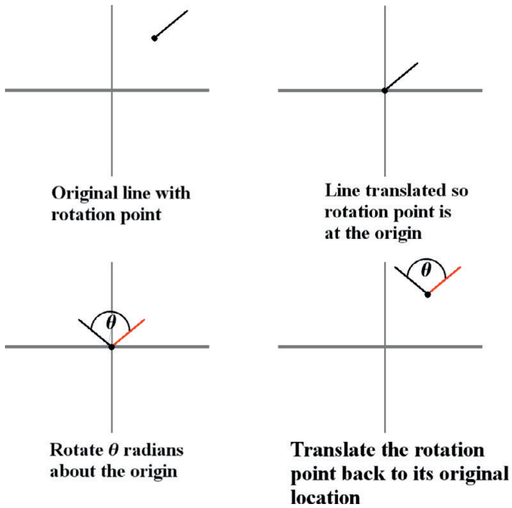

FIGURE 1.6 Rotate a line about a point.

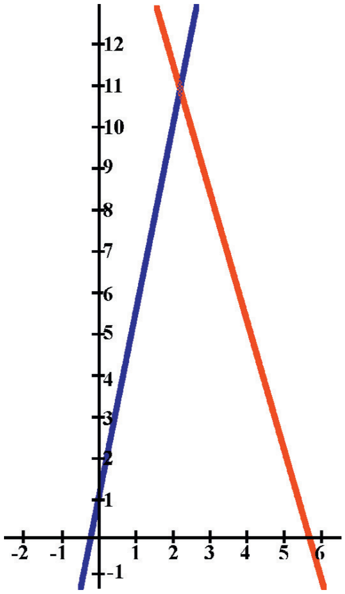

FIGURE 1.7 Rotation about an arbitrary point.

By computing the two points $F\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$ and $F\begin{bmatrix} 2 \\ 11 \\ 1 \end{bmatrix}$ on the rotated line and using the two point formula for the equation of a line, we obtain the equation $y = -2.9561x + 16.9121$.

Figure 1.7 is a plot of both the original and the rotated line. ■

1.3 POWERS OF MATRICES

Definition 1.10 (The identity matrix). The $n \times n$ matrix $I = [\delta_{ij}]$, defined by $\delta_{ij} = 1$ if $i = j$, $\delta_{ij} = 0$ if $i \neq j$, is called the $n \times n$ *identity* matrix of order n. In other words, the columns of the identity matrix of order n are the vectors

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}, e_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \\ 0 \end{bmatrix}, \ldots, e_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}.$$

For example, $I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. The identity matrix plays a critical role in linear algebra. When any $n \times n$ matrix A is multiplied by the identity matrix, either on the left or the right, the result is A. Thus, the identity matrix acts like 1 in the real number system. For example,

$$\begin{bmatrix} 2 & 6 & 1 \\ 7 & 2 & 9 \\ -1 & 5 & -4 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 2(1) + 6(0) + 1(0) & 2(0) + 6(1) + 1(0) & 2(0) + 6(0) + 1(1) \\ 7 & 2 & 9 \\ -1 & 5 & -4 \end{bmatrix} = \begin{bmatrix} 2 & 6 & 1 \\ 7 & 2 & 9 \\ -1 & 5 & -4 \end{bmatrix}.$$

Definition 1.11 (kth power of a matrix). If A is an $n \times n$ matrix, we define A^k as follows: $A^0 = I$ and $A^k = \underbrace{A \times A \times A \cdots A \times A}_{A \text{ occurs } k \text{ times}}$ for $k \geq 1$.

For example, $A^4 = A \times A \times A \times A$. Compute from left to right as follows:

$$A^2 = A \times A, \quad A^3 = (A)^2 \times A, \quad A^4 = (A)^3 \times A.$$

Example 1.9. The MATLAB exponentiation operator ^ applies to matrices.

```
>> A = [1 1;1 0]
A =
1 1
1 0
>> A^8
ans =
34 21
21 13
```

A is known as the *Fibonacci matrix*, since it generates elements from the famous Fibonacci sequence

$$0, 1, 1, 2, 3, 5, 8, 13, 21, 34, \ldots$$ ■

Example 1.10. Let $A = \begin{bmatrix} 7 & 4 \\ -9 & -5 \end{bmatrix}$. Let's investigate powers of A and see if we can find a formula for A^n.

$$\begin{aligned}
A^2 &= \begin{bmatrix} 7 & 4 \\ -9 & -5 \end{bmatrix}\begin{bmatrix} 7 & 4 \\ -9 & -5 \end{bmatrix} = \begin{bmatrix} 13 & 8 \\ -18 & -11 \end{bmatrix}, \\
A^3 &= \begin{bmatrix} 13 & 8 \\ -18 & -11 \end{bmatrix}\begin{bmatrix} 7 & 4 \\ -9 & -5 \end{bmatrix} = \begin{bmatrix} 19 & 12 \\ -27 & -17 \end{bmatrix}, \\
A^4 &= \begin{bmatrix} 19 & 12 \\ -27 & -17 \end{bmatrix}\begin{bmatrix} 7 & 4 \\ -9 & -5 \end{bmatrix} = \begin{bmatrix} 25 & 16 \\ -36 & -23 \end{bmatrix}, \\
A^5 &= \begin{bmatrix} 25 & 16 \\ -36 & -23 \end{bmatrix}\begin{bmatrix} 7 & 4 \\ -9 & -5 \end{bmatrix} = \begin{bmatrix} 31 & 20 \\ -45 & -29 \end{bmatrix}.
\end{aligned}$$

The elements in positions (1, 2) and (2, 1) follow a pattern. The element in position (1, 2) is always $4n$, and the element at position (2, 1) is always $-9n$. The element at (1, 1) is $6n + 1$, so we only need the pattern for the entry at (2, 2). It is always one (1) more than $-6n$, so it has the value $1 - 6n$. Here is the formula for A^n.

$$A^n = \begin{bmatrix} 1+6n & 4n \\ -9n & 1-6n \end{bmatrix} \quad \text{if } n \geq 1.$$

This is not a mathematical proof, just an example of pattern recognition. The result can be formally proved using mathematical induction (see Appendix B). ■

Our final example of matrix powers is a result from *graph theory*. A graph is a set of *vertices* and connections between them called *edges*. You have seen many graphs; for example, a map of the interstate highway system is a graph, as is the airline route map at the back of those boring magazines you find on airline flights. Consider the simple graph in Figure 1.8. A *path* from one vertex v to another vertex w is a sequence of edges that connect v and w. For instance, here are three paths from A to F: A-B-F, A-B-D-F, and A-B-C-E-B-F. The length of a path between v and w is the number of edges that must

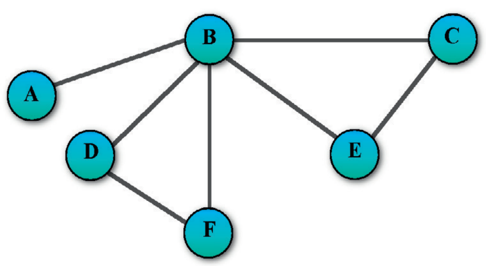

FIGURE 1.8 Undirected graph.

be crossed in moving from one to the other. For instance, in our three paths, the first has length 2, the second has length 3, and the third has length 5.

If a graph has n vertices, the *adjacency matrix* of the graph is an $n \times n$ matrix that specifies the location of edges. The concept is best illustrated by displaying the adjacency matrix for our six vertex graph, rather than giving a mathematical definition.

$$\text{Adj} = \begin{array}{c} \\ A \\ B \\ C \\ D \\ E \\ F \end{array} \begin{array}{c} \begin{array}{cccccc} A & B & C & D & E & F \end{array} \\ \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \end{bmatrix} \end{array}$$

A one (1) occurs in row A, column B, so there is an edge connecting A and B. Similarly, a one is in row E, column C, so there is an edge connecting E and C. There is no edge between A and D, so row A, column D contains zero (0).

There is a connection between the adjacency matrix of a graph and the number of possible paths between two vertices. Clearly, Adj^1 specifies all the paths of length 1 from one vertex to another (an edge).

If Adj *is the adjacency matrix for a graph, then* Adj^k *defines the number of possible paths of length k between any two vertices.* We will not attempt to prove this, but will use our graph as an example.

$$\text{Adj}^2 = \begin{bmatrix} 1 & 0 & 1 & 1 & 1 & 1 \\ 0 & 5 & 1 & 1 & 1 & 1 \\ 1 & 1 & 2 & 1 & 1 & 1 \\ 1 & 1 & 1 & 2 & 1 & 1 \\ 1 & 1 & 1 & 1 & 2 & 1 \\ 1 & 1 & 1 & 1 & 1 & 2 \end{bmatrix},$$

$$\text{Adj}^3 = \begin{bmatrix} 0 & 5 & 1 & 1 & 1 & 1 \\ 5 & 4 & 6 & 6 & 6 & 6 \\ 1 & 6 & 2 & 2 & 3 & 2 \\ 1 & 6 & 2 & 2 & 2 & 3 \\ 1 & 6 & 3 & 2 & 2 & 2 \\ 1 & 6 & 2 & 3 & 2 & 2 \end{bmatrix}.$$

By looking at Adj^2, we see that there is one path of length 2 between C and E, C-B-E, and two paths of length 2 connecting E to E (E-C-E, E-B-E). There are five (5) paths of length 3 between B and A (B-A-B-A, B-D-B-A, B-C-B-A, B-E-B-A, B-F-B-A). Note that if we reverse each path of length three from B to A, we have a path that starts at A and ends at B. Look carefully at Adj, Adj^2, and Adj^3 and notice that the entry at position (i, j) is always the same as the entry at (j, i). Such a matrix is termed *symmetric*. If you exchange rows and columns, the matrix remains the same. There are many applications of symmetric matrices in science and engineering.

1.4 NONSINGULAR MATRICES

Definition 1.12 (Nonsingular matrix). An $n \times n$ matrix A is called *nonsingular* or *invertible* if there exists an $n \times n$ matrix B such that

$$AB = BA = I.$$

If A does not have an inverse, A is called *singular*.

A matrix B such that $AB = BA = I$ is called an *inverse* of A. There can only be one inverse, as Theorem 1.3 shows.

Theorem 1.3. *A matrix A can have only one inverse.*

Proof. Assume that $AB = I$, $BA = I$, and $CA = AC = I$. Then,
$C(AB) = (CA)B$, and $CI = IB$, so $C = B$. □

When determining if B is the inverse of A, it is only necessary to verify $AB = I$ or $BA = I$. This is important because an procedure that computes the inverse of A need only to verify the product in one direction or the other.

Theorem 1.4. *If B is a matrix such that $BA = I$, then $AB = I$. Similarly, if $AB = I$, then $BA = I$.*

Proof. Assume that $BA = I$. Then $ABAB = AIB = AB$, and $ABAB - AB = 0$. Factor out AB, and $AB\,(AB - I) = 0$. Either $AB = 0$ or $AB - I = 0$. If $AB = 0$, then $ABA = 0\,(A) = 0$. But, $BA = I$, and so it follows that $A = 0$. The product of any matrix with the zero matrix is the zero matrix, so $BA = I$ is not possible. Thus, $AB - I = 0$, or $AB = I$. The fact that $AB = I$ implies $BA = I$ is handled in the same fashion. □

If we denote the inverse by A^{-1}, then

$$\left(A^{-1}\right) A = I,$$
$$A\left(A^{-1}\right) = I,$$

and it follows that

$$(A^{-1})^{-1} = A.$$

This says the inverse of A^{-1} is A itself.

The inverse has a number of other properties that play a role in developing results in linear algebra. For instance,

$$\left(B^{-1}A^{-1}\right)(AB) = B^{-1}\left(A^{-1}A\right)B = B^{-1}IB = B^{-1}B = I,$$

and so

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

By Theorem 1.4, we do not need to verify that $(AB)\left(B^{-1}A^{-1}\right) = I$.

Remark 1.2. The above result generalizes to a product of m nonsingular matrices: If $A_1, \ldots, A_m$ are nonsingular $n \times n$ matrices, then the product $A_1 \ldots A_m$ is also nonsingular. Moreover,

$$(A_1 \ldots A_m)^{-1} = A_m^{-1} \ldots A_1^{-1},$$

so the inverse of a product equals the product of the inverses *in the reverse order*.

Example 1.11. If A and B are $n \times n$ matrices satisfying $A^2 = B^2 = (AB)^2 = I$, show that $AB = BA$.
This says that A, B, and AB are each their own inverse, or $(AB)^{-1} = AB$, $B^{-1} = B$, and $A^{-1} = A$. Now,

$$AB = (AB)^{-1} = B^{-1}A^{-1} = BA,$$

and so $AB = BA$. ■

Normally, is it not true that $AB = BA$ for $n \times n$ matrices A and B. For instance,

$$\begin{bmatrix} 1 & 2 \\ -1 & 5 \end{bmatrix}\begin{bmatrix} 6 & 1 \\ -7 & 4 \end{bmatrix} = \begin{bmatrix} -8 & 9 \\ -41 & 19 \end{bmatrix},$$

and

$$\begin{bmatrix} 6 & 1 \\ -7 & 4 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ -1 & 5 \end{bmatrix} = \begin{bmatrix} 5 & 17 \\ -11 & 6 \end{bmatrix}.$$

Remark 1.3. We will show how to compute the inverse in Chapter 2; however, it is computationally expensive. The inverse is primarily a tool for developing other results.

A matrix having an inverse guarantees that a linear system has a unique solution.

Theorem 1.5. *If the coefficient matrix A of a system of n equations in n unknowns has an inverse, then the system $Ax = b$ has the unique solution $x = A^{-1}b$.*

Proof. **1.** (Uniqueness) Assume $Ax = b$. Then

$$\begin{aligned} A^{-1}(Ax) &= A^{-1}b, \\ \left(A^{-1}A\right)x &= A^{-1}b, \\ Ix &= A^{-1}b, \\ x &= A^{-1}b. \end{aligned}$$

2. (Existence) Assume $x = A^{-1}b$. Then

$$Ax = A\left(A^{-1}b\right) = \left(AA^{-1}\right)b = Ib = b.$$ □

A linear system $Ax = 0$ is said to be *homogeneous*. If A is nonsingular, then $x = A^{-1}0 = 0$, so the system has only 0 as its solution. It is said to have only the *trivial solution*.

Example 1.12. Consider the homogeneous system

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

A simple calculation verifies that $\begin{bmatrix} 1 & 3 \\ 1 & 2 \end{bmatrix}^{-1} = \begin{bmatrix} -2 & 3 \\ 1 & -1 \end{bmatrix}$, and so

$$\begin{gathered} \begin{bmatrix} -2 & 3 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} 1 & 3 \\ 1 & 2 \end{bmatrix} x = \begin{bmatrix} -2 & 3 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \\ Ix = x = 0. \end{gathered}$$ ■

There are some cases where there is an explicit formula for the inverse matrix. In particular, we can demonstrate a formula for the inverse of a 2×2 matrix subject to a condition.

Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ with $ad - bc \neq 0$ and let $B = (1/(ad - bc))\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$. Perform a direct calculation of AB.

$$\begin{aligned} \begin{bmatrix} a & b \\ c & d \end{bmatrix}\left(\frac{1}{ad-bc}\right)\begin{bmatrix} d & -b \\ -c & a \end{bmatrix} &= \left(\frac{1}{ad-bc}\right)\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \\ &= \left(\frac{1}{ad-bc}\right)\begin{bmatrix} ad-bc & -ab+ab \\ cd-cd & ad-bc \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I \end{aligned}$$

Remark 1.4. The expression $ad - bc$ is called the *determinant* of A and is denoted by $\det(A)$. Later we will see that A has an inverse if and only if $\det A \neq 0$.

The MATLAB function that computes the inverse of an $n \times n$ matrix is `inv`. If A is an $n \times n$ matrix, then

```
>> B = inv(A);
```

computes A^{-1}.

Example 1.13. The following MATLAB statements demonstrate the use of `inv` and verify that $(AB)^{-1} = B^{-1}A^{-1}$ for two particular matrices.

```
>> format rational;
>> A = [1 3 -1; 4 1 6; 0 2 3]
A =
1 3 -1
4 1 6
0 2 3
```

```
>> A_inv = inv(A)
A_inv =
9/53  11/53 -19/53
12/53 -3/53 10/53
-8/53 2/53  11/53

>> B = [1 4 0; 3 5 1; 2 -7 8]
B =
1  4 0
3  5 1
2 -7 8

>> B_inv = inv(B)
B_inv =
-47/41 32/41 -4/41
22/41 -8/41  1/41
31/41 -15/41 7/41

>> inv(A*B)
ans =
-7/2173 -621/2173 1169/2173
94/2173 268/2173 -487/2173
43/2173 400/2173 -662/2173

>> B_inv * A_inv
ans =
-7/2173 -621/2173 1169/2173
94/2173 268/2173 -487/2173
43/2173 400/2173 -662/2173
```

■

1.5 THE MATRIX TRANSPOSE AND SYMMETRIC MATRICES

There is another property of a matrix that we will use extensively, the matrix transpose.

Definition 1.13 (The transpose of a matrix). Let A be an $m\times n$ matrix. Then A^{T}, the *transpose* of A, is the matrix obtained by interchanging the rows and columns of A. In other words if $A = [a_{ij}]$, then $\left(A^{\mathrm{T}}\right)_{ij} = a_{ji}$. Consequently A^{T} is an $n\times m$ matrix.

For instance, if

$$A = \begin{bmatrix} 1 & 9 & 0 \\ 3 & 7 & 15 \\ 4 & 8 & 1 \\ -7 & 12 & 3 \end{bmatrix},$$

then

$$A^{\mathrm{T}} = \begin{bmatrix} 1 & 3 & 4 & -7 \\ 9 & 7 & 8 & 12 \\ 0 & 15 & 1 & 3 \end{bmatrix}.$$

Theorem 1.6. *The transpose operation has the following properties:*

1. $\left(A^{\mathrm{T}}\right)^{\mathrm{T}} = A$;
2. $(A \pm B)^{\mathrm{T}} = A^{\mathrm{T}} \pm B^{\mathrm{T}}$ *if A and B are $m\times n$;*
3. $(sA)^{\mathrm{T}} = sA^{\mathrm{T}}$ *if s is a scalar;*
4. $(AB)^{\mathrm{T}} = B^{\mathrm{T}}A^{\mathrm{T}}$ *if A is $m\times k$ and B is $k\times n$;*

5. *A is nonsingular, then* A^{T} *is also nonsingular and* $\left(A^{\mathrm{T}}\right)^{-1} = \left(A^{-1}\right)^{\mathrm{T}}$.
6. $x^{\mathrm{T}}x = x_1^2 + \cdots + x_n^2$ *if* $x = [x_1, \ldots, x_n]^{\mathrm{T}}$ *is a column vector.*

Proof. We will verify 5 and 6 and leave the remaining properties to the exercises.
Property 5: $A^{\mathrm{T}}\left(A^{-1}\right)^{\mathrm{T}} = \left(A^{-1}A\right)^{\mathrm{T}}$ by property 4. Therefore, $A^{\mathrm{T}}\left(A^{-1}\right)^{\mathrm{T}} = I^{\mathrm{T}} = I$.

Property 6: $x^{\mathrm{T}}x = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1^2 + \cdots + x_n^2.$ □

There is a frequently occurring class of matrices defined in terms of the transpose operation.

Definition 1.14 (Symmetric matrix). A matrix A is *symmetric* if $A^{\mathrm{T}} = A$. In other words A is square $(n \times n)$ and $a_{ji} = a_{ij}$ for all $1 \le i \le n, 1 \le j \le n$. Another way of looking this is that when the rows and columns are interchanged, the resulting matrix is A. For instance,

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

is a general 2×2 symmetric matrix.

Example 1.14. $A = \begin{bmatrix} 1 & 8 & 12 & 3 \\ 8 & 5 & 1 & 10 \\ 12 & 1 & 6 & 9 \\ 3 & 10 & 9 & 2 \end{bmatrix}$ is a symmetric matrix. Notice row 1 has the same entries as column 1, row 2 has the same entries as column 2, and so forth. ■

The following proposition proves a property of $A^{\mathrm{T}}A$ that is critical for the computation of what are termed singular values in Chapter 7.

Proposition 1.1. *If A is an* $m \times n$ *matrix, then* $A^{\mathrm{T}}A$ *is a symmetric matrix.*

Proof. $A^{\mathrm{T}}A$ is an $n \times n$ matrix, since A^{T} has dimension $n \times m$, and A has dimension $m \times n$. $A^{\mathrm{T}}A$ is symmetric, since

$$\left(A^{\mathrm{T}}A\right)^{\mathrm{T}} = A^{\mathrm{T}}\left(A^{\mathrm{T}}\right)^{\mathrm{T}} = A^{\mathrm{T}}A.$$ □

Example 1.15. In MATLAB, the transpose operator is $'$.

```
>> A = [1 8 -1; 3 -9 15; -1 5 3]

A =
     1     8    -1
     3    -9    15
    -1     5     3

>> A_TA = A'*A

A_TA =
    11   -24    41
   -24   170  -128
    41  -128   235

>> A_TA - A_TA'

ans =
     0     0     0
     0     0     0
     0     0     0
```

■

1.6 CHAPTER SUMMARY

Matrix Arithmetic

This chapter defines a matrix, introduces matrix notation, and presents matrix operations, including matrix multiplication. To multiply matrices A and B, the number of columns of A must equal the number of rows of B. It should be emphasized that multiplication of large matrices, most of whose elements are nonzero, is a time-consuming computation. When matrices in applications are very large, they normally consist primarily of zero entries, and are termed sparse. Although matrix multiplication obeys most of the laws of arithmetic, it is not commutative; that is, if A and B are $n \times n$ matrices, AB is rarely equal to BA.

A vector is a matrix having one row or one column. In this book, we will primarily use column vectors such as

$$\begin{bmatrix} -1 \\ 3 \\ 7 \end{bmatrix}.$$

The trace of an $n \times n$ matrix is the sum of its diagonal elements a_{ii}, $1 \le i \le n$, or trace $(A) = \sum_{i=1}^{n} a_{ii}$. The trace occurs in many matrix formulas, and we will encounter it in later chapters. It is important to note that even though $AB \neq BA$ in general, in fact trace (AB) = trace (BA).

A primary topic in this book is the solution of linear systems of equations, and we write them using matrix notation; for instance, the system

$$\begin{aligned} x_1 - x_2 + 5x_3 &= 1, \\ -2x_1 + 4x_2 + x_3 &= 0, \\ 7x_1 - 2x_2 - 6x_3 &= 8 \end{aligned}$$

using matrix notation is

$$\begin{bmatrix} 1 & -1 & 5 \\ -2 & 4 & 1 \\ 7 & -2 & -6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 8 \end{bmatrix}.$$

Linear Transformations

If A is an $m \times n$ matrix, a linear transformation is a mapping from $\mathbb{R}^n$ to $\mathbb{R}^m$ defined by $y = Ax$. We will deal with linear transformations throughout the remainder of this book. An excellent example is a two-dimensional linear transformation of the form

$$y = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

that rotates the vector $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ counter clockwise through angle θ. Such linear transformations perform rotation, displacement, and scaling of objects in computer graphics.

Powers of Matrices

There are numerous applications of matrix powers, A^k, $k \ge 0$. Given an undirected graph, powers of the adjacency matrix provide a count of the number of paths between any two vertices. We will see in Chapters 21 and 22 that a sequence of the form $Ax_0, A^2x_0, \ldots, A^{k-1}x_0$ forms the basis for series of important methods that solve linear systems and compute eigenvalues of large, sparse matrices. We will discuss solving linear systems in Chapter 2 and eigenvalues in Chapter 5.

Nonsingular Matrices

The inverse of a matrix is of great theoretical importance in linear algebra. An $n \times n$ matrix A has inverse A^{-1} if

$$A^{-1}A = AA^{-1} = I,$$

where I is the identity matrix

$$\begin{bmatrix} 1 & 0 & \dots & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & \ddots & \dots & 0 \\ \vdots & \vdots & \ddots & 1 & \vdots \\ 0 & 0 & \dots & 0 & 1 \end{bmatrix}.$$

When it exists, the inverse is unique, and the matrix is termed nonsingular. Not all matrices have an inverse, and such matrices are said to be singular. A linear system $Ax = b$ has a unique solution $x = A^{-1}b$ when A is nonsingular. A very important result is that a homogeneous system of the form $Ax = 0$ has only the solution $x = 0$ if A is nonsingular.

The Matrix Transpose and Symmetric Matrices

The transpose of an $m \times n$ matrix A, named A^{T}, is the $n \times m$ matrix obtained by exchanging the rows and columns of A. Theorem 1.6 lists properties of the transpose.

An important class of matrices are the symmetric matrices. A square matrix A is symmetric if $A^{\mathrm{T}} = A$, and this means that $a_{ij} = a_{ji}$, $1 \le i, j \le n$. Many problems in engineering and science involve symmetric matrices, and entire sections of this book deal with them. As you will see, when a problem involves a symmetric matrix, this normally leads to a faster and more accurate solution.

It is of the utmost importance that you remember the relationship $(AB)^{\mathrm{T}} = B^{\mathrm{T}}A^{\mathrm{T}}$, as we will use it again and again throughout this book. Here is an interesting fact we will use beginning in Chapter 7. If A is any $m \times n$ matrix, then $A^{\mathrm{T}}A$ is symmetric, since

$$\left(A^{\mathrm{T}}A\right)^{\mathrm{T}} = A^{\mathrm{T}}\left(A^{\mathrm{T}}\right)^{\mathrm{T}} = A^{\mathrm{T}}A.$$

1.7 PROBLEMS

1.1 For

$$A = \begin{bmatrix} 1 & 8 & -1 \\ 0 & 6 & -7 \\ 2 & 4 & 12 \end{bmatrix}, \quad B = \begin{bmatrix} 6 & -1 & 25 \\ 14 & -6 & 0 \\ -9 & 15 & 25 \end{bmatrix}$$

compute the following:

a. $A - B$

b. $8A$

c. $5A + 7B$

1.2 Using the matrices A and B from Problem 1.1, compute AB. Do not use a computer program. Do it with pencil and paper.

1.3 Let A, B, C, D be matrices defined by

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

$$C = \begin{bmatrix} -3 & -1 \\ 2 & 1 \\ 4 & 3 \end{bmatrix}, \quad D = \begin{bmatrix} 4 & -1 \\ 2 & 0 \end{bmatrix}.$$

Which of the following matrices are defined? Compute those matrices which are defined.

$$A + B, A + C, AB, BA, CD, DC, D^2.$$

1.4 Let $A = \begin{bmatrix} -1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$. Show that if B is a 3×2 matrix such that $AB = I$, then

$$B = \begin{bmatrix} a & b \\ -a-1 & 1-b \\ a+1 & b \end{bmatrix}$$

for suitable numbers a and b. Use the associative law to show that $(BA)^2 B = B$.

1.5 If $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, show that $A^2 - (a+d)A + (ad - bc)I_2 = 0$.

1.6 A square matrix $D = [d_{ij}]$ is called *diagonal* if $d_{ij} = 0$ for $i \neq j$; that is the *off-diagonal* elements are zero. Show that premultiplication of a matrix A by a diagonal matrix D results in matrix DA whose rows are the rows of A multiplied by the respective diagonal elements of D.

1.7 Write the following linear algebraic system in matrix form.

$$\begin{aligned} 5x_1 + 6x_2 - x_3 + 2x_4 &= 1, \\ -x_1 + 2x_2 + x_3 - 9x_4 &= 8, \\ 2x_1 - x_3 &= -3, \\ 3x_2 + 28x_3 - 2x_4 &= 0. \end{aligned}$$

1.8 Write this matrix equation as a system of equations.

$$\begin{bmatrix} 1 & 0 & 9 \\ -8 & 3 & 45 \\ 12 & -6 & 55 \end{bmatrix} x = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}.$$

1.9 Define the linear transformation $T_A : \mathbb{R}^5 \to \mathbb{R}^5$ using the matrix

$$A = \begin{bmatrix} 1 & 7 & 0 & 0 & 0 \\ 4 & 5 & 8 & 0 & 0 \\ 0 & 6 & 1 & 1 & 0 \\ 0 & 0 & 7 & 3 & -9 \\ 0 & 0 & 0 & 1 & 2 \end{bmatrix}.$$

A is termed a *tridiagonal matrix*, since the only non-zero entries are along the main diagonal, the diagonal below, and the diagonal above.

a. Compute $A \begin{bmatrix} 0 \\ 1 \\ -1 \\ 3 \\ 2 \end{bmatrix}$.

b. Compute $A \begin{bmatrix} 6 \\ 0 \\ 1 \\ 3 \\ 0 \end{bmatrix}$.

c. Compute the general product

$$\begin{bmatrix} a_{11} & a_{12} & 0 & 0 \\ a_{21} & a_{22} & a_{23} & 0 \\ 0 & a_{32} & a_{33} & a_{34} \\ 0 & 0 & a_{43} & a_{44} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}.$$

d. Propose a formula for the product $y = Ax$ of an $n \times n$ tridiagonal matrix A and an $n \times 1$ column vector x. Use the result of part (c) to help you formulate your answer.

1.10 Rotate the line $y = -x + 3$ 30° counterclockwise about the origin, and graph the two lines on the same set of axes.

1.11 Rotate the line $y = -x + 3$ 60° counterclockwise about the point $(4, -1)$, and graph the two lines on the same set of axes.

1.12 Let $A = \begin{bmatrix} 1 & 4 \\ -3 & 1 \end{bmatrix}$. Show that A is nonsingular by verifying that

$$A^{-1} = \begin{bmatrix} \frac{1}{13} & -\frac{4}{13} \\ \frac{3}{13} & \frac{1}{13} \end{bmatrix}.$$

1.13 Let $A = \begin{bmatrix} 1 & 3 \\ 2 & 0 \end{bmatrix}$

a. Find the inverse of the matrix

b. Use A^{-1} to solve the system

$$\begin{aligned} x_1 + 3x_2 &= 6, \\ 2x_1 - 9x_2 &= 1. \end{aligned}$$

1.14 If $A = \begin{bmatrix} 1 & 4 \\ -3 & 1 \end{bmatrix}$.

a. Verify that $A^2 - 2A + 13I = 0$

b. Show that $A^{-1} = -\frac{1}{13}(A - 2I)$.

1.15 Let $A = \begin{bmatrix} 1 & 1 & -1 \\ 0 & 0 & 1 \\ 2 & 1 & 2 \end{bmatrix}$. Verify that $A^3 = 3A^2 - 3A + I$.

1.16 Let A be an $n \times n$ matrix.

a. If $A^2 = 0$, prove that A is singular. Start by assuming A^{-1} exists. Compute $A^{-1}\,(A)^2$and deduce that A must be singular.

b. If $A^2 = A$ and $A \neq I$, prove that A is singular. Use the same logic as in part (a).

1.17 If $X = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}$ and $Y = \begin{bmatrix} -1 \\ 3 \\ 4 \end{bmatrix}$, find $XX^{\mathrm{T}}, X^{\mathrm{T}}X, YY^{\mathrm{T}}, Y^{\mathrm{T}}Y$.

1.18 For matrices A and B, show that $(AB)^{\mathrm{T}} = B^{\mathrm{T}}A^{\mathrm{T}}$.

$$A = \begin{bmatrix} 1 & 4 & -1 \\ 0 & 7 & 1 \\ 1 & 7 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 2 & 6 \\ 1 & -7 & 3 \\ 0 & 1 & 2 \end{bmatrix}.$$

1.19 If A is a symmetric $n \times n$ matrix and B is $n \times m$, prove that $B^{\mathrm{T}}AB$ is a symmetric $m \times m$ matrix.

1.20 Show that $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$ is its own inverse.

1.21 It is not usually the case for $n \times n$ matrices A and B that $AB = BA$. For instance,

$$A = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}, \quad B = \begin{bmatrix} 7 & 3 \\ 1 & 8 \end{bmatrix}, \quad AB = \begin{bmatrix} 9 & 19 \\ 3 & 24 \end{bmatrix}, \quad BA = \begin{bmatrix} 7 & 23 \\ 1 & 26 \end{bmatrix}.$$

Let A and B be $n \times n$ diagonal matrices:

$$A = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix}, \quad B = \begin{bmatrix} b_{11} & 0 & \cdots & 0 \\ 0 & b_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & b_{nn} \end{bmatrix}.$$

Show that $AB = BA$.

1.22 Prove the following formulas are satisfied by the trace matrix operator.

a. trace $(A + B) = \text{trace}\,(A) + \text{trace}\,(B)$

b. trace $(cA) = c\ \text{trace}\,(A)$, where c is a scalar.

1.23 For an arbitrary $n \times 1$ column vector and an $n \times n$ matrix A, show that $x^{\mathrm{T}}Ax$ is a real number. This is called a *quadratic form*. For $x = [1\ 3\ 9]^{\mathrm{T}}$, compute $x^{\mathrm{T}}Ax$ for the matrix

$$A = \begin{bmatrix} 1 & -8 & 3 \\ 4 & 0 & -1 \\ 3 & 5 & 7 \end{bmatrix}.$$

1.24 Prove the following properties of the matrix transpose operator.

a. $\left(A^{\mathrm{T}}\right)^{\mathrm{T}} = A$.

b. $(A \pm B)^{\mathrm{T}} = A^{\mathrm{T}} \pm B^{\mathrm{T}}$ if A and B are $m \times n$.

c. $(sA)^{\mathrm{T}} = sA^{\mathrm{T}}$ if s is a scalar.

1.25 Prove that $(AB)^{\mathrm{T}} = B^{\mathrm{T}}A^{\mathrm{T}}$ if A is $m \times n$ and B is $n \times p$. Hint: Use the definition of matrix multiplication and the fact that taking the transpose means element a_{ij} of A is the element at row j, column i of A^{T}.

1.7.1 MATLAB Problems

1.26 For this exercise, use the MATLAB command `format rational` so the computations are done using rational arithmetic. Find the inverse of matrices A and B.

$$A = \begin{bmatrix} 1 & 4 & 1 \\ 1 & 3 & 2 \\ -1 & 2 & 7 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 & 1 \\ 2 & 5 & 12 \\ -9 & 1 & 1 \end{bmatrix}$$

Verify that $(AB)^{-1} = B^{-1}A^{-1}$.

1.27 Use MATLAB to find the inverse of the matrix $A = \begin{bmatrix} 1 & 3 & -1 & -9 \\ 0 & 3 & 0 & 1 \\ 12 & 8 & -11 & 0 \\ 2 & 1 & 5 & 3 \end{bmatrix}$.

1.28 The $n \times n$ Hilbert matrices are defined by $H\,(i,j) = 1/(i + j - 1), 1 \leq i,j \leq n$. For instance, here is the 5×5 Hilbert matrix.

$$H = \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \frac{1}{6} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \frac{1}{6} & \frac{1}{7} \\ \frac{1}{4} & \frac{1}{5} & \frac{1}{6} & \frac{1}{7} & \frac{1}{8} \\ \frac{1}{5} & \frac{1}{6} & \frac{1}{7} & \frac{1}{8} & \frac{1}{9} \end{bmatrix}.$$

Systems of the form $Hx = b$, where H is a Hilbert matrix are notoriously difficult to solve because they are *ill-conditioned*. This means that a solution can change widely with only a small change in the elements of b or H. Chapter 10 discusses ill-conditioned matrices. The MATLAB command `hilb` builds an $n \times n$ Hilbert matrix. For instance to find the 6×6 Hilbert matrix, execute

```
>> H = hilb(6);
```

a. The command

```
format shortg
```

causes output of the best of fixed or floating point format with 5 digits. Using this format, compute the inverse of the 6×6 Hilbert matrix. What makes you suspicious that it is ill-conditioned?

b. The exact inverse of any Hilbert matrix consists entirely of integer entries. Using the Symbolic Toolbox will provide an exact answer. If your MATLAB distribution has this software, use the help system to determine how to use the commands `syms` and `sym`. Determine the exact value of H^{-1}.

1.29 **a.** Write a MATLAB `function t = tr(A)` that computes the trace of matrix A. Test to make sure A is a square matrix.

b. Use `tr` to compute the trace for the matrix of Problem 1.9 and the Hilbert matrix of order 15 (Problem 1.28). Verify your result by using the MATLAB command `trace`.

1.30 This problem uses the result of Problem 1.9(d).

a. Write a MATLAB `function y = triprod(A,x)` that forms the product of an $n \times n$ tridiagonal matrix A with an $n \times 1$ column vector x.

b. Test the function using the matrix and vectors specified in Problem 1.9, parts (a), and (b).