CHAPTER 2

Matrices and Systems of Linear Equations

Algebra is the intellectual instrument which has been created for rendering clear the quantitative aspects of the world. — Alfred North Whitehead

We will see in the later chapters that most problems in linear algebra can be reduced to questions regarding the solutions of systems of linear equations. In preparation for this, the next two chapters provide a detailed introduction to the theory and solution techniques for such systems. An example of a linear system of equations in the unknowns x_1, x_2, x_3 is

$$\begin{aligned} 3x_1 + 4x_2 - 7x_3 &= 5, \\ 2x_1 - 3x_2 + 9x_3 &= 7, \\ 7x_1 + 2x_2 - 3x_3 &= 4. \end{aligned}$$

We see that this system is completely determined by the array of numbers

$$\begin{bmatrix} 3 & 4 & -7 & 5 \\ 2 & -3 & 9 & 7 \\ 7 & 2 & -3 & 4 \end{bmatrix},$$

which contains the coefficients of the unknowns on the left-hand side of the system and the numbers appearing on the right-hand side of the system. Such an array is an example of a matrix. In this chapter we see that, in general, linear systems of equations are best represented in terms of matrices and that, once such a representation has been made, the set of all solutions to the system can be easily determined. In the first few sections of this chapter we therefore introduce the basics of matrix algebra. We then apply matrices to solve systems of linear equations. In Chapter 7, we will see how matrices also give a natural framework for formulating and solving systems of linear differential equations.

2.1 Matrices: Definitions and Notation

We begin our discussion of matrices with a definition.

DEFINITION 2.1.1

An $m \times n$ (read "m by n") **matrix** is a rectangular array of numbers arranged in m horizontal rows and n vertical columns. Matrices are usually denoted by uppercase letters, such as A and B. The entries in the matrix are called the **elements** of the matrix.

Example 2.1.2 The following are examples of a 2×3 and a 3×3 matrix, respectively:

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

□

We will use the index notation to denote the elements of a matrix. According to this notation, the element in the ith row and jth column of the matrix A will be denoted a_{ij}. Thus, for the matrices in the previous example we have

$$a_{13} = \frac{1}{5}, \qquad a_{22} = -\frac{3}{7}, \qquad b_{23} = -1, \qquad \text{and so on.}$$

Using the index notation, a general $m \times n$ matrix A is written

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

or, in a more abbreviated form, $A = [a_{ij}]$.

Remark The expression $m \times n$ representing the number of rows and columns of a general matrix A is sometimes informally called the **size** of the matrix A. The numbers m and n themselves are sometimes called the **dimensions**[1] of the matrix A.

Next we define what is meant by equality of matrices.

DEFINITION 2.1.3

Two matrices A and B are **equal**, written $A = B$, if

1. They both have the same size, $m \times n$.
2. All corresponding elements in the matrices are equal: $a_{ij} = b_{ij}$ for all i and j with $1 \le i \le m$ and $1 \le j \le n$.

[1] Be careful not to confuse this usage of the term with the *dimension* of a vector space, which will be introduced in Chapter 4.

According to Definition 2.1.3, even though the matrices

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

contain the same six numbers, and therefore store the same basic information, they are not equal as matrices.

Row Vectors and Column Vectors

Of particular interest to us in the future will be $1 \times n$ and $n \times 1$ matrices. For this reason we give them special names.

DEFINITION 2.1.4

A $1 \times n$ matrix is called a **row** n**-vector**. An $n \times 1$ matrix is called a **column** n**-vector**. The elements of a row or column n-vector are called the **components** of the vector.

Remarks

1. We can refer to the objects just defined simply as row vectors and column vectors if the value of n is clear from the context.
2. We will see later in this chapter that when a system of linear equations is written using matrices, the basic unknown in the reformulated system is a column vector. A similar formulation will also be given in Chapter 7 for systems of differential equations.

Example 2.1.5 The matrix $\mathbf{a} = \begin{bmatrix} \frac{2}{3} & -\frac{1}{5} & \frac{4}{7} \end{bmatrix}$ is a row 3-vector and

$$\mathbf{b} = \begin{bmatrix} 1 \\ -1 \\ 3 \\ 4 \end{bmatrix}$$

is a column 4-vector. □

As indicated here, we usually denote a row or column vector by a lowercase letter in **bold** print.

Associated with any $m \times n$ matrix are m row n-vectors and n column m-vectors. These are referred to as the **row vectors** of the matrix and the **column vectors** of the matrix, respectively.

Example 2.1.6 Associated with the matrix

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

are the row 4-vectors

$$\begin{bmatrix} -2 & 1 & 3 & 4 \end{bmatrix}, \qquad \begin{bmatrix} 1 & 2 & 1 & 1 \end{bmatrix}, \qquad \text{and} \quad \begin{bmatrix} 3 & -1 & 2 & 5 \end{bmatrix},$$

and the column 3-vectors

$$\begin{bmatrix} -2 \\ 1 \\ 3 \end{bmatrix}, \quad \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}, \quad \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}, \quad \text{and} \quad \begin{bmatrix} 4 \\ 1 \\ 5 \end{bmatrix}. \qquad \Box$$

Conversely, if $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ are each column m-vectors, then we let $[\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n]$ denote the $m \times n$ matrix whose column vectors are $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$. Similarly, if $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_m$ are each row n-vectors, then we write

$$\begin{bmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \\ \vdots \\ \mathbf{b}_m \end{bmatrix}$$

for the $m \times n$ matrix with row vectors $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_m$. The reader should observe that a list of vectors arranged in a row will always consist of column vectors, while a list of vectors arranged in a column will always consist of row vectors.

Example 2.1.7 If $\mathbf{a}_1 = \begin{bmatrix} \frac{1}{5} \\ \frac{2}{3} \end{bmatrix}$, $\mathbf{a}_2 = \begin{bmatrix} \frac{4}{7} \\ \frac{5}{9} \end{bmatrix}$, and $\mathbf{a}_3 = \begin{bmatrix} -\frac{1}{3} \\ \frac{3}{11} \end{bmatrix}$, then

$$[\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3] = \begin{bmatrix} \frac{1}{5} & \frac{4}{7} & -\frac{1}{3} \\ \frac{2}{3} & \frac{5}{9} & \frac{3}{11} \end{bmatrix}. \qquad \Box$$

DEFINITION 2.1.8

If we interchange the row vectors and column vectors in an $m \times n$ matrix A, we obtain an $n \times m$ matrix called the **transpose** of A. We denote this matrix by A^T. In index notation, the (i, j)th element of A^T, denoted a_{ij}^T, is given by

$$a_{ij}^T = a_{ji}.$$

Example 2.1.9 If

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

then

$$A^T = \begin{bmatrix} 1 & 0 \\ 2 & 3 \\ 6 & 4 \\ 2 & 7 \end{bmatrix}.$$

If

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

then

$$A^T = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 0 & 4 \\ 5 & 7 & 9 \end{bmatrix}.$$

□

Square Matrices

An $n \times n$ matrix is called a **square matrix**, since it has the same number of rows as columns. If A is a square matrix, then the elements a_{ii}, $1 \leq i \leq n$, make up the **main diagonal**, or **leading diagonal**, of the matrix. (See Figure 2.1.1 for the 3×3 case.)

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

Figure 2.1.1: The main diagonal of a 3×3 matrix.

The sum of the main diagonal elements of an $n \times n$ matrix A is called the **trace** of A and is denoted $\text{tr}(A)$. Thus,

$$\text{tr}(A) = a_{11} + a_{22} + \cdots + a_{nn}.$$

An $n \times n$ matrix A is said to be **lower triangular** if $a_{ij} = 0$ whenever $i < j$ (zeros everywhere above (i.e.. "northeast of") the main diagonal), and it is said to be **upper triangular** if $a_{ij} = 0$ whenever $i > j$ (zeros everywhere below (i.e., "southwest of") the main diagonal). The following are examples of an upper triangular and lower triangular matrix, respectively:

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

Observe that the transpose of a lower (upper) triangular matrix is an upper (lower) triangular matrix.

If every element on the main diagonal of a lower (upper) triangular matrix is a 1, the matrix is called a **unit lower (upper) triangular matrix**.

An $n \times n$ matrix $D = [d_{ij}]$ that has all *off-diagonal* elements equal to zero is called a **diagonal matrix**. Note that a matrix D is a diagonal matrix if and only if D is simultaneously upper and lower triangular. Such a matrix is completely determined by giving its main diagonal elements, since $d_{ij} = 0$ whenever $i \neq j$. Consequently, we can specify a diagonal matrix in the compact form

$$D = \text{diag}(d_1, d_2, \ldots, d_n),$$

where d_i denotes the diagonal element d_{ii}.

Example 2.1.10 The 4×4 diagonal matrix $D = \text{diag}(1, 2, 0, 3)$ is

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

□

The transpose naturally picks out two important types of square matrices as follows.

DEFINITION 2.1.11

1. A square matrix A satisfying $A^T = A$ is called a **symmetric matrix**.
2. If $A = [a_{ij}]$, then we let $-A$ denote the matrix with elements $-a_{ij}$. A square matrix A satisfying $A^T = -A$ is called a **skew-symmetric** (or **anti-symmetric**) **matrix**.

Example 2.1.12 The matrix

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

is symmetric, whereas

$$B = \begin{bmatrix} 0 & -1 & -5 & 3 \\ 1 & 0 & 1 & -2 \\ 5 & -1 & 0 & 7 \\ -3 & 2 & -7 & 0 \end{bmatrix}$$

is skew-symmetric. □

Notice that the main diagonal elements of the skew-symmetric matrix in the preceding example are all zero. This is true in general, since if A is a skew-symmetric matrix, then $a_{ij} = -a_{ji}$, which implies that when $i = j$, $a_{ii} = -a_{ii}$, so that $a_{ii} = 0$.

Matrix and Vector Functions

Later in the text we will be concerned with systems of two or more differential equations. The most effective way to study such systems, as it turns out, is to represent the system using matrices and vectors. However, we will need to allow the elements of the matrices and vectors that arise to contain *functions* of a single variable, not just real or complex numbers. This leads to the following definition, reminiscent of Definition 2.1.1.

DEFINITION 2.1.13

An $m \times n$ **matrix function** A is a rectangular array with m rows and n columns whose elements are functions of a single real variable t.

Example 2.1.14 Here are two examples of matrix functions:

$$A(t) = \begin{bmatrix} t^3 & t - \cos t & 5 \\ e^{t^2} & \ln(t+1) & te^t \end{bmatrix} \quad \text{and} \quad B(t) = \begin{bmatrix} 5 - t + t^2 & \sin(e^{2t}) \\ -1 & \tan t \\ 6 & 6 - t \end{bmatrix}.$$

A matrix function $A(t)$ is defined only for real values of t such that *all* elements in $A(t)$ assume a well-defined value. The function A is defined only for real values of t with

$t > -1$, since $\ln(t+1)$ is defined only for $t > -1$. The reader should determine the values of t for which the matrix function B is defined. □

Remark It is possible, of course, to consider matrix functions of more than one variable. However, this will not be particularly relevant for our purposes in this text.

Finally in this section, we have the following special type of matrix function.

DEFINITION 2.1.15

An $n \times 1$ matrix function is called a **column n-vector function**.

For instance, $\begin{bmatrix} t^2 \\ -6te^t \end{bmatrix}$ is a column 2-vector function.[2]

Exercises for 2.1

Key Terms

Matrices, Elements, Size (dimensions) of a matrix, Row vector, Column vector, Square matrix, Main diagonal, Trace, Lower (Upper) triangular matrix, Unit lower (upper) triangular matrix, Diagonal matrix, Symmetric matrix, Skew-symmetric matrix, Matrix function, Column n-vector function.

Skills

- Be able to determine the elements of a matrix.
- Be able to identify the size (i.e., dimensions) of a matrix.
- Be able to identify the row and column vectors of a matrix.
- Be able to determine the components of a row or column vector.
- Be able to say whether or not two given matrices are equal.
- Be able to find the transpose of a matrix.
- Be able to compute the trace of a square matrix.
- Be able to recognize square matrices that are upper triangular, lower triangular, or diagonal.
- Be able to recognize square matrices that are symmetric or skew-symmetric.
- Be able to determine the values of the variable t such that a matrix function A is defined.

True-False Review

For Questions 1–10, decide if the given statement is **true** or **false**, and give a brief justification for your answer. If true, you can quote a relevant definition or theorem from the text. If false, provide an example, illustration, or brief explanation of why the statement is false.

1. A diagonal matrix must be both upper triangular and lower triangular.

2. An $m \times n$ matrix has m column vectors and n row vectors.

3. If A is a symmetric matrix, then so is A^T.

4. The trace of a matrix is the product of the elements along the main diagonal.

5. A skew-symmetric matrix must have zeros along the main diagonal.

6. A matrix that is both symmetric and skew-symmetric cannot contain any nonzero elements.

7. The matrix functions

$$\begin{bmatrix} \sqrt{t} & 3t^2 \\ \frac{1}{|t|} & \sin 2t \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} -2+t & \ln t \\ e^{\sin t} & -3 \end{bmatrix}$$

are defined for exactly the same values of t.

[2] We could, of course, also speak of **row n-vector functions** as the $1 \times n$ matrix functions, but we will not need them in this text.

8. The matrix function

$$\begin{bmatrix} \cos t & t^2 \\ -2 & -t \\ e^t & \dfrac{1}{\sqrt{t-3}} \end{bmatrix}$$

is defined for all positive real numbers t.

9. Any matrix of numbers is a matrix function defined for all real values of the variable t.

10. If A and B are matrix functions such that the matrices $A(0)$ and $B(0)$ are the same, then we should consider A and B to be the same matrix function.

Problems

1. If

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

determine a_{31}, a_{24}, a_{14}, a_{32}, a_{21}, and a_{34}.

For Problems 2–6, write the matrix with the given elements. In each case, specify the dimensions of the matrix.

2. $a_{11} = 1, a_{21} = -1, a_{12} = 5, a_{22} = 3.$

3. $a_{11} = 2, a_{12} = 1, a_{13} = -1, a_{21} = 0, a_{22} = 4, a_{23} = -2.$

4. $a_{11} = -1, a_{41} = -5, a_{31} = 1, a_{21} = 1.$

5. $a_{11} = 1, a_{31} = 2, a_{42} = -1, a_{32} = 7, a_{13} = -2, a_{23} = 0, a_{33} = 4, a_{21} = 3, a_{41} = -4, a_{12} = -3, a_{22} = 6, a_{43} = 5.$

6. $a_{12} = -1, a_{13} = 2, a_{23} = 3, a_{ji} = -a_{ij}$, $1 \le i \le 3, 1 \le j \le 3.$

For Problems 7–9, determine tr(A) for the given matrix.

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

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

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

For Problems 10–12, write the column vectors and row vectors of the given matrix.

10. $A = \begin{bmatrix} 1 & -1 \\ 3 & 5 \end{bmatrix}.$

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

12. $A = \begin{bmatrix} 2 & 10 & 6 \\ 5 & -1 & 3 \end{bmatrix}.$

13. If $\mathbf{a}_1 = [1 \quad 2]$, $\mathbf{a}_2 = [3 \quad 4]$, and $\mathbf{a}_3 = [5 \quad 1]$, write the matrix

$$A = \begin{bmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \mathbf{a}_3 \end{bmatrix},$$

and determine the column vectors of A.

14. If

$$\mathbf{b}_1 = \begin{bmatrix} 2 \\ -1 \\ 4 \end{bmatrix}, \quad \mathbf{b}_2 = \begin{bmatrix} 5 \\ 7 \\ -6 \end{bmatrix},$$

$$\mathbf{b}_3 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{b}_4 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix},$$

write the matrix $B = [\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3, \mathbf{b}_4]$ and determine the row vectors of B.

15. If $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_p$ are each column q-vectors, what are the dimensions of the matrix that has $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_p$ as its column vectors?

For Problems 16–20, give an example of a matrix of the specified form.

16. 3×3 diagonal matrix.

17. 4×4 upper triangular matrix.

18. 4×4 skew-symmetric matrix.

19. 3×3 upper triangular symmetric matrix.

20. 3×3 lower triangular skew-symmetric matrix.

For Problems 21– 24, give an example of a matrix function of the specified form.

21. 2×3 matrix function defined only for values of t with $-2 \le t < 3$.

22. 4×2 matrix function A such that

$$A(0) = A(1) \neq A(2).$$

23. 1×5 matrix function A that is nonconstant such that all elements of $A(t)$ are positive for all t in $\mathbb{R}$.

24. 2×1 matrix function A that is nonconstant such that all elements of $A(t)$ are in $[0, 1]$ for every t in $\mathbb{R}$.

25. Construct *distinct* matrix functions A and B defined on all of $\mathbb{R}$ such that $A(0) = B(0)$ and $A(1) = B(1)$.

26. Prove that a symmetric upper triangular matrix is diagonal.

27. Determine all elements of the 3×3 skew-symmetric matrix A with $a_{21} = 1, a_{31} = 3, a_{23} = -1$.

2.2 Matrix Algebra

In the previous section we introduced the general idea of a matrix. The next step is to develop the algebra of matrices. Unless otherwise stated, we assume that all elements of the matrices that appear are real or complex numbers.

Addition and Subtraction of Matrices and Multiplication of a Matrix by a Scalar

Addition and subtraction of matrices is defined only for matrices with the same dimensions. We begin with addition.

DEFINITION 2.2.1

If A and B are both $m \times n$ matrices, then we define **addition** (or the **sum**) of A and B, denoted by $A + B$, to be the $m \times n$ matrix whose elements are obtained by adding corresponding elements of A and B. In index notation, if $A = [a_{ij}]$ and $B = [b_{ij}]$, then $A + B = [a_{ij} + b_{ij}]$.

Example 2.2.2 We have

$$\begin{bmatrix} 2 & -1 & 3 \\ 4 & -5 & 0 \end{bmatrix} + \begin{bmatrix} -1 & 0 & 5 \\ -5 & 2 & 7 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 8 \\ -1 & -3 & 7 \end{bmatrix}. \qquad \square$$

Properties of Matrix Addition: If A and B are both $m \times n$ matrices, then

$$A + B = B + A \qquad \text{(matrix addition is commutative)},$$
$$A + (B + C) = (A + B) + C \qquad \text{(matrix addition is associative)}.$$

Both of these properties follow directly from Definition 2.2.1.

In order that we can model oscillatory physical phenomena, in much of the later work we will need to use complex as well as real numbers. Throughout the text we will use the term **scalar** to mean a real or complex number.

DEFINITION 2.2.3

If A is an $m \times n$ matrix and s is a scalar, then we let sA denote the matrix obtained by multiplying every element of A by s. This procedure is called **scalar multiplication**. In index notation, if $A = [a_{ij}]$, then $sA = [sa_{ij}]$.

Example 2.2.4 If $A = \begin{bmatrix} 2 & -1 \\ 4 & 6 \end{bmatrix}$, then $5A = \begin{bmatrix} 10 & -5 \\ 20 & 30 \end{bmatrix}$. $\square$