Basic Linear Algebra

In this chapter, we study the topics in linear algebra that will be needed in the rest of the book. We begin by discussing the building blocks of linear algebra: matrices and vectors. Then we use our knowledge of matrices and vectors to develop a systematic procedure (the Gaussian-Jordan method) for solving linear equations, which we then use to invert matrices. We close the chapter with an introduction to determinants.

The material covered in this chapter will be used in our study of linear and nonlinear programming.

2.1 Matrices and Vectors

Matrices

**Definition** A matrix is any rectangular array of numbers.

For example,

\[
\begin{bmatrix}
1 & 2 \\
3 & 4
\end{bmatrix},
\begin{bmatrix}
1 & 2 & 3 \\
4 & 5 & 6
\end{bmatrix},
\begin{bmatrix}
1 \\
-2
\end{bmatrix},
\begin{bmatrix}
2 & 1
\end{bmatrix}
\]

are all matrices.

If a matrix \( A \) has \( m \) rows and \( n \) columns, we call \( A \) an \( m \times n \) matrix. We refer to \( m \times n \) as the order of the matrix. A typical \( m \times n \) matrix \( A \) may be written as

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

**Definition** The number in the \( i \)th row and \( j \)th column of \( A \) is called the \( ij \)th element of \( A \) and is written \( a_{ij} \).

For example, if

\[
A = \begin{bmatrix}
1 & 2 & 3 \\
4 & 5 & 6 \\
7 & 8 & 9
\end{bmatrix}
\]

then \( a_{11} = 1, a_{23} = 6, \) and \( a_{31} = 7. \)
Sometimes we will use the notation $A = [a_{ij}]$ to indicate that $A$ is the matrix whose $ij$th element is $a_{ij}$.

**Definition**

Two matrices $A = [a_{ij}]$ and $B = [b_{ij}]$ are equal if and only if $A$ and $B$ are of the same order and for all $i$ and $j$, $a_{ij} = b_{ij}$.

For example, if

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

then $A = B$ if and only if $x = 1$, $y = 2$, $w = 3$, and $z = 4$.

**Vectors**

Any matrix with only one column (that is, any $m \times 1$ matrix) may be thought of as a **column vector**. The number of rows in a column vector is the **dimension** of the column vector. Thus,

$$\begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

may be thought of as a $2 \times 1$ matrix or a two-dimensional column vector. $R^m$ will denote the set of all $m$-dimensional column vectors.

In analogous fashion, we can think of any vector with only one row (a $1 \times n$ matrix as a **row vector**. The dimension of a row vector is the number of columns in the vector. Thus, $[9 \quad 2 \quad 3]$ may be viewed as a $1 \times 3$ matrix or a three-dimensional row vector. In this book, vectors appear in boldface type: for instance, vector $v$. An $m$-dimensional vector (either row or column) in which all elements equal zero is called a **zero vector** (written $0$). Thus,

$$\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

are two-dimensional zero vectors.

Any $m$-dimensional vector corresponds to a directed line segment in the $m$-dimensional plane. For example, in the two-dimensional plane, the vector

$$u = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

corresponds to the line segment joining the point

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix}$$
to the point

$$\begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

The directed line segments corresponding to

$$u = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad v = \begin{bmatrix} 1 \\ -3 \end{bmatrix}, \quad w = \begin{bmatrix} -1 \\ -2 \end{bmatrix}$$

are drawn in Figure 1.
**The Scalar Product of Two Vectors**

An important result of multiplying two vectors is the *scalar product*. To define the scalar product of two vectors, suppose we have a row vector \( \mathbf{u} = [u_1 \ u_2 \ \cdots \ u_n] \) and a column vector

\[
\mathbf{v} = \begin{bmatrix}
    v_1 \\
    v_2 \\
    \vdots \\
    v_n
\end{bmatrix}
\]

of the same dimension. The *scalar product* of \( \mathbf{u} \) and \( \mathbf{v} \) (written \( \mathbf{u} \cdot \mathbf{v} \)) is the number

\[ u_1v_1 + u_2v_2 + \cdots + u_nv_n. \]

For the scalar product of two vectors to be defined, the first vector must be a row vector and the second vector must be a column vector. For example, if

\[
\mathbf{u} = [1 \ 2 \ 3] \quad \text{and} \quad \mathbf{v} = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}
\]

then \( \mathbf{u} \cdot \mathbf{v} = 1(2) + 2(1) + 3(2) = 10. \) By these rules for computing a scalar product, if

\[
\mathbf{u} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad \text{and} \quad \mathbf{v} = [2 \ 3]
\]

then \( \mathbf{u} \cdot \mathbf{v} \) is not defined. Also, if

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

then \( \mathbf{u} \cdot \mathbf{v} \) is not defined because the vectors are of two different dimensions.

Note that two vectors are perpendicular if and only if their scalar product equals 0. Thus, the vectors \([1 \ -1]\) and \([1 \ 1]\) are perpendicular.

We note that \( \mathbf{u} \cdot \mathbf{v} = \|\mathbf{u}\| \|\mathbf{v}\| \cos \theta \), where \( \|\mathbf{u}\| \) is the length of the vector \( \mathbf{u} \) and \( \theta \) is the angle between the vectors \( \mathbf{u} \) and \( \mathbf{v} \).
Matrix Operations

We now describe the arithmetic operations on matrices that are used later in this book.

The Scalar Multiple of a Matrix

Given any matrix $A$ and any number $c$ (a number is sometimes referred to as a scalar), the matrix $cA$ is obtained from the matrix $A$ by multiplying each element of $A$ by $c$. For example,

$$A = \begin{bmatrix} 1 & 2 \\ -1 & 0 \end{bmatrix}, \quad \text{then} \quad 3A = \begin{bmatrix} 3 & 6 \\ -3 & 0 \end{bmatrix}$$

For $c = -1$, scalar multiplication of the matrix $A$ is sometimes written as $-A$.

Addition of Two Matrices

Let $A = [a_{ij}]$ and $B = [b_{ij}]$ be two matrices with the same order (say, $m \times n$). Then the matrix $C = A + B$ is defined to be the $m \times n$ matrix whose $ij$th element is $a_{ij} + b_{ij}$. Thus, to obtain the sum of two matrices $A$ and $B$, we add the corresponding elements of $A$ and $B$. For example, if

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

then

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

This rule for matrix addition may be used to add vectors of the same dimension. For example, if $u = [1 \ 2]$ and $v = [2 \ 1]$, then $u + v = [1 + 2 \ 2 + 1] = [3 \ 3]$. Vectors may be added geometrically by the parallelogram law (see Figure 2).

We can use scalar multiplication and the addition of matrices to define the concept of a line segment. A glance at Figure 1 should convince you that any point $u$ in the $m$-dimensional plane corresponds to the $m$-dimensional vector $u$ formed by joining the origin to the point $u$. For any two points $u$ and $v$ in the $m$-dimensional plane, the line segment joining $u$ and $v$ (called the line segment $uv$) is the set of all points in the $m$-dimensional plane that correspond to the vectors $cu + (1 - c)v$, where $0 \leq c \leq 1$ (Figure 3). For example, if $u = (1, 2)$ and $v = (2, 1)$, then the line segment $uv$ consists
of the points corresponding to the vectors \(c[1 \ 2] + (1 - c)[2 \ 1] = [2 - c \ 1 + c]\),
where \(0 \leq c \leq 1\). For \(c = 0\) and \(c = 1\), we obtain the endpoints of the line segment \(uv\);
for \(c = \frac{1}{2}\), we obtain the midpoint \((0.5u + 0.5v)\) of the line segment \(uv\).

Using the parallelogram law, the line segment \(uv\) may also be viewed as the points
 corresponding to the vectors \(u + c(v - u)\), where \(0 \leq c \leq 1\) (Figure 4). Observe that for \(c = 0\),
we obtain the vector \(u\) (corresponding to point \(u\)), and for \(c = 1\), we obtain the
vector \(v\) (corresponding to point \(v\)).

**The Transpose of a Matrix**

Given any \(m \times n\) matrix

\[
A = \begin{bmatrix}
a_{11} & a_{12} & \ldots & a_{1n} \\
a_{21} & a_{22} & \ldots & a_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
a_{m1} & a_{m2} & \ldots & a_{mn}
\end{bmatrix}
\]

the transpose of \(A\) (written \(A^T\)) is the \(n \times m\) matrix

\[
A^T = \begin{bmatrix}
a_{11} & a_{21} & \ldots & a_{m1} \\
a_{12} & a_{22} & \ldots & a_{m2} \\
\vdots & \vdots & \ddots & \vdots \\
a_{1n} & a_{2n} & \ldots & a_{mn}
\end{bmatrix}
\]
Thus, $A^T$ is obtained from $A$ by letting row 1 of $A$ be column 1 of $A^T$, letting row 2 of $A$ be column 2 of $A^T$, and so on. For example,

$$
\text{if } A = \begin{bmatrix}
1 & 2 & 3 \\
4 & 5 & 6
\end{bmatrix}, \text{ then } \quad A^T = \begin{bmatrix}
1 & 2 \\
4 & 5 \\
3 & 6
\end{bmatrix}
$$

Observe that $(A^T)^T = A$. Let $B = [1 \quad 2]$; then

$$
B^T = \begin{bmatrix}
1 \\
2
\end{bmatrix} \quad \text{and} \quad (B^T)^T = \begin{bmatrix}
1 & 2
\end{bmatrix} = B
$$

As indicated by these two examples, for any matrix $A$, $(A^T)^T = A$.

**Matrix Multiplication**

Given two matrices $A$ and $B$, the matrix product of $A$ and $B$ (written $AB$) is defined if and only if

$$
\text{Number of columns in } A = \text{number of rows in } B
$$

For the moment, assume that for some positive integer $r$, $A$ has $r$ columns and $B$ has $r$ rows. Then for some $m$ and $n$, $A$ is an $m \times r$ matrix and $B$ is an $r \times n$ matrix.

**Definition**

The **matrix product** $C = AB$ of $A$ and $B$ is the $m \times n$ matrix $C$ whose $ij$th element is determined as follows:

$$
c_{ij} = \text{scalar product of row } i \text{ of } A \times \text{column } j \text{ of } B
$$

If Equation (1) is satisfied, then each row of $A$ and each column of $B$ will have the same number of elements. Also, if (1) is satisfied, then the scalar product in Equation (2) will be defined. The product matrix $C = AB$ will have the same number of rows as $A$ and the same number of columns as $B$.

**Example 1** Matrix Multiplication

Compute $C = AB$ for

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

**Solution** Because $A$ is a $2 \times 3$ matrix and $B$ is a $3 \times 2$ matrix, $AB$ is defined, and $C$ will be a $2 \times 2$ matrix. From Equation (2),

$$
c_{11} = [1 & 1 & 2] \begin{bmatrix}
1 \\
2 \\
1
\end{bmatrix} = 1(1) + 1(2) + 2(1) = 5
$$

$$
c_{12} = [1 & 1 & 2] \begin{bmatrix}
1 \\
3 \\
2
\end{bmatrix} = 1(1) + 1(3) + 2(2) = 8
$$

$$
c_{21} = [2 & 1 & 3] \begin{bmatrix}
1 \\
1
\end{bmatrix} = 2(1) + 1(2) + 3(1) = 7
$$

16

Chapter 2 Basic Linear Algebra
\[ c_{22} = \begin{bmatrix} 2 & 1 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \end{bmatrix} = 2(1) + 1(3) + 3(2) = 11 \]

\[ C = AB = \begin{bmatrix} 5 & 8 \\ 7 & 11 \end{bmatrix} \]

**Example 2: Column Vector Times Row Vector**

Find \( AB \) for

\[ A = \begin{bmatrix} 3 \\ 4 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 2 \end{bmatrix} \]

**Solution**

Because \( A \) has one column and \( B \) has one row, \( C = AB \) will exist. From Equation (2), we know that \( C \) is a \( 2 \times 2 \) matrix with

\[ c_{11} = 3(1) = 3 \quad c_{21} = 4(1) = 4 \]
\[ c_{12} = 3(2) = 6 \quad c_{22} = 4(2) = 8 \]

Thus,

\[ C = \begin{bmatrix} 3 & 6 \\ 4 & 8 \end{bmatrix} \]

**Example 3: Row Vector Times Column Vector**

Compute \( D = BA \) for the \( A \) and \( B \) of Example 2.

**Solution**

In this case, \( D \) will be a \( 1 \times 1 \) matrix (or a scalar). From Equation (2),

\[ d_{11} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 3 \\ 4 \end{bmatrix} = 1(3) + 2(4) = 11 \]

Thus, \( D = 11 \). In this example, matrix multiplication is equivalent to scalar multiplication of a row and column vector.

Recall that if you multiply two real numbers \( a \) and \( b \), then \( ab = ba \). This is called the **commutative property of multiplication**. Examples 2 and 3 show that for matrix multiplication it may be that \( AB \neq BA \). Matrix multiplication is not necessarily commutative. (In some cases, however, \( AB = BA \) will hold.)

**Example 4: Undefined Matrix Product**

Show that \( AB \) is undefined if

\[ A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 1 \end{bmatrix} \]

**Solution**

This follows because \( A \) has two columns and \( B \) has three rows. Thus, Equation (1) is not satisfied.
TABLE 1
Balloons of Crude Oil Required to Produce 1 Gallon
of Gasoline

<table>
<thead>
<tr>
<th>Crude Oil</th>
<th>Premium Unleaded</th>
<th>Regular Unleaded</th>
<th>Regular Lead</th>
<th>Created</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$\frac{3}{4}$</td>
<td>$\frac{1}{3}$</td>
<td>$\frac{1}{4}$</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>$\frac{1}{2}$</td>
<td>$\frac{1}{3}$</td>
<td>$\frac{1}{2}$</td>
<td></td>
</tr>
</tbody>
</table>

Many computations that commonly occur in operations research (and other branches of mathematics) can be concisely expressed by using matrix multiplication. To illustrate this, suppose an oil company manufactures three types of gasoline: premium unleaded, regular unleaded, and regular leaded. These gasolines are produced by mixing two types of crude oil: crude oil 1 and crude oil 2. The number of gallons of crude oil required to manufacture 1 gallon of gasoline is given in Table 1.

From this information, we can find the amount of each type of crude oil needed to manufacture a given amount of gasoline. For example, if the company wants to produce 10 gallons of premium unleaded, 6 gallons of regular unleaded, and 5 gallons of regular leaded, then the company’s crude oil requirements would be

\[
\text{Crude 1 required} = \left(\frac{3}{4}\right) (10) + \left(\frac{1}{3}\right) (6) + \left(\frac{1}{4}\right) (5) = 12.75 \text{ gallons}
\]
\[
\text{Crude 2 required} = \left(\frac{1}{2}\right) (10) + \left(\frac{1}{3}\right) (6) + \left(\frac{1}{2}\right) (5) = 8.25 \text{ gallons}
\]

More generally, we define

\[ p_U = \text{gallons of premium unleaded produced} \]
\[ r_U = \text{gallons of regular unleaded produced} \]
\[ r_L = \text{gallons of regular leaded produced} \]
\[ c_1 = \text{gallons of crude 1 required} \]
\[ c_2 = \text{gallons of crude 2 required} \]

Then the relationship between these variables may be expressed by

\[ c_1 = \left(\frac{3}{4}\right) p_U + \left(\frac{1}{3}\right) r_U + \left(\frac{1}{4}\right) r_L \]
\[ c_2 = \left(\frac{1}{2}\right) p_U + \left(\frac{1}{3}\right) r_U + \left(\frac{1}{2}\right) r_L \]

Using matrix multiplication, these relationships may be expressed by

\[
\begin{bmatrix}
  c_1 \\
  c_2
\end{bmatrix} =
\begin{bmatrix}
  \frac{3}{4} & \frac{1}{3} & \frac{1}{4} \\
  \frac{1}{2} & \frac{1}{3} & \frac{1}{2}
\end{bmatrix}
\begin{bmatrix}
  p_U \\
  r_U \\
  r_L
\end{bmatrix}
\]

Properties of Matrix Multiplication

To close this section, we discuss some important properties of matrix multiplication. In what follows, we assume that all matrix products are defined.

1. Row i of \(AB\) = (row i of \(A\))B. To illustrate this property, let

\[
A = \begin{bmatrix}
  1 & 1 & 2 \\
  2 & 1 & 3
\end{bmatrix} \text{ and } B = \begin{bmatrix}
  1 & 1 \\
  2 & 3 \\
  1 & 2
\end{bmatrix}
\]

Then row 2 of the 2 \(\times\) 2 matrix \(AB\) is equal to

\[
\begin{bmatrix}
  2 & 1 \\
  3 & 2
\end{bmatrix}
\]
This answer agrees with Example 1.

2 Column $j$ of $AB = A$ (column $j$ of $B$). Thus, for $A$ and $B$ as given, the first column of $AB$ is

$$
\begin{bmatrix}
1 & 1 \\
2 & 3 \\
1 & 2
\end{bmatrix}
\begin{bmatrix}
1 & 2 \\
2 & 3 \\
1 & 2
\end{bmatrix}
= \begin{bmatrix}
7 & 11
\end{bmatrix}
$$

Properties 1 and 2 are helpful when you need to compute only part of the matrix $AB$.

3 Matrix multiplication is associative. That is, $A(BC) = (AB)C$. To illustrate, let

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

Then $AB = \begin{bmatrix} 10 & 13 \end{bmatrix}$ and $(AB)C = 10(2) + 13(1) = 33$.

On the other hand,

$$BC = \begin{bmatrix} 7 \\ 13 \end{bmatrix}$$

so $A(BC) = 1(7) + 2(13) = 33$. In this case, $A(BC) = (AB)C$ does hold.

4 Matrix multiplication is distributive. That is, $A(B + C) = AB + AC$ and $(B + C)D = BD + CD$.

**Matrix Multiplication with EXCEL**

Using the EXCEL MMULT function, it is easy to multiply matrices. To illustrate, let's use EXCEL to find the matrix product $AB$ that we found in Example 1 (see Figure 5 and file Mmult.xls). We proceed as follows:

**Step 1** Enter $A$ and $B$ in D2:F3 and D5:E7, respectively.

**Step 2** Select the range (D9:E10) in which the product $AB$ will be computed.

**Step 3** In the upper left-hand corner (D9) of the selected range, type the formula

= MMULT(D2:F3,D5:E7).

Then hit **Control Shift Enter** (not just enter), and the desired matrix product will be computed. Note that MMULT is an **array** function and not an ordinary spreadsheet function. This explains why we must preselect the range for $AB$ and use Control Shift Enter.
PROBLEMS

Group A

1 For \( A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \) and \( B = \begin{bmatrix} 1 & 2 \\ 0 & -1 \\ 1 & 2 \end{bmatrix} \), find:
   \( a \ - A \ b \ 3A \ c \ A + 2B \ d \ A^T \ e \ B^T \ f \ AB \ g \ BA \)

2 Only three brands of beer (beer 1, beer 2, and beer 3) are available for sale in Metropolis. From time to time, people try one or another of these brands. Suppose that at the beginning of each month, people change the beer they are drinking according to the following rules:
   30\% of the people who prefer beer 1 switch to beer 2.
   20\% of the people who prefer beer 1 switch to beer 3.
   30\% of the people who prefer beer 2 switch to beer 3.
   10\% of the people who prefer beer 3 switch to beer 1.

   For \( i = 1, 2, 3 \), let \( x_i \) be the number who prefer beer \( i \) at the beginning of this month and \( y_i \) be the number who prefer beer \( i \) at the beginning of next month. Use matrix multiplication to relate the following:

\[
\begin{bmatrix}
  y_1 \\
  y_2 \\
  y_3
\end{bmatrix} = \begin{bmatrix}
  x_1 \\
  x_2 \\
  x_3
\end{bmatrix}
\]

Group B

3 Prove that matrix multiplication is associative.

4 Show that for any two matrices \( A \) and \( B \), \((AB)^T = B^T A^T\).

5 An \( n \times n \) matrix \( A \) is symmetric if \( A = A^T \).
   \( a \) Show that for any \( n \times n \) matrix, \( AA^T \) is a symmetric matrix.
   \( b \) Show that for any \( n \times n \) matrix \( A \), \((A + A^T) \) is a symmetric matrix.

6 Suppose that \( A \) and \( B \) are both \( n \times n \) matrices. Show that computing the matrix product \( AB \) requires \( n^3 \) multiplications and \( n^3 - n^2 \) additions.

7 The trace of a matrix is the sum of its diagonal elements.
   \( a \) For any two matrices \( A \) and \( B \), show that trace \((A + B) = \text{trace } A + \text{trace } B \).
   \( b \) For any two matrices \( A \) and \( B \) for which the products \( AB \) and \( BA \) are defined, show that \( \text{trace } AB = \text{trace } BA \).

2.2 Matrices and Systems of Linear Equations

Consider a system of linear equations given by

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

(3)

In Equation (3), \( x_1, x_2, \ldots, x_n \) are referred to as variables, or unknowns, and the \( a_{ij} \)'s and \( b_i \)'s are constants. A set of equations such as (3) is called a linear system of \( m \) equations in \( n \) variables.

**DEFINITION** A solution to a linear system of \( m \) equations in \( n \) unknowns is a set of values for the unknowns that satisfies each of the system's \( m \) equations.

To understand linear programming, we need to know a great deal about the properties of solutions to linear equation systems. With this in mind, we will devote much effort to studying such systems.

We denote a possible solution to Equation (3) by an \( n \)-dimensional column vector \( \mathbf{x} \), in which the \( i \)th element of \( \mathbf{x} \) is the value of \( x_i \). The following example illustrates the concept of a solution to a linear system.

20
Solution to Linear System

Show that

\[ x = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \]

is a solution to the linear system

\[
\begin{align*}
x_1 + 2x_2 &= 5 \\
2x_1 - x_2 &= 0
\end{align*}
\] (4)

and that

\[ x = \begin{bmatrix} 3 \\ 1 \end{bmatrix} \]

is not a solution to linear system (4).

**Solution**

To show that

\[ x = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \]

is a solution to Equation (4), we substitute \( x_1 = 1 \) and \( x_2 = 2 \) in both equations and check that they are satisfied: \( 1 + 2(2) = 5 \) and \( 2(1) - 2 = 0 \).

The vector

\[ x = \begin{bmatrix} 3 \\ 1 \end{bmatrix} \]

is not a solution to (4), because \( x_1 = 3 \) and \( x_2 = 1 \) fail to satisfy \( 2x_1 - x_2 = 0 \).

Using matrices can greatly simplify the statement and solution of a system of linear equations. To show how matrices can be used to compactly represent Equation (3), let

\[
A = \begin{bmatrix}
a_{11} & a_{12} & \cdots & a_{1n} \\
a_{21} & a_{22} & \cdots & a_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
a_{m1} & a_{m2} & \cdots & a_{mn}
\end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}
\]

Then (3) may be written as

\[ Ax = b \] (5)

Observe that both sides of Equation (5) will be \( m \times 1 \) matrices (or \( m \times 1 \) column vectors). For the matrix \( Ax \) to equal the matrix \( b \) (or for the vector \( Ax \) to equal the vector \( b \)), their corresponding elements must be equal. The first element of \( Ax \) is the scalar product of row 1 of \( A \) with \( x \). This may be written as

\[
\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n
\]

This must equal the first element of \( b \) (which is \( b_1 \)). Thus, (5) implies that \( a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \). This is the first equation of (3). Similarly, (5) implies that the scalar
product of row $i$ of $A$ with $x$ must equal $b_i$, and this is just the $i$th equation of (3). Our discussion shows that (3) and (5) are two different ways of writing the same linear system. We call (5) the **matrix representation** of (3). For example, the matrix representation of (4) is

$$
\begin{bmatrix}
1 & 2 \\
2 & -1
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2
\end{bmatrix} =
\begin{bmatrix}
5 \\
0
\end{bmatrix}
$$

Sometimes we abbreviate (5) by writing

$$A|b$$  \hspace{1cm} (6)

If $A$ is an $m \times n$ matrix, it is assumed that the variables in (6) are $x_1, x_2, \ldots, x_n$. Then (6) is still another representation of (3). For instance, the matrix

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

represents the system of equations

$$
x_1 + 2x_2 + 3x_3 = 2 \\
x_2 + 2x_3 = 3 \\
x_1 + x_2 + x_3 = 1
$$

**PROBLEM**

**Group A**

1. Use matrices to represent the following system of equations in two different ways:

$$
x_1 - x_3 = 4 \\
2x_1 + x_2 = 6 \\
x_1 + 3x_2 = 8
$$

2.3 The Gauss–Jordan Method for Solving Systems of Linear Equations

We develop in this section an efficient method (the Gauss–Jordan method) for solving a system of linear equations. Using the Gauss–Jordan method, we show that any system of linear equations must satisfy one of the following three cases:

**Case 1** The system has no solution.

**Case 2** The system has a unique solution.

**Case 3** The system has an infinite number of solutions.

The Gauss–Jordan method is also important because many of the manipulations used in this method are used when solving linear programming problems by the simplex algorithm (see Chapter 4).

**Elementary Row Operations**

Before studying the Gauss–Jordan method, we need to define the concept of an **elementary row operation** (ERO). An ERO transforms a given matrix $A$ into a new matrix $A'$ via one of the following operations.
Type 1 ERO

A' is obtained by multiplying any row of A by a nonzero scalar. For example, if

\[
A = \begin{bmatrix}
1 & 2 & 3 & 4 \\
1 & 3 & 5 & 6 \\
0 & 1 & 2 & 3
\end{bmatrix}
\]

then a Type 1 ERO that multiplies row 2 of A by 3 would yield

\[
A' = \begin{bmatrix}
1 & 2 & 3 & 4 \\
3 & 9 & 15 & 18 \\
0 & 1 & 2 & 3
\end{bmatrix}
\]

Type 2 ERO

Begin by multiplying any row of A (say, row i) by a nonzero scalar c. For some \(j \neq i\), let row j of \(A' = c(\text{row } i \text{ of } A) + \text{row } j \text{ of } A\), and let the other rows of \(A'\) be the same as the rows of \(A\).

For example, we might multiply row 2 of A by 4 and replace row 3 of A by 4(row 2 of A) + row 3 of A. Then row 3 of \(A'\) becomes

\[
4 \begin{bmatrix} 1 & 3 & 5 & 6 \end{bmatrix} + \begin{bmatrix} 0 & 1 & 2 & 3 \end{bmatrix} = \begin{bmatrix} 4 & 13 & 22 & 27 \end{bmatrix}
\]

and

\[
A' = \begin{bmatrix}
1 & 2 & 3 & 4 \\
1 & 3 & 5 & 6 \\
4 & 13 & 22 & 27
\end{bmatrix}
\]

Type 3 ERO

Interchange any two rows of A. For instance, if we interchange rows 1 and 3 of A, we obtain

\[
A' = \begin{bmatrix}
0 & 1 & 2 & 3 \\
1 & 3 & 5 & 6 \\
1 & 2 & 3 & 4
\end{bmatrix}
\]

Type 1 and Type 2 EROs formalize the operations used to solve a linear equation system. To solve the system of equations

\[
x_1 + x_2 = 2 \\
2x_1 + 4x_2 = 7
\]

we might proceed as follows. First replace the second equation in (7) by \(-2(\text{first equation in (7)}) + \text{second equation in (7)}\). This yields the following linear system:

\[
x_1 + x_2 = 2 \\
2x_2 = 3 \tag{7.1}
\]

Then multiply the second equation in (7.1) by \(\frac{1}{2}\), yielding the system

\[
x_1 + x_2 = 2 \\
x_2 = \frac{3}{2} \tag{7.2}
\]

Finally, replace the first equation in (7.2) by \(-1[\text{second equation in (7.2)}] + \text{first equation in (7.2)}\). This yields the system:

\[
2.3 \text{ The Gauss-Jordan Method for Solving Systems of Linear Equations} \quad 23
\]
\[ x_1 = \frac{1}{2} \]
\[ x_2 = \frac{3}{2} \]  

(7.3)

System (7.3) has the unique solution \( x_1 = \frac{1}{2} \) and \( x_2 = \frac{3}{2} \). The systems (7), (7.1), (7.2), and (7.3) are equivalent in that they have the same set of solutions. This means that \( x_1 = \frac{1}{2} \) and \( x_2 = \frac{3}{2} \) is also the unique solution to the original system, (7).

If we view (7) in the augmented matrix form \( [A|b] \), we see that the steps used to solve (7) may be seen as Type 1 and Type 2 EROs applied to \( [A|b] \). Begin with the augmented matrix version of (7):

\[
\begin{bmatrix}
1 & 1 & 2 \\
2 & 4 & 7
\end{bmatrix}
\]  

(7′)

Now perform a Type 2 ERO by replacing row 2 of (7′) by \(-2\) (row 1 of (7′)) + row 2 of (7′). The result is

\[
\begin{bmatrix}
1 & 1 & 2 \\
0 & 2 & 3
\end{bmatrix}
\]  

(7.1′)

which corresponds to (7.1). Next, we multiply row 2 of (7.1′) by \(\frac{1}{2}\) (a Type 1 ERO), resulting in

\[
\begin{bmatrix}
1 & 1 & 2 \\
0 & 1 & \frac{3}{2}
\end{bmatrix}
\]  

(7.2′)

which corresponds to (7.2). Finally, perform a Type 2 ERO by replacing row 1 of (7.2′) by \(-1\) (row 2 of (7.2′)) + row 1 of (7.2′). The result is

\[
\begin{bmatrix}
1 & 0 & \frac{1}{2} \\
0 & 1 & \frac{3}{2}
\end{bmatrix}
\]  

(7.3′)

which corresponds to (7.3). Translating (7.3′) back into a linear system, we obtain the system \( x_1 = \frac{1}{2} \) and \( x_2 = \frac{3}{2} \), which is identical to (7.3).

**Finding a Solution by the Gauss–Jordan Method**

The discussion in the previous section indicates that if the matrix \( A'|b' \) is obtained from \( [A|b] \) via an ERO, the systems \( Ax = b \) and \( A'x = b' \) are equivalent. Thus, any sequence of EROs performed on the augmented matrix \( [A|b] \) corresponding to the system \( Ax = b \) will yield an equivalent linear system.

The Gauss–Jordan method solves a linear equation system by utilizing EROs in a systematic fashion. We illustrate the method by finding the solution to the following linear system:

\[
\begin{align*}
2x_1 + 2x_2 + x_3 &= 9 \\
2x_1 - x_2 + 2x_3 &= 6 \\
x_1 - x_2 + 2x_3 &= 5
\end{align*}
\]

(8)

The augmented matrix representation is

\[
[A|b] = \begin{bmatrix}
2 & 2 & 1 & 9 \\
2 & -1 & 2 & 6 \\
1 & -1 & 2 & 5
\end{bmatrix}
\]

(8′)

Suppose that by performing a sequence of EROs on (8′) we could transform (8′) into...
\[
\begin{bmatrix}
1 & 0 & 0 & 1 \\
0 & 1 & 0 & 2 \\
0 & 0 & 1 & 3
\end{bmatrix}
\]

We note that the result obtained by performing an ERO on a system of equations can also be obtained by multiplying both sides of the matrix representation of the system of equations by a particular matrix. This explains why EROs do not change the set of solutions to a system of equations.

Matrix \( \begin{bmatrix} \end{bmatrix} \) corresponds to the following linear system:

\[
x_1 = 1 \\
x_2 = 2 \\
x_3 = 3
\]

System (9) has the unique solution \( x_1 = 1, x_2 = 2, x_3 = 3 \). Because (9') was obtained from (8') by a sequence of EROs, we know that (8) and (9) are equivalent linear systems. Thus, \( x_1 = 1, x_2 = 2, x_3 = 3 \) must also be the unique solution to (8). We now show how we can use EROs to transform a relatively complicated system such as (8) into a relatively simple system like (9). This is the essence of the Gauss-Jordan method.

We begin by using EROs to transform the first column of (8') into

\[
\begin{bmatrix}
1 \\
0 \\
0
\end{bmatrix}
\]

Then we use EROs to transform the second column of the resulting matrix into

\[
\begin{bmatrix}
0 \\
1 \\
0
\end{bmatrix}
\]

Finally, we use EROs to transform the third column of the resulting matrix into

\[
\begin{bmatrix}
0 \\
0 \\
1
\end{bmatrix}
\]

As a final result, we will have obtained (9'). We now use the Gauss-Jordan method to solve (8). We begin by using a Type 1 ERO to change the element of (8') in the first row and first column into a 1. Then we add multiples of row 1 to row 2 and then to row 3 (these are Type 2 EROs). The purpose of these Type 2 EROs is to put zeros in the rest of the first column. The following sequence of EROs will accomplish these goals.

**Step 1** Multiply row 1 of (8') by \( \frac{1}{2} \). This Type 1 ERO yields

\[
A_1|b_1 = \begin{bmatrix}
1 & 1 & \frac{1}{2} & \frac{9}{2} \\
2 & -1 & 2 & 6 \\
1 & -1 & 2 & 5
\end{bmatrix}
\]

**Step 2** Replace row 2 of \( A_1|b_1 \) by \(-2(\text{row 1 of } A_1|b_1) + \text{row 2 of } A_1|b_1\). The result of this Type 2 ERO is

\[
A_2|b_2 = \begin{bmatrix}
1 & 1 & \frac{1}{2} & \frac{9}{2} \\
0 & -3 & 1 & -3 \\
1 & -1 & 2 & 5
\end{bmatrix}
\]

2.3 The Gauss-Jordan Method for Solving Systems of Linear Equations

25
Step 3 Replace row 3 of $A_2|b_2$ by $-1$ (row 1 of $A_2|b_2$ + row 3 of $A_2|b_2$). The result of this Type 2 ERO is

$$A_3|b_3 = \begin{bmatrix} 1 & 1 & \frac{1}{2} & 9 \frac{1}{2} \\ 0 & -3 & 1 & -3 \\ 0 & -2 & \frac{1}{2} & 1 \end{bmatrix}$$

The first column of $A_3$ has now been transformed into

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

By our procedure, we have made sure that the variable $x_1$ occurs in only a single equation and in that equation has a coefficient of 1. We now transform the second column of $A_3|b_3$ into

$$\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

We begin by using a Type 1 ERO to create a 1 in row 2 and column 2 of $A_3|b_3$. Then we use the resulting row 2 to perform the Type 2 EROs that are needed to put zeros in the rest of column 2. Steps 4–6 accomplish these goals.

Step 4 Multiply row 2 of $A_3|b_3$ by $\frac{1}{3}$. The result of this Type 1 ERO is

$$A_4|b_4 = \begin{bmatrix} 1 & 1 & \frac{1}{3} & \frac{5}{2} \\ 0 & 1 & -\frac{1}{3} & 1 \\ 0 & -2 & \frac{2}{3} & \frac{1}{2} \end{bmatrix}$$

Step 5 Replace row 1 of $A_4|b_4$ by $-1$ (row 2 of $A_4|b_4$) + row 1 of $A_4|b_4$. The result of this Type 2 ERO is

$$A_5|b_5 = \begin{bmatrix} 1 & 0 & \frac{5}{6} & \frac{7}{2} \\ 0 & 1 & -\frac{1}{3} & 1 \\ 0 & -2 & \frac{2}{3} & \frac{1}{2} \end{bmatrix}$$

Step 6 Replace row 3 of $A_5|b_5$ by 2 (row 2 of $A_5|b_5$) + row 3 of $A_5|b_5$. The result of this Type 2 ERO is

$$A_6|b_6 = \begin{bmatrix} 1 & 0 & \frac{5}{6} & \frac{7}{2} \\ 0 & 1 & -\frac{1}{3} & 1 \\ 0 & 0 & \frac{5}{6} & \frac{5}{2} \end{bmatrix}$$

Column 2 has now been transformed into

$$\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

Observe that our transformation of column 2 did not change column 1.

To complete the Gauss–Jordan procedure, we must transform the third column of $A_6|b_6$ into

$$\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$
We first use a Type 1 ERO to create a 1 in the third row and third column of \( A_6 | b_6 \). Then we use Type 2 EROs to put zeros in the rest of column 3. Steps 7–9 accomplish these goals.

**Step 7** Multiply row 3 of \( A_6 | b_6 \) by \( \frac{6}{5} \). The result of this Type 1 ERO is

\[
A_7 | b_7 = \begin{bmatrix} 1 & 0 & \frac{5}{6} & \frac{7}{5} \\ 0 & 1 & -\frac{1}{3} & 1 \\ 0 & 0 & 1 & 3 \end{bmatrix}
\]

**Step 8** Replace row 1 of \( A_7 | b_7 \) by \( \frac{5}{3} \) (row 3 of \( A_7 | b_7 \)) + row 1 of \( A_7 | b_7 \). The result of this Type 2 ERO is

\[
A_8 | b_8 = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & -\frac{1}{5} & 1 \\ 0 & 0 & 1 & 3 \end{bmatrix}
\]

**Step 9** Replace row 2 of \( A_8 | b_8 \) by \( \frac{1}{3} \) (row 3 of \( A_8 | b_8 \)) + row 2 of \( A_8 | b_8 \). The result of this Type 2 ERO is

\[
A_9 | b_9 = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{bmatrix}
\]

\( A_9 | b_9 \) represents the system of equations

\[
\begin{align*}
x_1 & = 1 \\
x_2 & = 2 \\
x_3 & = 3 
\end{align*}
\]

(9)

Thus, (9) has the unique solution \( x_1 = 1, x_2 = 2, x_3 = 3 \). Because (9) was obtained from (8) via EROs, the unique solution to (8) must also be \( x_1 = 1, x_2 = 2, x_3 = 3 \).

The reader might be wondering why we defined Type 3 EROs (interchanging of rows). To see why a Type 3 ERO might be useful, suppose you want to solve

\[
\begin{align*}
2x_2 + x_3 &= 6 \\
x_1 + x_2 - x_3 &= 2 \\
2x_1 + x_2 + x_3 &= 4
\end{align*}
\]

(10)

To solve (10) by the Gauss–Jordan method, first form the augmented matrix

\[
A | b = \begin{bmatrix} 0 & 2 & 1 & 6 \\ 1 & 1 & -1 & 2 \\ 2 & 1 & 1 & 4 \end{bmatrix}
\]

The 0 in row 1 and column 1 means that a Type 1 ERO cannot be used to create a 1 in row 1 and column 1. If, however, we interchange rows 1 and 2 (a Type 3 ERO), we obtain

\[
\begin{bmatrix} 1 & 1 & -1 & 2 \\ 0 & 2 & 1 & 6 \\ 2 & 1 & 1 & 4 \end{bmatrix}
\]

(10')

Now we may proceed as usual with the Gauss–Jordan method.

2.3 The Gauss–Jordan Method for Solving Systems of Linear Equations
Special Cases: No Solution or an Infinite Number of Solutions

Some linear systems have no solution, and some have an infinite number of solutions. The following two examples illustrate how the Gauss–Jordan method can be used to recognize these cases.

**Example 6. Linear System with No Solution**

Find all solutions to the following linear system:

\[
\begin{align*}
x_1 + 2x_2 &= 3 \\
2x_1 + 4x_2 &= 4
\end{align*}
\]  

\( (11) \)

**Solution** We apply the Gauss–Jordan method to the matrix

\[
A|b = \begin{bmatrix}
1 & 2 & 3 \\
2 & 4 & 4
\end{bmatrix}
\]

We begin by replacing row 2 of \( A|b \) by \(-2\) row 1 of \( A|b \) + row 2 of \( A|b \). The result of this Type 2 ERO is

\[
\begin{bmatrix}
1 & 2 & 3 \\
0 & 0 & -2
\end{bmatrix}
\]  

\( (12) \)

We would now like to transform the second column of \( (12) \) into

\[
\begin{bmatrix}
0 \\
1
\end{bmatrix}
\]

but this is not possible. System \( (12) \) is equivalent to the following system of equations:

\[
\begin{align*}
x_1 + 2x_2 &= 3 \\
0x_1 + 0x_2 &= -2
\end{align*}
\]  

\( (12') \)

Whatever values we give to \( x_1 \) and \( x_2 \), the second equation in \( (12') \) can never be satisfied. Thus, \( (12') \) has no solution. Because \( (12') \) was obtained from \( (11) \) by use of EROs, \( (11) \) also has no solution.

Example 6 illustrates the following idea: *If you apply the Gauss–Jordan method to a linear system and obtain a row of the form \([0 \ 0 \ \cdots \ 0\ c] \ (c \neq 0)\), then the original linear system has no solution.*

**Example 7. Linear System with Infinite Number of Solutions**

Apply the Gauss–Jordan method to the following linear system:

\[
\begin{align*}
x_1 + x_2 &= 1 \\
x_2 + x_3 &= 3 \\
x_1 + 2x_2 + x_3 &= 4
\end{align*}
\]  

\( (13) \)

**Solution** The augmented matrix form of \( (13) \) is

\[
A|b = \begin{bmatrix}
1 & 1 & 0 & 1 \\
0 & 1 & 1 & 3 \\
1 & 2 & 1 & 4
\end{bmatrix}
\]
We begin by replacing row 3 (because the row 2, column 1 value is already 0) of \( A b \) by \\
\(-1(\text{row 1 of } A b) + \text{row 3 of } A b\). The result of this Type 2 ERO is
\\
\[ A_1 b_1 = \begin{bmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 3 \\ 0 & 1 & 1 & 3 \end{bmatrix} \]  
(14)
\\

Next we replace row 1 of \( A_1 b_1 \) by \\
\(-1(\text{row 2 of } A_1 b_1) + \text{row 1 of } A_1 b_1\). The result of this Type 2 ERO is
\\
\[ A_2 b_2 = \begin{bmatrix} 1 & 0 & -1 & -2 \\ 0 & 1 & 1 & 3 \\ 0 & 1 & 1 & 3 \end{bmatrix} \]
\\

Now we replace row 3 of \( A_2 b_2 \) by \\
\(-1(\text{row 2 of } A_2 b_2) + \text{row 3 of } A_2 b_2\). The result of this Type 2 ERO is
\\
\[ A_3 b_3 = \begin{bmatrix} 1 & 0 & -1 & -2 \\ 0 & 1 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix} \]
\\

We would now like to transform the third column of \( A_3 b_3 \) into
\\
\[ \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \]
\\

but this is not possible. The linear system corresponding to \( A_3 b_3 \) is
\\
\[ x_1 - x_3 = -2 \]  
(14.1)
\\
\[ x_2 + x_3 = 3 \]  
(14.2)
\\
\[ 0x_1 + 0x_2 + 0x_3 = 0 \]  
(14.3)
\\
Suppose we assign an arbitrary value \( k \) to \( x_3 \). Then (14.1) will be satisfied if \( x_1 - k = -2 \), or \( x_1 = k - 2 \). Similarly, (14.2) will be satisfied if \( x_2 + k = 3 \), or \( x_2 = 3 - k \). Of course, (14.3) will be satisfied for any values of \( x_1 \), \( x_2 \), and \( x_3 \). Thus, for any number \( k \), \( x_1 = k - 2 \), \( x_2 = 3 - k \), \( x_3 = k \) is a solution to (14). Thus, (14) has an infinite number of solutions (one for each number \( k \)). Because (14) was obtained from (13) via EROs, (13) also has an infinite number of solutions. A more formal characterization of linear systems that have an infinite number of solutions will be given after the following summary of the Gauss–Jordan method.

**Summary of the Gauss–Jordan Method**

**Step 1** To solve \( Ax = b \), write down the augmented matrix \( A b \).

**Step 2** At any stage, define a current row, current column, and current entry (the entry in the current row and column). Begin with row 1 as the current row, column 1 as the current column, and \( a_{11} \) as the current entry. (a) If \( a_{11} \) (the current entry) is nonzero, then use EROs to transform column 1 (the current column) to
\\
\[ \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \]