Matrix & Determinant

Matrix: A matrix is a rectangular array of numbers (real or complex) enclosed in a pair of either parentheses or brackets. For example,

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

Matrices are usually represented by uppercase letters, such as *A*, *B*, *C*, *X*, *Y*, *Z*, etc.. The size of a matrix (or the dimension of a matrix or the order of a matrix) is determined by the number of rows and columns it contains. Since matrix *A* has *m* rows and *n* columns, it is of order $m\times n$ (read "*m* by *n*") and is written as $A_{m\times n}$.

Types of matrices

One of the most frequently used types of matrices is a square matrix. A **square matrix** is a matrix in which the number of rows is equal to the number of columns; that is, $m=n$. An example of a square matrix is $A_{2\times 2}=\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$. If in a matrix $m\neq n$, then that matrix is called a **rectangular matrix**. An example of a rectangular matrix is $B_{2\times 3}=\begin{pmatrix} 1 & 2 & 6 \\ 7 & 5 & 9 \end{pmatrix}$.

Another important type of matrix is a null matrix. A **null matrix** is a matrix in which all the elements are zeros. Examples of a square null matrix and a rectangular null matrix are $C_{2\times 2}=\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ and $D_{2\times 3}=\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ respectively.

Another type of matrix is a diagonal matrix. A square matrix with all non-diagonal elements equal to zero is called a **diagonal matrix**. That is, only the diagonal entries of the square matrix can be non-zero. $A=\begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix}$ and $B=\begin{pmatrix} 3 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 5 \end{pmatrix}$ are diagonal matrices.

A diagonal matrix with all its main diagonal elements are equal is called a **scalar matrix**. $A=\begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix}$ and $B=\begin{pmatrix} a & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & a \end{pmatrix}$ are scalar matrices.

A **triangular matrix** is a matrix whose elements off the main diagonal are all zeros. If $a_{ij}=0$ for all $i<j$ in a matrix (that is, if the elements in a matrix above the main diagonal are all zeros), then

that matrix is called a **lower triangular matrix**. If $a_{ij}=0$ for all $i > j$ in a matrix (that is, if the elements in a matrix below the main diagonal are all zeros), then that matrix is called an **upper triangular matrix**. Notice that for a matrix to be either triangular, or lower or upper triangular, the matrix must be a square matrix. The following matrices, *A*, *B*, and *C*, are examples of triangular, lower triangular, and upper triangular matrices, respectively.

$$A=\begin{pmatrix}1 & 0 & 0\\ 0 & 2 & 0\\ 0 & 0 & 3\end{pmatrix}, B=\begin{pmatrix}1 & 0 & 0\\ 2 & 3 & 0\\ 4 & 5 & 6\end{pmatrix}, \text{ and } C=\begin{pmatrix}1 & 2 & 3\\ 0 & 4 & 5\\ 0 & 0 & 6\end{pmatrix}$$

Another important type of matrix is called an identity matrix or a unit matrix. Imagine what happens if, in a scalar matrix, the elements on the main diagonal are all 1's. Such a matrix is called an **identity or a unit matrix** and is normally denoted by I. Examples of identity matrices are $I_{2\times 2}=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$ and $I_{3\times 3}=\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}$.

The last three types of matrices we discuss here are row matrix, column matrix, and equal matrices. A **row matrix** is a matrix in which there is only one row. A **column matrix** is a matrix in which there is only one column. Examples of row matrix and column matrix are $A_{1\times 2}=\begin{pmatrix}1 & 2\end{pmatrix}$ and $B_{2\times 1}=\begin{pmatrix}1\\ 2\end{pmatrix}$ respectively. Notice that these row and column matrices are nothing other than the row and column vectors, respectively. Two matrices are said to be **equal** only if their corresponding elements are equal and they are of the same order. For example, suppose we have four matrices: $A=\begin{pmatrix}1 & 2\\ 3 & 4\end{pmatrix}$, $B=\begin{pmatrix}1 & 2\\ 3 & 4\end{pmatrix}$, $C=\begin{pmatrix}5 & 6\\ 7 & 8\end{pmatrix}$ and $D=\begin{pmatrix}1 & 2\end{pmatrix}$. Since *A* and *B* have the same order (2×2) and since their corresponding elements are equal, *A*=*B*. Although *A*, *B*, and *C* are of the same order, their corresponding elements are not equal. Therefore, $A=B\neq C$. Since the orders of *A* and *B* are different from that of *D*, $A=B\neq D$.

Matrix operations: scalar multiplication

We often need to use *scalar multiplication* and its properties in other matrix operations. Therefore, we first discuss the multiplication of one or more matrices by one or more scalars. Suppose that we have a matrix $\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ and a scalar s. What we mean by scalar multiplication is that we multiply every element of the matrix $\mathbf{A}$ by s, and we write it as $s\mathbf{A} = s\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} s.a_{11} & s.a_{12} \\ s.a_{21} & s.a_{22} \end{bmatrix}$. Notice that, for scalar multiplication, the matrix need not be a square matrix. Notice also that the reverse is also possible. Two of the important *properties of scalar multiplication* of matrices are the following.

Property I. Suppose that we have two scalars s and t, and a matrix $\mathbf{A}$. Then $(s \pm t)\times\mathbf{A} = s \times \mathbf{A} \pm t \times \mathbf{A}$.

Property II. Suppose that we have a scalar s and two matrices $\mathbf{A}$ and $\mathbf{B}$ of the same order. Then $s \times (\mathbf{A} + \mathbf{B}) = s \times \mathbf{A} + s \times \mathbf{B}$.

Matrix operations: addition and subtraction

Assume that we have two matrices of the same order: $\mathbf{A}_{ij} = [a_{ij}]$ and $\mathbf{B}_{ij} = [b_{ij}]$. Then a new matrix $\mathbf{C}_{ij} = [c_{ij}]$, where $[c_{ij}] = [a_{ij} + b_{ij}]$ for all i and j, and $i = 1, 2, 3, \dots, m$ and $j = 1, 2, 3, \dots, n$, can be generated. What this means is that we add the (i, j)th element of $\mathbf{A}$ to the corresponding (i, j)th element of $\mathbf{B}$ to obtain the (i, j)th of $\mathbf{C}$. As an example consider two matrices of order 2×2: $\mathbf{A}_{2\times 2} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ and $\mathbf{B}_{2\times 2} = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}$. Then the sum, or addition, of the two matrices $\mathbf{A}_{ij}$ and $\mathbf{B}_{ij}$ is given as

$$\mathbf{C}_{2\times 2} = \mathbf{A}_{2\times 2} + \mathbf{B}_{2\times 2} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} + \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} = \begin{bmatrix} 1+5 & 2+6 \\ 3+7 & 4+8 \end{bmatrix} = \begin{bmatrix} 6 & 8 \\ 10 & 12 \end{bmatrix}$$

Subtraction of two or more matrices is similar to the addition of two or more matrices. Suppose that, as in the case of addition, we have two matrices of the same order: $\mathbf{A}_{ij} = [a_{ij}]$ and $\mathbf{B}_{ij} = [b_{ij}]$. Then a new matrix $\mathbf{C}_{ij} = [c_{ij}]$, where $[c_{ij}] = [a_{ij} - b_{ij}]$ for all i and j, and $i = 1, 2, 3, \dots, m$ and $j = 1, 2, 3, \dots, n$, can be generated. What this means is that we subtract the (i, j)th element of $\mathbf{B}$ from the corresponding (i, j)th element of $\mathbf{A}$ to obtain the (i, j)th of $\mathbf{C}$. Notice that, in subtraction, we are in fact adding the negative of $\mathbf{B}$ (which is nothing but a scalar multiple of $\mathbf{B}$ with the scalar being -1) to the corresponding element of $\mathbf{A}$. As an example, consider our last two matrices $\mathbf{A}_{2\times 2}$ and $\mathbf{B}_{2\times 2}$. Then the difference between $\mathbf{A}$ and $\mathbf{B}$ is given as

$$\mathbf{C}_{2\times 2} = \mathbf{A}_{2\times 2} - 1 \times \mathbf{B}_{2\times 2} = \mathbf{A}_{2\times 2} - \mathbf{B}_{2\times 2} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} - \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} = \begin{bmatrix} 1-5 & 2-6 \\ 3-7 & 4-8 \end{bmatrix}$$

$$= \begin{bmatrix} -4 & -4 \\ -4 & -4 \end{bmatrix} = (-4)\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

Notice that for addition (or subtraction) of two matrices, the matrices must be of the same order. This is called the *conformability condition* for addition (or subtraction) of matrices. Notice also that the resulting matrix **C** will be of the same order as those of the original matrices. The important *properties of matrix addition* (or of *matrix subtraction*) are presented below, supposing that we have four matrices of the same order: **A**, **B**, **C**, and **0** (null matrix).

Property I. Matrix addition is commutative: $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$.
Property II. Matrix addition is associative: $(\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C})$.
Property III. Existence of additive inverse: $\mathbf{A} + (-\mathbf{A}) = \mathbf{0}$.
Property IV. Existence of additive identity: $\mathbf{A} + \mathbf{0} = \mathbf{A}$.

Matrix operation: multiplication

Let there be two matrices: $\mathbf{A}_{\boldsymbol{m}\times\boldsymbol{p}} = [a_{ij}]_{m\times p}$ and $\mathbf{B}_{\boldsymbol{p}\times\boldsymbol{n}} = [b_{ij}]_{p\times n}$. Then the product of **A** and **B**, denoted by **C**, is given by $\mathbf{AB} = \mathbf{C} = [c_{ij}]_{m\times n}$, where $c_{ij} = a_{i1}.b_{1j} + a_{i2}.b_{2j} + a_{i3}.b_{3j} + \ldots\ldots\ldots + a_{ip}.b_{pj}$, $i = 1, 2, 3, \ldots, m$ and $j = 1, 2, 3, \ldots, n$. This implies that the (i, j)th element of **C** is obtained by multiplying the ith row of **A** and the jth column of **B** and summing the result. Suppose that the two matrices are: $\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}_{2\times 3}$ and $\mathbf{B} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{bmatrix}_{3\times 2}$.

Assume that we multiply every element in the first row of **A** by the corresponding elements in the first column of **B**, multiply every element in the first row of **A** by the corresponding elements in the second column of **B**, multiply every element in the second row of **A** by the corresponding elements in the first column of **B**, and multiply every element in the second row of **A** by the corresponding elements in the second column of **B**. When we take the sums of these four products and write it in the form of a matrix, we obtain matrix **C** (which is the product of **A** and **B**):

$$\mathbf{C} = \mathbf{A}.\mathbf{B} = \begin{bmatrix} a_{11} \times b_{11} + a_{12} \times b_{21} + a_{13} \times b_{31} & a_{11} \times b_{12} + a_{12} \times b_{22} + a_{13} \times b_{32} \\ a_{21} \times b_{11} + a_{22} \times b_{21} + a_{23} \times b_{31} & a_{21} \times b_{12} + a_{22} \times b_{22} + a_{23} \times b_{32} \end{bmatrix}_{2\times 2}$$

Notice that these are nothing but the scalar products of vectors contained in the matrices **A** and **B**.

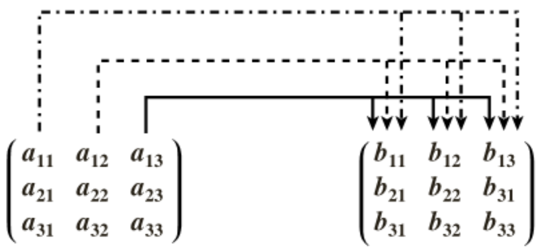

Figure 2.3.1

As an example of finding the products of two matrices, we shall use the matrices **A** and **B**: $\mathbf{A} = \begin{bmatrix} 1 & 2 & 2 \\ 1 & 2 & 2 \end{bmatrix}_{2\times 3}$ and $\mathbf{B} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \\ 2 & 1 \end{bmatrix}_{3\times 2}$. Therefore, the product of **A** and **B** is

$$\mathbf{A.B} = \begin{bmatrix} 1\times 2+2\times 1+2\times 2 & 1\times 1+2\times 2+2\times 1 \\ 1\times 2+2\times 1+2\times 2 & 1\times 1+2\times 2+2\times 1 \end{bmatrix}_{2\times 2} = \begin{bmatrix} 8 & 7 \\ 8 & 7 \end{bmatrix}_{2\times 2}$$

Suppose that our matrices are $\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}_{3\times 3}$ and $\mathbf{B} = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix}_{3\times 3}$. Then, following a procedure similar to the one above or as illustrated in Figure 2.3.1, we obtain the product of **A** and **B** as

$$\mathbf{AB} = \begin{bmatrix} a_{11}\times b_{11}+a_{12}\times b_{21}+a_{13}\times b_{31} & a_{11}\times b_{12}+a_{12}\times b_{22}+a_{13}\times b_{32} & a_{11}\times b_{13}+a_{12}\times b_{23}+a_{13}\times b_{33} \\ a_{21}\times b_{11}+a_{22}\times b_{21}+a_{23}\times b_{31} & a_{21}\times b_{12}+a_{22}\times b_{22}+a_{23}\times b_{32} & a_{31}\times b_{13}+a_{32}\times b_{23}+a_{33}\times b_{33} \\ a_{31}\times b_{11}+a_{32}\times b_{21}+a_{33}\times b_{31} & a_{31}\times b_{12}+a_{32}\times b_{22}+a_{33}\times b_{32} & a_{31}\times b_{13}+a_{32}\times b_{23}+a_{33}\times b_{33} \end{bmatrix}_{3\times 3}$$

As an example of finding the product of two 3×3 matrices, suppose that the matrices are $\mathbf{A} = \begin{bmatrix} 2 & 2 & 1 \\ 2 & 1 & 1 \\ 2 & 2 & 1 \end{bmatrix}_{3\times 3}$ and $\mathbf{B} = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 2 & 2 \\ 1 & 1 & 2 \end{bmatrix}_{3\times 3}$. Then their product will be

$$\mathbf{A.B} = \begin{bmatrix} 2\times 1+2\times 1+1\times 1 & 2\times 1+2\times 2+1\times 1 & 2\times 2+2\times 2+1\times 2 \\ 2\times 1+1\times 1+1\times 1 & 2\times 1+1\times 2+1\times 1 & 2\times 2+1\times 2+1\times 2 \\ 2\times 1+2\times 1+1\times 1 & 2\times 1+2\times 2+1\times 1 & 2\times 2+2\times 2+1\times 2 \end{bmatrix}_{3\times 3} = \begin{bmatrix} 5 & 7 & 10 \\ 4 & 5 & 8 \\ 5 & 7 & 10 \end{bmatrix}_{3\times 3}$$

The reader would have noticed that in the multiplication of $\mathbf{A}$ by $\mathbf{B}$, the number of columns of $\mathbf{A}$ was equal to the number of rows of $\mathbf{B}$. This is called the *conformability condition for the multiplication of matrices*. This implies that two matrices ($\mathbf{A}$ and $\mathbf{B}$) are said to be conformable for multiplication (if we multiply $\mathbf{A}$ by $\mathbf{B}$) only if the number of columns in $\mathbf{A}$ is equal to the number of rows in $\mathbf{B}$. Otherwise the product will not exist. This points to the fact that the order of $\mathbf{A.B}$ will be equal to the number of rows of $\mathbf{A}$ and the number of columns of $\mathbf{B}$. In our first example above, the order of $\mathbf{A}$ was 2×3 and the order of $\mathbf{B}$ was 3×2 giving us the product $\mathbf{A.B}$ whose order was 2×2. In the second example, $\mathbf{A}$ and $\mathbf{B}$ were of the same order (3×3) yielding us $\mathbf{A.B}$ whose order was 3×3.

Similarly, if we multiply $\mathbf{B}$ by $\mathbf{A}$, then the number of rows in $\mathbf{B}$ must be equal to the number of columns in $\mathbf{A}$, and the order of the resulting matrix will be equal to the number of rows of $\mathbf{B}$ and the number of columns of $\mathbf{A}$. Therefore, one needs to check whether the matrices are conformable for multiplication before one carries out multiplication. Notice that those two square matrices are always (in both ways) conformable for multiplication. We state below the *properties of matrix multiplication*, assuming that we have five matrices, $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, an identity matrix ($\mathbf{I}$), and a null matrix ($\mathbf{0}$), such that they are conformable for the operations indicated below.

Property I. Matrix multiplication is associative: $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$.

Property II. Matrix multiplication is distributive over addition and subtraction: $\mathbf{A}(\mathbf{B} \pm \mathbf{C}) = \mathbf{AB} \pm \mathbf{AC}$ and $(\mathbf{B} \pm \mathbf{C})\mathbf{A} = \mathbf{BA} \pm \mathbf{CA}$.

Property III. Matrix multiplication is *not always* commutative: $\mathbf{AB} = \mathbf{BA}$ or $\mathbf{AB} \neq \mathbf{BA}$ (that is, $\mathbf{AB}$ *may or may not be equal to* $\mathbf{BA}$, even if these two operations are conformable or even if the products exist).

Property IV. $\mathbf{AI} = \mathbf{A} = \mathbf{IA}$.

Property V. $\mathbf{A0} = \mathbf{0} = \mathbf{0A}$.

Transpose of a matrix and powers of square matrices

Suppose that we have the matrix $\mathbf{A} = [a_{ij}]_{m \times n}$. Then the transpose of this matrix, denoted by $\mathbf{A}^{\mathbf{T}}$, is defined as $\mathbf{A}^{\mathbf{T}} = [b_{ji}]_{n \times m}$, where $b_{ji} = a_{ij}$ for all $i = 1, 2, 3, \ldots, m$ and $j = 1, 2, 3, \ldots, n$. What this means is that we obtain $\mathbf{A}^{\mathbf{T}}$ by interchanging the rows and columns of $\mathbf{A}$; that is, the rows of $\mathbf{A}^{\mathbf{T}}$ will be columns of $\mathbf{A}$. This implies that the transposition of a matrix reverses the order of the matrix; that is, if $\mathbf{A}$ is of order $m \times n$, then $\mathbf{A}^{\mathbf{T}}$ will be of order $n \times m$. Notice that the transposes of matrices $\mathbf{I}$ and $\mathbf{0}$ will be the same; i.e., $\mathbf{I} = \mathbf{I}^{\mathbf{T}}$ and $\mathbf{0} = \mathbf{0}^{\mathbf{T}}$.

As an example, consider the matrices $\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}_{2\times3}$ and $\mathbf{B} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}_{3\times3}$. Then,

$\mathbf{A}^{\mathbf{T}} = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}_{3\times2}$ and $\mathbf{B}^{\mathbf{T}} = \begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix}_{3\times3}$. The reader will have noticed that if a square matrix (such as $\mathbf{B}$ above) is transposed, the diagonal elements remain unchanged.

Two matrices related to the transpose of a matrix are *symmetric matrix* and *skew symmetric matrix*. A square matrix, say $\mathbf{C}$, is said to be a symmetric matrix if $\mathbf{C}^{\mathbf{T}} = \mathbf{C}$. Notice that $\mathbf{I}$ and $\mathbf{0}$ are two examples of a symmetric matrices. Other examples include ${\mathbf{C}^{\mathbf{T}}}_{\mathbf{1}} = \mathbf{C}_1 = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$, ${\mathbf{C}^{\mathbf{T}}}_{\mathbf{2}} = \mathbf{C}_2 = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$, ${\mathbf{C}^{\mathbf{T}}}_{\mathbf{3}} = \mathbf{C}_3 = \begin{bmatrix} a & b & c \\ b & d & e \\ c & e & f \end{bmatrix}$, and ${\mathbf{C}^{\mathbf{T}}}_{\mathbf{4}} = \mathbf{C}_4 = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{bmatrix}$.

A square matrix, say $\mathbf{C}$, is said to be a skew symmetric matrix if $\mathbf{C}^{\mathbf{T}} = -\mathbf{C}$. Examples of a skew symmetric matrix include ${\mathbf{C}^{\mathbf{T}}}_{\mathbf{1}} = \begin{bmatrix} 0 & -b \\ b & 0 \end{bmatrix}$ given $\mathbf{C}_1 = \begin{bmatrix} 0 & b \\ -b & 0 \end{bmatrix}$, and ${\mathbf{C}^{\mathbf{T}}}_{\mathbf{2}} = \begin{bmatrix} 0 & b & c \\ -b & 0 & e \\ -c & -e & 0 \end{bmatrix}$ given $\mathbf{C}_2 = \begin{bmatrix} 0 & -b & -c \\ b & 0 & -e \\ c & e & 0 \end{bmatrix}$.

Property I. Let there be a matrix $\mathbf{A}$ and scalar s, then $(s.\ \mathbf{A})^{\mathbf{T}} = s.\ (\mathbf{A})^{\mathbf{T}}$.

Property II. For any matrix $\mathbf{A}$, then $(\mathbf{A}^{\mathbf{T}})^{\mathbf{T}} = \mathbf{A}$.

Property III. Let there be two matrices, $\mathbf{A}$ and $\mathbf{B}$, of the same order, then $(\mathbf{A} \pm \mathbf{B})^{\mathbf{T}} = \mathbf{A}^{\mathbf{T}} \pm \mathbf{B}^{\mathbf{T}}$.

Property IV. Let there be two matrices, $\mathbf{A}$ and $\mathbf{B}$, and their product $\mathbf{A.B}$ exists, then $(\mathbf{A.B})^{\mathbf{T}} = \mathbf{B}^{\mathbf{T}}.\mathbf{A}^{\mathbf{T}}$.

We now consider the *powers of a square matrix*. Let $\mathbf{A}$ be a square matrix. Then, we define $\mathbf{A.A} = \mathbf{A}^{\mathbf{2}}$; $\mathbf{A.A}^{\mathbf{2}} = \mathbf{A}^{\mathbf{3}}$; $\mathbf{A.A}^{\mathbf{3}} = \mathbf{A}^{\mathbf{4}}$;…; and $\mathbf{A.A}^{\boldsymbol{n-1}} = \mathbf{A}^{\boldsymbol{n}}$, where n is any positive integer. Notice that $\mathbf{A}^{\boldsymbol{m}}.\mathbf{A}^{\boldsymbol{n}} = \mathbf{A}^{\boldsymbol{m+n}}$, and $(\mathbf{A}^{\boldsymbol{m}})^{\boldsymbol{n}} = \mathbf{A}^{\boldsymbol{m.n}}$.

Orthogonal Matrix

A matrix $A = (a_{ij})_{n\times n}$ is said to be a orthogonal matrix if $AA^T = A^T A = I$.

Example: $A = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} & -\frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{2}{3} & -\frac{1}{3} \end{pmatrix}$.

2.4.1 Meaning and evaluation of determinants

So far we were dealing with vectors and matrices. Some students may wonder if we can associate a number with every square matrix. Yes, we can, and it is called its *determinant*. But, how can one find such a number? We discuss below the methods of finding the values that we associate with square matrices of different orders (that is, *evaluating determinants*).

Suppose that we have a 2×2 matrix given by $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$. Then the determinant of $\mathbf{A}$, denoted by $|\mathbf{A}|$, is written as $|\mathbf{A}| = \begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix}$. We can find the value we associate with $\mathbf{A}$ (that is, $|\mathbf{A}|$) by finding the products of the elements on the *main diagonal* and subtracting from this the product of the *off-diagonal* elements. The product of the elements on the main diagonal in the present example is $1 \times 4 = 4$ and the product of the off-diagonal elements is $2 \times 3 = 6$. Therefore, the determinant is $|\mathbf{A}| = \begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} = 4 - 6 = -2$.

In general, if we have a 2×2 matrix $\mathbf{B} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$, then the determinant of $\mathbf{B}$ is given by

$$|\mathbf{B}| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = (a_{11} \times a_{22}) - (a_{21} \times a_{12}) \tag{2.4.1}$$

Now suppose we have a matrix of order 3×3 given by

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \longrightarrow \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{matrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{matrix}$$

One easy way of evaluating the determinant of this third-order matrix is to write it as given alongside the matrix $\mathbf{A}$. Notice that the last two columns above are the first two columns of $\mathbf{A}$. The next step is to find the product of the elements on the downward sloping arrows, and sum these products, which is given by $a_{11} \times a_{22} \times a_{33} + a_{12} \times a_{23} \times a_{31} + a_{13} \times a_{21} \times a_{32}$. The third step is to find the product of the elements on the upward sloping arrows, and sum them too, which is given by $a_{31} \times a_{22} \times a_{13} + a_{32} \times a_{23} \times a_{11} + a_{33} \times a_{21} \times a_{12}$. The final step is to subtract the latter from the former. Therefore, the determinant of $\mathbf{A}$ is given by

$$\begin{aligned} |\mathbf{A}| &= (a_{11} \times a_{22} \times a_{33} + a_{12} \times a_{23} \times a_{31} + a_{13} \times a_{21} \times a_{32}) \\ &\quad - (a_{31} \times a_{22} \times a_{13} + a_{32} \times a_{23} \times a_{11} + a_{33} \times a_{21} \times a_{12}) \end{aligned} \tag{2.4.2}$$

As an example, consider the following 3×3 matrix:

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \longrightarrow \begin{bmatrix} \mathbf{1} & \mathbf{2} & \mathbf{3} \\ \mathbf{4} & \mathbf{5} & \mathbf{6} \\ \mathbf{7} & \mathbf{8} & \mathbf{9} \end{bmatrix} \begin{matrix} \mathbf{1} & \mathbf{2} \\ \mathbf{4} & \mathbf{5} \\ \mathbf{7} & \mathbf{8} \end{matrix}$$

As before, we write the matrix as given above alongside the matrix $\mathbf{A}$. First, we find the product of elements on the downward sloping arrows, and sum these products. This sum is equal to $1 \times 5 \times 9 + 2 \times 6 \times 7 + 3 \times 4 \times 8 = 225$. Second, we obtain the product of elements on the upward sloping arrows, which is equal to $7 \times 5 \times 3 + 8 \times 6 \times 1 + 9 \times 4 \times 2 = 225$. The difference between these two sums is $225 - 225 = 0$. Therefore, the determinant is zero; that is, $|\mathbf{A}| = 0$. If the determinant of a matrix is zero, then the determinant is said to *vanish*. It is important to notice that the above method using arrows, popularly called *Sarrus' rule*, works *only in the case of* 3×3 *matrices*. Notice that the determinant of the matrix $\mathbf{A} = [3]$ is given by $|\mathbf{A}| = 3$. Also notice that determinants are defined only for square matrices. We will consider later the evaluation of determinants of square matrices of order greater than 3.

2.4.2 Properties of determinants

Determinants obey some important properties. These properties are the following.

Property I. The determinant of a square matrix, $\mathbf{A}$, and the determinant of its transpose are equal, that is, $|\mathbf{A}| = |\mathbf{A}^{\mathbf{T}}|$. This implies that the determinant of a matrix remains unchanged if the rows and columns of that matrix are interchanged.

Property II. If two rows (or columns) of a determinant, $|\mathbf{A}|$, are interchanged, then the sign of the determinant changes but the determinant remains the same.

Property III. If two rows (or columns) of a determinant, $|\mathbf{A}|$, are identical, then the determinant is zero.

Property IV. If any one row (or one column) of a determinant is multiplied by a scalar s, then the determinant will be multiplied by the scalar s.

Property V. If a multiple of one row (or column) is added to any row (or column), then the determinant remains unchanged.

Property VI. The determinant of an identity matrix, $\mathbf{I}$, of any order is equal to 1.

Property VII. Suppose we have two matrices, $\mathbf{A}$ and $\mathbf{B}$. Also suppose that the product $\mathbf{AB}$ exists. Then $|\mathbf{AB}| = |\mathbf{A}|.|\mathbf{B}|$.

2.4.3 *Sub-matrices, minors, co-factors, Laplace expansion, and determinants of matrices of order greater than 3*

So far we were evaluating the determinants of square matrices of order less than or equal to 3 and we applied Sarrus' rule for this. Since Sarrus' rule is not applicable to evaluate determinants of order greater than 3, we need to find out a method that can deal with such determinants. There exists another method, called the method of *expansion by co-factors*, to evaluate determinants of any order. However, to discuss this method, we need to explain few concepts such as *sub-matrices*, *minors*, and *co-factors*. We shall do this first.

For simplicity, we begin our exposition of the concepts with a 3×3 matrix $\mathbf{A}$ we used in the previous section. Suppose, now, that we discard the first row and the first column of $\mathbf{A}$ (one can discard any row and any column); that is, we discard $i = 1$ and $j = 1$. Then the matrix we obtain is called a sub-matrix (of order 2×2) and we denote it by $\mathbf{S}_{11} = \begin{bmatrix} 5 & 6 \\ 8 & 9 \end{bmatrix}$, where $\mathbf{S}$ represents the *sub-matrix*, and the two subscripts (1's) represent the numbers of discarded rows and columns. One can generate a number of sub-matrices like this depending upon the rows and columns discarded. For example, two other sub-matrices of $\mathbf{A}$ are $\mathbf{S}_{12} = \begin{bmatrix} 4 & 6 \\ 7 & 9 \end{bmatrix}$ and $\mathbf{S}_{13} = \begin{bmatrix} 4 & 5 \\ 7 & 8 \end{bmatrix}$.

We are now ready to define a sub-matrix formally. Assume that we have a square matrix $\mathbf{A} = (a_{ij})_{n \times n}$. The square matrix that can be generated if we discard the ith row and jth column of $\mathbf{A}$, denoted by $\mathbf{S}_{ij}$, is called the ij^{th} sub-matrix of $\mathbf{A}$. We can generate, in total, nine sub-matrices from a 3×3 matrix such as $\mathbf{A}$ above. Therefore, all the sub-matrices of order two that can be generated from $\mathbf{A}$ above are $\mathbf{S_{11}}$, $\mathbf{S_{12}}$, $\mathbf{S_{13}}$, $\mathbf{S_{21}}$, $\mathbf{S_{22}}$, $\mathbf{S_{23}}$, $\mathbf{S_{31}}$, $\mathbf{S_{32}}$, and $\mathbf{S_{33}}$.

Now suppose that we take the determinant of a sub-matrix, say $\mathbf{S}_{ij}$, and denote it by $|\mathbf{S}_{ij}|$. This determinant is called the ij^{th} *minor* of $\mathbf{A}$, and we denote it by $\mathbf{M}_{ij} = |\mathbf{S}_{ij}|$. For example, $\mathbf{M_{11}} = |\mathbf{S_{11}}| = (5 \times 9 - 8 \times 6) = 45 - 48 = -3$. Similarly, $\mathbf{M}_{12} = |\mathbf{S_{12}}| = (4 \times 9 - 7 \times 6) = 36 - 42 = -6$, and $\mathbf{M_{13}} = |\mathbf{S_{13}}| = (4 \times 8 - 7 \times 5) = 32 - 35 = -3$. Therefore, given $\mathbf{A}$, we will have nine minors in total: $\mathbf{M_{11}} = -3$, $\mathbf{M_{12}} = -6$, $\mathbf{M_{13}} = -3$, $\mathbf{M_{21}} = -6$, $\mathbf{M_{22}} = -12$, $\mathbf{M_{23}} = -6$, $\mathbf{M_{31}} = -3$, $\mathbf{M_{32}} = -6$, and $\mathbf{M_{33}} = -3$.

Assume now that we multiply $\mathbf{M_{11}}$ by $(+1)$, $\mathbf{M_{12}}$ by (-1), $\mathbf{M_{13}}$ by 1, and so on. Notice that we are multiplying each minor by the scalar $(+1)$ or the scalar (-1) depending upon the sum of the subscripted numbers of the corresponding minor. If the sum is even, the scalar is $+1$; if the sum is odd, the scalar is (-1). In general, we can write the scalar as $(-1)^{(i+j)}$. If we multiply $\mathbf{M}_{ij}$ by $(-1)^{(i+j)}$, then it is called the *signed minor* of $\mathbf{M}_{ij}$. The signed minor of $\mathbf{M}_{ij}$ (that is, $(-1)^{(i+j)}\,\mathbf{M}_{ij}$) is also called the *co-factor* of $\mathbf{M}_{ij}$, and is denoted by $\mathbf{C}_{ij}$. Therefore, the co-factor of $\mathbf{M_{11}}$ is $\mathbf{C_{11}} = (-1)^{(i+j)}\ \mathbf{M_{11}} = (-1)^{(1+1)}\ \mathbf{M_{11}} = (-1)^2\ \mathbf{M_{11}} = (1)\ \mathbf{M_{11}} = \mathbf{M_{11}} = -3$; of $\mathbf{M_{12}}$ is $\mathbf{C_{12}} = (-1)\ \mathbf{M_{12}} = 6$; and of $\mathbf{M_{13}}$ is $\mathbf{C_{13}} = (1)\ \mathbf{M_{13}} = -3$. Similarly, $\mathbf{C_{21}} = (-1)\ \mathbf{M_{21}} = 6$; $\mathbf{C_{22}} = (1)\ \mathbf{M_{22}} = -12$; $\mathbf{C_{23}} = (-1)\ \mathbf{M_{23}} = 6$; $\mathbf{C_{31}} = (1)\ \mathbf{M_{31}} = -3$; $\mathbf{C_{32}} = (-1)\ \mathbf{M_{32}} = 6$; and $\mathbf{C_{33}} = (1)\ \mathbf{M_{33}} = -3$.

We are now ready to use the concepts of co-factors to evaluate determinants of order 3×3. Consider, for example, our last 3×3 matrix $\mathbf{A}$ to find $|\mathbf{A}|$ using co-factors. Then $|\mathbf{A}|$ is defined as $|\mathbf{A}| = 1\times\ \mathbf{C_{11}} + 2\times\ \mathbf{C_{12}} + 3\times\ \mathbf{C_{13}}$. Since $\mathbf{C_{11}} = (-1)^{(i+j)}\ \mathbf{M_{11}} = (-1)^{(1+1)}\ \mathbf{M_{11}} = (-1)^2\ \mathbf{M_{11}} = (1)\ \mathbf{M_{11}} = (1)\ \mathbf{M_{11}}$; $\mathbf{C_{12}} = (-1)\ \mathbf{M_{12}}$; and $\mathbf{C_{13}} = (1)\ \mathbf{M_{13}}$, we can write $|\mathbf{A}| = 1 \times (1)\ \mathbf{M_{11}} + 2 \times (-1)\ \mathbf{M_{12}} + 3 \times (1)\ \mathbf{M_{13}}$. This implies that $|\mathbf{A}| = 1 \times (1) \times \begin{vmatrix} 5 & 6 \\ 8 & 9 \end{vmatrix} + 2 \times (-1) \times \begin{vmatrix} 4 & 6 \\ 7 & 9 \end{vmatrix} + 3 \times (1) \times \begin{vmatrix} 4 & 5 \\ 7 & 8 \end{vmatrix} = 1 \times \begin{vmatrix} 5 & 6 \\ 8 & 9 \end{vmatrix} - 2 \times \begin{vmatrix} 4 & 6 \\ 7 & 9 \end{vmatrix} + 3 \begin{vmatrix} 4 & 5 \\ 7 & 8 \end{vmatrix}$. Since $\begin{vmatrix} 5 & 6 \\ 8 & 9 \end{vmatrix} = -3$, $\begin{vmatrix} 4 & 6 \\ 7 & 9 \end{vmatrix} = -6$, and $\begin{vmatrix} 4 & 5 \\ 7 & 8 \end{vmatrix} = -3$, the last result simplifies to $|\mathbf{A}| = 1 \times -3 + -2 \times -6 + 3 \times -3 = -3 + 12 - 9 = 12 - 12 = 0$. This was precisely the result we obtained when we used Sarrus' rule in the previous section.

Let us now generalize the idea with the help of a general matrix, $\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$.

Then we define $|\mathbf{A}|$ as

$$\begin{aligned} |\mathbf{A}| &= a_{11} \times (1) \times \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} + a_{12} \times (-1) \times \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} \\ &\quad + a_{13} \times (1) \times \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} = a_{11}\mathbf{C_{11}} + a_{12}\mathbf{C_{12}} + a_{13}\mathbf{C_{13}} \end{aligned} \tag{2.4.3}$$

Equation (2.4.3) implies that $|\mathbf{A}| = \sum_j^3 a_{ij}C_{ij}$. If we evaluate (that is, expand) the last matrix using the j^{th} column, then we can write $|\mathbf{A}| = \sum_i^3 a_{ij}C_{ij}$. Notice that if the matrix is an $n \times n$ matrix, then we have $|\mathbf{A}| = \sum_j^n a_{ij}C_{ij}$ if we use the j^{th} column and $|\mathbf{A}| = \sum_i^n a_{ij}C_{ij}$ if we use the i^{th} row. Also notice that $\sum_j^n a_{ij}C_{ij} = \sum_i^n a_{ij}C_{ij}$; that is, the determinant remains the same irrespective of the row or column of the matrix that is used to evaluate it. This method of evaluating (or expanding) a determinant, using co-factors, is called the *Laplace expansion* of a determinant.

We now consider evaluating the determinant of a 4×4 matrix, $\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{bmatrix}$.

Following exactly the process of Laplace expansion as detailed above, we obtain

$$|\mathbf{A}| = 1\times(1)\times\begin{vmatrix} 6 & 7 & 8 \\ 10 & 11 & 12 \\ 14 & 15 & 16 \end{vmatrix} + 2\times(-1)\times\begin{vmatrix} 5 & 7 & 8 \\ 9 & 11 & 12 \\ 13 & 15 & 16 \end{vmatrix}$$

$$+3\times(1)\times\begin{vmatrix} 5 & 6 & 8 \\ 9 & 10 & 12 \\ 13 & 14 & 16 \end{vmatrix} + 4\times(-1)\times\begin{vmatrix} 5 & 6 & 7 \\ 9 & 10 & 11 \\ 13 & 14 & 15 \end{vmatrix} = 0.$$

We described in the last section concepts like sub-matrices, minors, and co-factors. Let us now construct a matrix using the co-factors called the *co-factor matrix*. Given the 3×3 matrix $\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$, the co-factors we obtained in the previous section were $\mathbf{C_{11}} = -3$, $\mathbf{C_{12}} = 6$, $\mathbf{C_{13}} = -3$, $\mathbf{C_{21}} = 6$, $\mathbf{C_{22}} = -12$, $\mathbf{C_{23}} = 6$, $\mathbf{C_{31}} = -3$, $\mathbf{C_{32}} = 6$, and $\mathbf{C_{33}} = -3$. Therefore, the co-factor matrix denoted by $\mathbf{C_{FA}}$, where $\mathbf{A}$ represents the original matrix in our present example, will be

$$\mathbf{C_{FA}} = \begin{bmatrix} C_{11} & C_{12} & C_{13} \\ C_{21} & C_{22} & C_{23} \\ C_{31} & C_{32} & C_{33} \end{bmatrix} = \begin{bmatrix} -3 & 6 & -3 \\ 6 & -12 & 6 \\ -3 & 6 & -3 \end{bmatrix} \tag{2.4.4}$$

Notice that if our original square matrix were of order 2×2, its co-factor matrix would be of order 2×2; if it were of order 4×4, its co-factor matrix would be of order 4×4; and so on. If we transpose the co-factor matrix, we obtain what is called the *adjoint matrix*, denoted by $\mathbf{A_{adj}}$. Therefore, the adjoint matrix in the case of our present example will be

$$\mathbf{A_{adj}} = (\mathbf{C_{FA}})^{\mathbf{T}} = \begin{bmatrix} -3 & 6 & -3 \\ 6 & -12 & 6 \\ -3 & 6 & -3 \end{bmatrix} \tag{2.4.5}$$

Notice that in this case $\mathbf{C_{FA}} = (\mathbf{C_{FA}})^{\mathbf{T}}$ because they are symmetric matrices.

Systems of linear equations:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= d_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= d_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 &= d_3 \end{aligned} \tag{2.3.4}$$

One can now write the coefficients, the variables, and the constants in the above system using matrices as $\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$, and $\mathbf{d} = \begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix}$. Therefore, the system in equation (2.3.4) can be written in matrix form as $\mathbf{Ax} = \mathbf{d}$. Suppose now that we have an $m \times n$ system of the form

$$
\begin{aligned}
&a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1n}x_n = d_1 \\
&a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \cdots + a_{2n}x_n = d_2 \\
&a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + \cdots + a_{3n}x_n = d_3 \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
&a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + \cdots + a_{mn}x_n = d_m
\end{aligned}
\tag{2.3.5}
$$

The above $m \times n$ system can be written, using matrices, as $\mathbf{Ax} = \mathbf{d}$, where

$$
\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \ldots & a_{1n} \\ a_{21} & a_{22} & a_{21} & \ldots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \ldots & a_{3n} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & a_{m3} & \ldots. & a_{mn} \end{bmatrix}, \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ .. \\ x_n \end{bmatrix}, \text{ and } \mathbf{d} = \begin{bmatrix} d_1 \\ d_2 \\ d_3 \\ .. \\ d_m \end{bmatrix}
$$

2.4.5 Solution of SSLEs: Cramer's rule

In Section 1.6.5 we used graphical, substitution, and elimination methods to solve SSLEs. In Sections 2.3.7 and 2.3.8 we used elementary row operations or the Gauss–Jordan elimination method and the Gauss elimination method, respectively, to solve SSLEs. Notice that each method has its own merits and demerits.

However, in this section, we discuss another popular method called *Cramer's rule*. For this, we use the $m \times n$ (where $m = n$) SSLEs given in equation (2.3.5). We expressed there the system as the matrix equation $\mathbf{Ax} = \mathbf{d}$. Now suppose that we replace the first column of $\mathbf{A}$ by $\mathbf{d}$ and denote the resulting matrix by $\mathbf{A_1}$. When we continue the replacement of the other columns of $\mathbf{A}$ by $\mathbf{d}$, we obtain n such new matrices (including $\mathbf{A_1}$) each denoted by $\mathbf{A_1}$, $\mathbf{A_2}$, $\mathbf{A_3}$,…, and $\mathbf{A_n}$. We know that the determinants of these matrices will be $|\mathbf{A_1}|$, $|\mathbf{A_2}|$, $|\mathbf{A_3}|$,…, and $|\mathbf{A_n}|$, respectively. Notice that the determinant of the original coefficient matrix is $|\mathbf{A}|$. It can now be shown that

$$
x_1 = |\mathbf{A_1}| \div |\mathbf{A}|, x_2 = |\mathbf{A_2}| \div |\mathbf{A}|, x_3 = |\mathbf{A_3}| \div |\mathbf{A}|, \ldots, x_j = |\mathbf{A_j}| \div |\mathbf{A}|, \ldots, x_n = |\mathbf{A_n}| \div |\mathbf{A}|
\tag{2.4.6}
$$

Let us now state Cramer's rule formally. Let $\mathbf{A}$ be an $n \times n$ matrix and $|\mathbf{A}| \neq 0$. Then the unique solution to the system $\mathbf{Ax} = \mathbf{d}$ is given by $x_1 = |\mathbf{A_1}| \div |\mathbf{A}|$, $x_2 = |\mathbf{A_2}| \div |\mathbf{A}|$, $x_3 = |\mathbf{A_3}| \div |\mathbf{A}|$,…, $x_j = |\mathbf{A_j}| \div |\mathbf{A}|$,…, $x_n = |\mathbf{A_n}| \div |\mathbf{A}|$, where $|\mathbf{A_j}|$ denotes the determinant of the adjoint matrix obtained by replacing the jth column of the coefficient matrix $\mathbf{A}$ by the vector of the constants ($\mathbf{d}$). Notice that we can use either Sarrus' rule if $\mathbf{A}$ (or $\mathbf{A}_j$)is of order 3×3 or Laplace expansion if $\mathbf{A}$ is of order larger than 3×3 to find the determinants of $\mathbf{A}$ (or of $\mathbf{A}_j$).

Example 3. Suppose that prices of three goods sold by a company are interrelated by the system $-4x_1 + 3x_2 + 3x_3 = 10$, $3x_1 - 4x_2 + 3x_3 = 10$, and $3x_1 + 3x_2 - 4x_3 = 10$, where x_1, x_2, and x_3 denote price in dollars of goods 1, 2, and 3, respectively. Find the prices that solve the system using Cramer's rule.

Solution. We first write the above system in matrix form as $\mathbf{Ax} = \mathbf{d}$, where $\mathbf{A} = \begin{bmatrix} -4 & 3 & 3 \\ 3 & -4 & 3 \\ 3 & 3 & -4 \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$, and $\mathbf{d} = \begin{bmatrix} 10 \\ 10 \\ 10 \end{bmatrix}$. Following Cramer's rule, we have $x_j = |\mathbf{A}_j| \div |\mathbf{A}|$, where $|\mathbf{A}_j|$ denotes the determinant of the coefficient matrix whose jth column is replaced by the constant vector $\mathbf{d}$. As above, we can obtain that $|\mathbf{A}_1| = 490$, $|\mathbf{A}_2| = 490$, and $|\mathbf{A}_3| = 490$. The determinant of the coefficient matrix is $|\mathbf{A}| = 98$. Therefore, $x_1 = |\mathbf{A}_1| \div |\mathbf{A}| = 490/98 = 5$, $x_2 = |\mathbf{A}_2| \div |\mathbf{A}| = 490/98 = 5$, and $x_3 = |\mathbf{A}_3| \div |\mathbf{A}| = 490/98 = 5$. Therefore, the prices that solve the system are $x_1 = \$5$, $x_2 = \$5$, and $x_3 = \$5$.

2.5 Inverse of a matrix

2.5.1 Meaning of inverse

We know from elementary algebra that $a \times (1/a) = a \times a^{-1} = a^{1-1} = a^0 = 1$. Here we call a^{-1} the *multiplicative inverse* of a. We shall now attempt to extend this result in algebra to matrices. Suppose that we have two square matrices of the same order: matrix $\mathbf{A}$ and identity matrix $\mathbf{I}$. Can we now find a matrix $\mathbf{B}$ such that $\mathbf{A}.\mathbf{B} = \mathbf{I}$? If we can, then $\mathbf{B}$ is called the *inverse* of $\mathbf{A}$, and is written as $\mathbf{A}^{-1}$ (that is, $\mathbf{B} = \mathbf{A}^{-1}$). Then $\mathbf{A}$ is said to be *invertible*.

Since $\mathbf{A}.\mathbf{B} = \mathbf{I}$, $\mathbf{A}$ is the inverse of $\mathbf{B}$ and $\mathbf{B}$ is the inverse of $\mathbf{A}$; that is, $\mathbf{A}$ and $\mathbf{B}$ are the inverses of each other. Notice that inverses are defined only for square matrices. However, how does one find the inverse in the first place? Two methods are generally used to find the inverse of a square matrix. One of these methods uses the determinant and the adjoint of the given matrix. The other method uses the Gauss–Jordan method discussed in Section 2.3.8. Let us discuss each of these below.

2.5.2 Finding inverse using determinant and adjoint matrix

Before finding the inverse, we need to check whether the inverse of a matrix exists or not. Above, we used the matrix equation $\mathbf{AB} = \mathbf{I}$. We need to know whether $\mathbf{A}$ is invertible or not. We know, from the properties of determinants, that $|\mathbf{AB}| = |\mathbf{A}|.|\mathbf{B}| = |\mathbf{I}|$. It is clear from this equation that if $|\mathbf{A}| = 0$, then we will obtain a contradictory result that $|\mathbf{A}|.|\mathbf{B}| = |\mathbf{I}| = 0.|\mathbf{B}| = \mathbf{0} = |\mathbf{I}|$. The only way to avoid this contradiction is that $|\mathbf{A}| \neq 0$. This suggests that the condition for $\mathbf{A}$ to have an inverse is that its determinant, $|\mathbf{A}|$, must be nonzero or nonvanishing. Therefore, one can check whether a matrix is invertible or not by checking its determinant. Notice that a square matrix will have only one inverse.

We are now ready to apply the determinant of a matrix and its adjoint to find the inverse of that matrix. We defined the adjoint of matrix $\mathbf{A}$ (denoted by $\mathbf{A_{adj}}$) as the transpose of the co-factor matrix, $\mathbf{C_{FA}}$. That is, $\mathbf{A_{adj}} = (\mathbf{C_{FA}})^{\mathbf{T}}$. One can show that the product of $\mathbf{A}$ and the adjoint of $\mathbf{A}$ is equal to the product of the determinant of $\mathbf{A}$ and the identity matrix $\mathbf{I}$ of the same order; that is, $\mathbf{A}.\mathbf{A_{adj}} = |\mathbf{A}|.\mathbf{I}$. Now dividing both sides of this equation by $|\mathbf{A}|$ yields $(\mathbf{A}.\mathbf{A_{adj}})/|\mathbf{A}| = |\mathbf{A}|.\mathbf{I}/|\mathbf{A}| = \mathbf{I}$. And pre-multiplying both sides of the last equation by $\mathbf{A^{-1}}$ gives $\mathbf{A^{-1}}(\mathbf{A}.\mathbf{A_{adj}})/|\mathbf{A}| = \mathbf{A^{-1}}\mathbf{I}$. Since $\mathbf{A^{-1}}\mathbf{A} = \mathbf{I}$, $(\mathbf{I}.\mathbf{A_{adj}})/|\mathbf{A}| = \mathbf{A_{adj}}/|\mathbf{A}|$ and $\mathbf{A^{-1}}\mathbf{I} = \mathbf{A^{-1}}$, the last equation can be written as $\mathbf{A_{adj}}/|\mathbf{A}| = \mathbf{A^{-1}}$. Therefore, the inverse of $\mathbf{A}$ is given by

$$\mathbf{A^{-1}} = \mathbf{A_{adj}} \div |\mathbf{A}| \tag{2.5.1}$$

As an example, consider the 3×3 matrix $\mathbf{A} = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 1 & 2 \\ 2 & 2 & 1 \end{bmatrix}$. The determinant of this matrix is $|\mathbf{A}| = 3$. The co-factor matrix of $\mathbf{A}$ is $\mathbf{C_{FA}} = \begin{bmatrix} C_{11} & C_{12} & C_{13} \\ C_{21} & C_{22} & C_{23} \\ C_{31} & C_{32} & C_{33} \end{bmatrix} = \begin{bmatrix} -3 & 3 & 0 \\ 0 & -1 & 2 \\ 3 & -1 & -1 \end{bmatrix}$, which implies that $\mathbf{A_{adj}} = (\mathbf{C_{FA}})^{\mathbf{T}} = \begin{bmatrix} -3 & 0 & 3 \\ 3 & -1 & -1 \\ 0 & 2 & -1 \end{bmatrix}$. Therefore, $\mathbf{A^{-1}} = \mathbf{A_{adj}} \div |\mathbf{A}| =$

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

2.5.5 Solution of SSLEs using inverse

Suppose, as in Section 2.4.5, we have an $m \times n$ (with $m = n$) SSLEs represented by the matrix equation $\mathbf{Ax} = \mathbf{d}$, where $\mathbf{A}$ represents the $n \times n$ coefficient matrix, $\mathbf{x}$ represents the $n \times 1$ variable vector, and $\mathbf{d}$ represents the $n \times 1$ constant vector. Let us now *post-multiply* $\mathbf{A}$ and *pre-multiply* $\mathbf{d}$ of the equation $\mathbf{Ax} = \mathbf{d}$ by the inverse of $\mathbf{A}$, $\mathbf{A^{-1}}$. These yield $\mathbf{AA^{-1}x} = \mathbf{A^{-1}d}$. Since $\mathbf{A\,A^{-1}} = \mathbf{I}$, the last equation simplifies to

$$\mathbf{Ix} = \mathbf{x} = \mathbf{A^{-1}d} \tag{2.5.2}$$

Equation (2.5.2) is a very important result in linear algebra. It states that the solution of a SSLEs is given by the product of the inverse of the matrix of coefficients and the vector of the constants of the system. This is called the *method of solving SSLEs using inverse*.

Example 2. Suppose that the three markets in an economy are related by the prices x_1, x_2, and x_3. Also suppose that the relationships among these prices are given by $-2x_1 + 2x_2 + 2x_3 - 25 = 0$, $2x_1 - 2x_2 + 2x_3 - 25 = 0$, and $2x_1 + 2x_2 - 2x_3 - 25 = 0$. Find the prices in dollars that solve the system using inverse.

Solution. As above, we first write the system in matrix form as $\mathbf{Ax} = \mathbf{d}$, where $\mathbf{A} = \begin{bmatrix} -2 & 2 & 2 \\ 2 & -2 & 2 \\ 2 & 2 & -2 \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$, and $\mathbf{d} = \begin{bmatrix} 25 \\ 25 \\ 25 \end{bmatrix}$. The solution to this system can be obtained by applying equation (2.5.2): $\mathbf{x} = \mathbf{A}^{-1}\mathbf{d}$. But, for this we need to find $\mathbf{A}^{-1}$. We can find that $\mathbf{A}^{-1} = \begin{bmatrix} 0 & 1/4 & 1/4 \\ 1/4 & 0 & 1/4 \\ 1/4 & 1/4 & 0 \end{bmatrix}$. Then, application of equation (2.5.2) yields $\mathbf{x} = \mathbf{A}^{-1}\mathbf{d} = \begin{bmatrix} 0 & 1/4 & 1/4 \\ 1/4 & 0 & 1/4 \\ 1/4 & 1/4 & 0 \end{bmatrix} \begin{bmatrix} 25 \\ 25 \\ 25 \end{bmatrix} = \begin{bmatrix} 12.5 \\ 12.5 \\ 12.5 \end{bmatrix}$. Therefore, the prices that solve the system are $x_1 = x_2 = x_3 = \$12.5$.