1.2 Matrices

From the last topic, we must have observed that, rather than the unknowns, their coefficients are involved in the various steps of the elimination method.

A compact form of writing a linear system without the unknowns is a *matrix*. Apart from being more convenient, a matrix lends itself more easily to mathematical manipulations and, in fact, the notion that has been developed on matrices is one of the most interesting and most important theoretical aspects of Mathematics.

A matrix is a formally defined as being a rectangular array of mn real or complex numbers arranged in m horizontal rows and n vertical columns.

For example, the linear system represented by

$$\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \dots\dots\dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots\dots\dots + a_{2n}x_n = b_2 \\ \quad . \qquad\quad . \qquad\qquad\qquad . \qquad . \\ \quad . \qquad\quad . \qquad\qquad\qquad . \qquad . \\ \quad . \qquad\quad . \qquad\qquad\qquad . \qquad . \\ a_{m1}x_1 + a_{m2}x_2 + \dots\dots\dots + a_{mn}x_n = b_m \end{array}$$

has *coefficient matrix*

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

In short form, we write $A = [\, a_{ij} \,]$, that is, the matrix A consists of elements of the form a_{ij} (found on the i^{th} row and the j^{th} column).

A matrix which has only one row of elements is sometimes known as *a row vector* and a matrix consisting only of one column of elements is called *a column vector*.

We will now proceed to some definitions which are essential to know before we start any rigorous treatment of matrices.

DEFINITION

A matrix is said to be of order $(m \times n)$ if it has m horizontal rows and n vertical columns.

EXAMPLE

$$\begin{bmatrix} 1 & 0 & -4 & 6 \\ -2 & 1 & 3 & 1 \\ 4 & 5 & 2 & -3 \end{bmatrix}$$
$$(3 \times 4)$$

DEFINITION

A matrix of order $(n \times n)$ is known as a *square* matrix.

EXAMPLE

$$\begin{bmatrix} 1 & 2 \\ 3 & -4 \end{bmatrix} \qquad \begin{bmatrix} 2 & -1 & 0 \\ 1 & 3 & 2 \\ 4 & 0 & -3 \end{bmatrix}$$
$$(2 \times 2) \qquad (3 \times 3)$$

DEFINITION

The *main diagonal* of a square matrix A is the one consisting of the elements a_{11}, a_{22}, a_{33}, ……, a_{nn}.

DEFINITION

A square matrix A is also a *diagonal* matrix if $a_{ij} = 0,\ i \neq j$.

EXAMPLE

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

DEFINITION

A square matrix A is also a *scalar* matrix if $a_{ij} = c,\ i = j$.
$$a_{ij} = 0,\ i \neq j .$$

EXAMPLE

$$\begin{bmatrix} -4 & 0 & 0 & 0 \\ 0 & -4 & 0 & 0 \\ 0 & 0 & -4 & 0 \\ 0 & 0 & 0 & -4 \end{bmatrix}$$

DEFINITION

Two matrices A and B are said to be *equal* if $[\, a_{ij} \,] = [\, b_{ij} \,]$.
It is fairly obvious that two equal matrices should have the same order.

Operations on matrices

Addition

The sum of two matrices A and B is possible only if they have the same order.
If $A + B = C$, then $c_{ij} = a_{ij} + b_{ij}$.

EXAMPLE

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

Scalar multiplication

Given a matrix $A = [\, a_{ij} \,]$, then, for any scalar r,

$$r\,[\, a_{ij} \,] = [\, ra_{ij} \,]$$

that is, when a matrix is multiplied by a scalar, all its elements are multiplied by the scalar.

Transpose

If $A = [\, a_{ij} \,]$, then $A^T = [\, a_{ji} \,]$, where A^T is the transpose of A.

Thus, the transpose of a matrix A is obtained by interchanging its rows and columns.

EXAMPLE

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

$$(2\times 3) \qquad\qquad (3\times 2)$$

Dot or **Inner Product**

DEFINITION

The dot product or inner product of two n-vectors **a** and **b** where

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

is given as

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

EXAMPLE

$$\begin{bmatrix} 2 & -3 & 1 & 4 \end{bmatrix} \begin{bmatrix} -1 \\ 4 \\ 2 \\ 0 \end{bmatrix} = \text{(-2 + -12 + 2 + 0) = -12 .}$$

Matrix multiplication

DEFINITION

If A is a matrix of order $(m \times n)$ and B is a matrix of order $(n \times p)$ then the *product* of A and B, denoted by AB is the matrix C of order $(m \times p)$, defined by

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + ...+ a_{in}b_{nj} = \sum_{k=1}^{n} a_{ik}b_{kj}$$

This can be clearly illustrated by the following:

$$\begin{bmatrix} a_{11} & a_{12} & .... & a_{1n} \\ a_{21} & a_{22} & .... & a_{2n} \\ . & . & & . \\ . & . & & . \\ a_{i1} & a_{i2} & .... & a_{in} \\ . & . & & . \\ . & . & & . \\ a_{m1} & a_{m2} & .... & a_{mn} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & ..... & b_{1j} & ..... & b_{1p} \\ b_{21} & b_{22} & ..... & b_{2j} & ..... & b_{2p} \\ . & . & & . & & . \\ . & . & & . & & . \\ b_{n1} & b_{n2} & ..... & b_{nj} & ..... & b_{np} \end{bmatrix}$$

$$row_i(A) \qquad\qquad col_j(B)$$

$$= \begin{bmatrix} c_{11} & c_{12} & ..... & c_{1p} \\ c_{21} & c_{22} & ..... & c_{2p} \\ . & . & & . \\ & & c_{ij} & \\ . & . & & . \\ c_{m1} & c_{m2} & ..... & c_{mp} \end{bmatrix}$$

$$row_i(A)\ \ col_j(B) = \sum_{k=1}^{n} a_{ik}b_{kj}$$

EXAMPLE

$$\begin{bmatrix} 1 & -3 & 2 \\ 4 & 0 & -1 \end{bmatrix} \begin{bmatrix} 2 & -1 & 3 & 2 \\ 1 & 0 & -4 & 3 \\ 5 & -2 & -3 & 1 \end{bmatrix} = \begin{bmatrix} 9 & -5 & 9 & -5 \\ 3 & -2 & 15 & 7 \end{bmatrix}$$

$$(2\times 3) \qquad (3\times 4) \qquad (2\times 4)$$

We just note that the number of columns of the first matrix should be equal to the number of rows of the second matrix.

The matrix BA

It is also important to check the properties of the matrix *BA* with respect to the matrix *AB*. Four different possibilities exist:

(1) *BA* is undefined.
(2) *BA* is defined but its size differs from that of *AB*.
(3) $AB = BA$.
(4) $AB \neq BA$.

The following examples illustrate perfectly each situation.

EXAMPLE 1

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

The product *AB* is possible whereas *BA* cannot be computed because the number of columns of *B* is not equal to the number of rows of *A*.

EXAMPLE 2

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

$$(3\times 2) \qquad\qquad (2\times 3)$$

It is left as an exercise to check that

$$AB = \begin{bmatrix} -2 & 2 & 7 \\ -15 & -6 & 0 \\ 6 & 0 & -6 \end{bmatrix} \text{ and that } BA = \begin{bmatrix} -2 & 8 \\ -3 & -12 \end{bmatrix}$$

$$(3\times 3) \qquad\qquad (2\times 2)$$

EXAMPLE 3A

One of the matrices is the zero matrix (*trivial* case), which is quite natural !

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

EXAMPLE 3B

One matrix is the *inverse* of the other (we will deal with inverses later on).

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

EXAMPLE 3C

The product of two *symmetric* matrices (to be dealt with later on)

$$\begin{bmatrix} -1 & 2 \\ 2 & -1 \end{bmatrix}\begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix} = \begin{bmatrix} -5 & 7 \\ 7 & -5 \end{bmatrix} \text{ and } \begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix}\begin{bmatrix} -1 & 2 \\ 2 & -1 \end{bmatrix} = \begin{bmatrix} -5 & 7 \\ 7 & -5 \end{bmatrix}$$

EXAMPLE 4

$$\begin{bmatrix} 1 & 2 \\ -3 & 1 \end{bmatrix}\begin{bmatrix} 3 & 0 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 4 \\ -10 & 2 \end{bmatrix}$$

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

Matrix – Vector Product Decomposition

Let $A = \begin{bmatrix} a_{11} & a_{12} & \cdots\cdots & a_{1n} \\ a_{21} & a_{22} & \cdots\cdots & a_{2n} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ a_{m1} & a_{m2} & \cdots\cdots & a_{mn} \end{bmatrix}$ be an $(m \times n)$ matrix and $\mathbf{c} = \begin{bmatrix} c_1 \\ c_2 \\ \cdot \\ \cdot \\ \cdot \\ c_n \end{bmatrix}$ be an *n*-vector.

Therefore,

$$\begin{bmatrix} a_{11} & a_{12} & \cdots\cdots & a_{1n} \\ a_{21} & a_{22} & \cdots\cdots & a_{2n} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ a_{m1} & a_{m2} & \cdots\cdots & a_{mn} \end{bmatrix}\begin{bmatrix} c_1 \\ c_2 \\ \cdot \\ \cdot \\ \cdot \\ c_n \end{bmatrix} = \begin{bmatrix} row_1(A)\bullet \mathrm{c} \\ row_2(A)\bullet \mathrm{c} \\ \cdot \\ \cdot \\ \cdot \\ row_m(A)\bullet \mathrm{c} \end{bmatrix} = \begin{bmatrix} a_{11}c_1 + a_{12}c_2 + \dots\dots + a_{1n}c_n \\ a_{21}c_1 + a_{22}c_2 + \dots\dots + a_{2n}c_n \\ \cdot \\ \cdot \\ \cdot \\ a_{m1}c_1 + a_{m2}c_2 + \dots\dots + a_{mn}c_n \end{bmatrix}$$

$$(m \times n) \qquad (n \times 1) \qquad (m \times 1)$$

The right hand side $(m \times 1)$ matrix can be *decomposed* as

$$c_1\begin{bmatrix} a_{11} \\ a_{21} \\ \cdot \\ \cdot \\ \cdot \\ a_{mi} \end{bmatrix} + c_2\begin{bmatrix} a_{12} \\ a_{22} \\ \cdot \\ \cdot \\ \cdot \\ a_{m2} \end{bmatrix} + \dots\dots\dots + c_3\begin{bmatrix} a_{1n} \\ a_{2n} \\ \cdot \\ \cdot \\ \cdot \\ a_{mn} \end{bmatrix} = c_1 \cdot col(A) + c_2 \cdot col_2(A) + \dots\dots + c_n \cdot col_n(A)$$

The above matrix - vector product can therefore be expressed as a *linear combination* of the columns of A.

EXAMPLE 1

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

$$= \begin{bmatrix} 1 \\ 4 \end{bmatrix} + \begin{bmatrix} 0 \\ 6 \end{bmatrix} + \begin{bmatrix} 1 \\ -2 \end{bmatrix} = \begin{bmatrix} 2 \\ 8 \end{bmatrix}$$

EXAMPLE 2

Consider the product of $A = \begin{bmatrix} 1 & 3 \\ 0 & 1 \\ -2 & 2 \end{bmatrix}$ and $B = \begin{bmatrix} -1 & 2 & 4 \\ 2 & 3 & 1 \end{bmatrix}$;

$$col_1(AB) = (-1) \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix} + (2) \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 5 \\ 2 \\ 4 \end{bmatrix}$$

$$col_1(AB) = (2) \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix} + (3) \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 11 \\ 3 \\ 2 \end{bmatrix}$$

$$col_1(AB) = (4) \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix} + (1) \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 7 \\ 1 \\ -6 \end{bmatrix}$$

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

Partitioned matrices

If we start with an $(m \times n)$ matrix A = $[\,a_{ij}\,]$ and we delete some of its rows and columns, we have a *submatrix* of *A*.

EXAMPLE

If $A = \begin{bmatrix} 2 & 1 & -4 & 0 \\ -1 & -2 & 3 & 2 \\ -3 & 0 & 1 & -1 \end{bmatrix}$, then, by deleting its 2nd row and 3rd column, we have the

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

A matrix can be *partitioned* into submatrices by drawing horizontal lines between rows and vertical lines between columns. This can be done in several ways.

EXAMPLE

$$A = \left[\begin{array}{cc|ccc|c} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} & a_{26} \\ \hline a_{31} & a_{32} & a_{33} & a_{34} & a_{35} & a_{36} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} & a_{46} \end{array}\right] \text{ can also be written as } \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \end{bmatrix}$$

where, for example, $A_{21} = \begin{bmatrix} a_{31} & a_{32} \\ a_{41} & a_{42} \end{bmatrix}$.

NOTE

The augmented matrix of a linear system is a partitioned matrix.

Partitioned matrices can also undergo matrix multiplication by a process known as *block multiplication*.

EXAMPLE

$$\text{Let } A = \left[\begin{array}{rr|rr} 1 & 0 & 4 & 1 \\ -2 & 1 & 3 & 1 \\ \hline 0 & 2 & -1 & 3 \\ 2 & 1 & 4 & 2 \end{array}\right] = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \text{ and}$$

$$B = \left[\begin{array}{rrr|rrr} 3 & 1 & -1 & 1 & 0 & 2 \\ 2 & 0 & 1 & -3 & -2 & 1 \\ \hline 5 & 1 & -2 & 0 & 1 & 3 \\ 3 & 4 & -1 & 2 & -1 & 1 \end{array}\right] = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$$

If $AB = C = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix}$, then, for example, $C_{11} = A_{11}B_{11} + A_{12}B_{21}$, that is,

$$C_{11} = \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix}\begin{bmatrix} 3 & 1 & -1 \\ 2 & 0 & -1 \end{bmatrix} + \begin{bmatrix} 4 & 1 \\ 3 & 1 \end{bmatrix}\begin{bmatrix} 5 & 1 & -2 \\ 3 & 4 & -1 \end{bmatrix}$$

$$= \begin{bmatrix} 3 & 1 & -1 \\ -4 & -2 & 3 \end{bmatrix} + \begin{bmatrix} 23 & 8 & -9 \\ 18 & 7 & -7 \end{bmatrix} = \begin{bmatrix} 26 & 9 & -10 \\ 14 & 5 & --4 \end{bmatrix}$$

$$\text{so that } C = \left[\begin{array}{rrr|r} 26 & 9 & -10 & \\ 14 & 5 & -4 & \\ \hline & & & \\ & & & \end{array}\right]$$

The rest is left as an exercise.
Try and double-check by multiplying the entire matrices directly.

NOTE

The advantage of partitioned matrices becomes obvious when we deal with matrices which exceed the memory of a computer. Obviously, *the partitioning is done so that the products are defined*.

Occasionally, we will make use of the *summation notation*. This is a very compact and useful notation.

$$\sum_{i=1}^{n} a_i \text{ means } a_1 + a_2 + \dots\dots + a_n$$

The variable i is known as the *index* and it acts as a counter. The constants a and b are the lower and upper limits of summation respectively and they are necessarily integers. The index starts at a and is incremented by 1 until it reaches b. Meanwhile, each value that it takes is substituted in the general term (in the above example, a_i) so that each substitution generates one term of the series.

NOTE

The letter used for the index can be changed so that

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

The summation satisfies the following properties:

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

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

We saw earlier on that a dot product $\mathbf{a} \bullet \mathbf{b}$ can be expressed using the summation notation as

$$\mathbf{a} \bullet \mathbf{b} = a_1 b_1 + a_2 b_2 + \dots\dots + a_n b_n = \sum_{i=1}^{n} a_i b_i$$

We can also form double sums like $\sum_{j=1}^{3} \sum_{i=1}^{4} a_{ij}$. This can be expanded by first summing on i and sum the resulting expression on j.

Thus, $\sum_{j=1}^{3} \sum_{i=1}^{4} a_{ij} = \sum_{j=1}^{3} (a_{1j} + a_{2j} + a_{3j} + a_{4j})$

$$= (a_{11} + a_{21} + a_{31} + a_{41}) + (a_{12} + a_{22} + a_{32} + a_{42}) + (a_{13} + a_{23} + a_{33} + a_{43})$$