"JUST THE MATHS"

UNIT NUMBER

9.1

MATRICES 1
(Definitions & elementary matrix algebra)

by

A.J.Hobson

UNIT 9.1 - MATRICES 1
DEFINITIONS AND ELEMENTARY MATRIX ALGEBRA

9.1.1 INTRODUCTION

(a) Presentation of Data
Sets of numerical information can often be presented as a rectangular **"array"** of numbers. For example, a football results table might look like this:

TEAM	PLAYED	WON	DRAWN	LOST
Blackburn	22	11	6	5
Burnley	22	9	6	7
Chelsea	21	7	8	6
Leicester	21	6	8	7
Stoke	21	6	6	9

If the headings are taken for granted, we write simply

$$\begin{bmatrix} 22 & 11 & 6 & 5 \\ 22 & 9 & 6 & 7 \\ 21 & 7 & 8 & 6 \\ 21 & 6 & 8 & 7 \\ 21 & 6 & 6 & 9 \end{bmatrix}$$

and this symbol is called a **"matrix"**.

Note:
The mould in which printers once cast **type** was called a matrix. In mathematics, the word "matrix" signifies that we have spaces into which **numbers** can be placed.

(b) Presentation of Algebraic Results
In two-dimensional geometry, the **"vector"**, $\underline{OP_0}$, joining the origin to the point $P_0(x_0, y_0)$ can be moved to the position $\underline{OP_1}$ joining the origin to the point $P_1(x_1, y_1)$ by means of a **"reflection"**, a **"rotation"**, a **"magnification"** or a combination of such operations. It can be shown that the relationship between the co-ordinates of P_0 and P_1 in any such operation is given by

$$\begin{aligned} x_1 &= ax_0 + by_0, \\ y_1 &= cx_0 + dy_0. \end{aligned}$$

ILLUSTRATIONS

1. The equations

$$\begin{aligned} x_1 &= -x_0, \\ y_1 &= y_0 \end{aligned}$$

   represent a reflection in the y-axis.

2. The equations

$$\begin{aligned} x_1 &= kx_0, \\ y_1 &= ky_0 \end{aligned}$$

   represent a magnification when $\mid k \mid > 1$ and a contraction when $\mid k \mid < 1$.

3. The equations

$$\begin{aligned} x_1 &= x_0 \cos \theta - y_0 \sin \theta, \\ y_1 &= x_0 \sin \theta + y_0 \cos \theta \end{aligned}$$

   represent a rotation of $\underline{\text{OP}}_0$ through an angle θ in a counter-clockwise direction.

Such an operation is called a **"linear transformation"** and is completely specified by the coefficients a, b, c, d in the correct positions. When referring to a linear transformation, it is therefore more convenient to write the matrix symbol

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

Sometimes, we want to show that a transformation

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

operates on the vector $\underline{\text{OP}}_0$, transforming it into the vector $\underline{\text{OP}}_1$. We write

$$\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \cdot \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}.$$

The two new symbols are still called matrices but can be given a special name as we see later.

9.1.2 DEFINITIONS

(i) A matrix is **a rectangular array of numbers** arranged in rows (horizontally), columns (vertically) and enclosed in brackets.

ILLUSTRATIONS:

$$\begin{bmatrix} 1 & 3 & 7 \\ 9 & 9 & 10 \end{bmatrix}, \begin{bmatrix} 10 & 16 & 17 \\ 3 & 5 & 11 \\ 4 & 0 & 10 \end{bmatrix}, \begin{bmatrix} 1 & 5 & 7 \end{bmatrix}.$$

Note:
The rows are counted from the top to the bottom of the matrix and the columns are counted from the left to the right of the matrix.

(ii) Any number within the array is called an **"element"** of the matrix. The term $ij-th$ element refers to the element lying in the i-th row and the j-th column of the matrix.

(iii) If a matrix has m rows and n columns, it is called a **"matrix of order $m \times n$"** or simply an **"$m \times n$ matrix"**. It clearly has mn elements.

(iv) A matrix of order $m \times m$ is called a **"square matrix"**.

Note:
A matrix of order 1×1 is considered to be the same as a single number. This assumption has particular significance in the definition of a matrix product (see later).

(v) A matrix of order $m \times 1$ is called a **"column vector"** and a matrix of order $1 \times n$ is called a **"row vector"**.

Note:
Matrices of order 2×1, 1×2, 3×1 and 1×3 have easy physical interpretations as vectors. But, in abstract linear algebra, there are vectors of **any** dimension; so the names "column vector" and "row vector" are retained no matter how many elements there are.

(vi) An arbitrary matrix whose elements and order do not have to be specified may be denoted by a single capital letter such as A,B,C, etc.

An arbitrary matrix of order $m \times n$ may be denoted fully by the symbol

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{bmatrix}$$

where, in each double-subscript, the first number is the row number and the second number is the column number. In a situation where the matrix has already been denoted by a single capital letter, the subsequent use of the full notation should include the corresponding small letter on which to attach the double-subscripts.

If the matrix is square, the elements $a_{11}, a_{22}, a_{33}, ...a_{mm}$ are called the **"diagonal elements"** and their sum is called the **"trace"** of the matrix. This line of elements is called the **"leading diagonal"** of the matrix.

An abbreviated form of the full notation is $[a_{ij}]_{m \times n}$

EXAMPLE
A matrix $A = [a_{ij}]_{2 \times 3}$ is such that $a_{ij} = i^2 + 2j$. Write out A in full.

Solution

This is an artificially contrived example, but will serve to illustrate the use of the double-subscript notation. We obtain

$$A = \begin{bmatrix} 3 & 5 & 7 \\ 6 & 8 & 10 \end{bmatrix}.$$

(vii) Given a matrix, A, of order $m \times n$, the matrix of order $n \times m$ obtained from A by writing the rows as columns is called the **"transpose"** of A and is denoted by A^T - some books use A' or $\tilde{A}$.

ILLUSTRATION

If

$$A = \begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \\ a_3 & b_3 \end{bmatrix},$$

then,

$$A^T = \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix}$$

(viii) A matrix, A, is said to be **"symmetric"** if $A = A^T$.

ILLUSTRATION

$$A = \begin{bmatrix} a & h & g \\ h & b & f \\ g & f & c \end{bmatrix} \text{ is symmetric.}$$

4

(ix) A matrix, A, is said to be **"skew-symmetric"** if the elements of A^T are minus the corresponding elements of A itself. This will mean that the leading diagonal elements must be zero.

ILLUSTRATION

$$A = \begin{bmatrix} 0 & a & -b \\ -a & 0 & c \\ b & -c & 0 \end{bmatrix} \text{ is skew} - \text{symmetric.}$$

(x) A matrix is said to be **"diagonal"** if the elements which are not on the leading diagonal have value zero while the elements on the leading diagonal are not all equal to zero.

9.1.3 THE ALGEBRA OF MATRICES (Part One)

An **"Algebra"** (coming from the Arabic word AL-JABR) refers to any mathematical system which uses the concepts of equality, addition, subtraction, multiplication and division. For example, the algebra of numbers is what we normally call **"arithmetic"**; but algebraical concepts can be applied to other mathematical systems of which matrices is one.

In meeting a new mathematical system for the first time, the concepts of equality, addition, subtraction, multiplication and division need to be properly defined, and that is the purpose of the present section. In some cases, the definitions are fairly obvious, but they need to be made without contradicting ideas already established in the system of numbers which matrices depend on.

1. Equality

Unlike a determinant, a matrix does **not** have a numerical value; so the use of equality here is not that which is made in elementary arithmetic. Rather we use it to mean "is the same as".

Two matrices are said to be equal if they have **the same order** and also **pairs of elements in corresponding positions are equal in value**.

In symbols, we could say that

$$[a_{ij}]_{m \times n} = [b_{ij}]_{m \times n}$$

provided

$$a_{ij} = b_{ij}.$$

For example,

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

implies that $x = 1, y = 2$ and $z = 3$.

We shall meet this type of conclusion later in a discussion on the solution, by matrices, of a set of simultaneous equations (see Unit 9.3).

2. Addition and Subtraction

The scientific applications of matrices require us to define the sum and difference of two matrices only when they have the same order.

The sum of two matrices of order $m \times n$ is formed by adding together the pairs of elements in corresponding positions. Similarly, the difference of two matrices of order $m \times n$ is formed by subtracting the pairs of elements in corresponding positions.

In symbols, we could say that

$$[a_{ij}]_{m \times n} \pm [b_{ij}]_{m \times n} = [c_{ij}]_{m \times n},$$

where

$$c_{ij} = a_{ij} \pm b_{ij}.$$

Note:
It may occur, in a calculation, that a matrix has to be subtracted from itself, in which case we would obtain another matrix of the same order but whose elements are all zero. This type of matrix is called a **"null matrix"** and may be denoted, for short, by $[0]_{m \times n}$ or just $[0]$ when the order is understood.

ILLUSTRATION
A grocer has two shops and, in each shop, he sells apples, oranges and bananas. The sales, in kilogrammes, of each fruit for the two shops on two separate days are represented by the matrices

$$\begin{bmatrix} 36 & 25 & 10 \\ 20 & 30 & 15 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 40 & 30 & 12 \\ 22 & 35 & 20 \end{bmatrix}$$

where the rows refer to the two shops and the columns refer to apples, oranges and bananas respectively.

The total sales, in kilogrammes, of each fruit for the two shops on both days together are represented by the matrix

$$\begin{bmatrix} 76 & 55 & 22 \\ 42 & 65 & 35 \end{bmatrix}.$$

The differences in sales of each fruit for the two shops between the second day and the first day are represented by the matrix

$$\begin{bmatrix} 4 & 5 & 2 \\ 2 & 5 & 5 \end{bmatrix}$$

3. Additive Identities and Additive Inverses

For the sake of completeness, we mention here that

(a) A null matrix, when added to any other matrix of the same order, leaves that other matrix identically the same as it was to start with. For this reason, a null matrix behaves as an **"additive identity"**.

(b) If any matrix A is added to the matrix $-$A (that is, the matrix obtained from A by reversing the signs of all the elements) the result obtained is the corresponding null matrix. For this reason $-$A can be called the **"additive inverse"** of A.

9.1.4 EXERCISES

1. State the order of each of the following matrices:

   (a) $\begin{bmatrix} 1 & 3 & 0 \\ -1 & 2 & 4 \end{bmatrix}$; (b) $\begin{bmatrix} 5 \\ 6 \end{bmatrix}$; (c) $\begin{bmatrix} 3 & 5 & 4 \end{bmatrix}$.

2. Two of the following matrices are equal to each other. Which are they ?

   (a) $\begin{bmatrix} \frac{1}{2} & \frac{1}{\sqrt{3}} \\ 1 & 0 \end{bmatrix}$; (b) $\begin{bmatrix} \sin 60° & \tan 30° \\ \tan 45° & \sin 90° \end{bmatrix}$; (c) $\begin{bmatrix} \cos 60° & \frac{\sqrt{3}}{3} \\ \sin 90° & \cos 90° \end{bmatrix}$.

3. Determine the values of x, y and z given that

   $$\begin{bmatrix} x+y & y+z \\ x+z & 3x-2y \end{bmatrix} = \begin{bmatrix} 5 & 7 \\ 6 & 0 \end{bmatrix}.$$

4. If A $= \begin{bmatrix} 1 & -3 & 2 \\ 4 & 0 & 1 \end{bmatrix}$, B $= \begin{bmatrix} 4 & 3 & 2 \\ -1 & 0 & 2 \end{bmatrix}$ and C $= \begin{bmatrix} -1 & 4 & -2 \\ 0 & 1 & 0 \end{bmatrix}$,

   determine the elements of the following matrices:

   (a) A + B; (b) (A + B) + C; (c) A + (B + C);

   (d) A - B - C; (e) $A^T + B^T$; (f) $(A + B)^T$.

5. Determine which pairs of the following matrices can be added and, for those which can, state the sum:

$$A = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, B = \begin{bmatrix} 4 \\ 2 \\ 1 \end{bmatrix}, C = [4 \quad 2 \quad 1], D = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix},$$

$$E = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, F = \begin{bmatrix} 2 & 3 \\ 5 & 0 \end{bmatrix}, G = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, H = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, J = [3 \quad 3 \quad 3].$$

9.1.5 ANSWERS TO EXERCISES

1. (a) 2×3; (b) 2×1; (c) 1×3.

2. (a) and (c) are equal to each other.

3. $x = 2, y = 3$ and $z = 4$.

4. (a) $\begin{bmatrix} 5 & 0 & 4 \\ 3 & 0 & 3 \end{bmatrix}$; (b) $\begin{bmatrix} 4 & 4 & 2 \\ 3 & 1 & 3 \end{bmatrix}$; (c) $\begin{bmatrix} 4 & 4 & 2 \\ 3 & 1 & 3 \end{bmatrix}$;

   (d) $\begin{bmatrix} -2 & -10 & 2 \\ 5 & -1 & -1 \end{bmatrix}$; (e) $\begin{bmatrix} 5 & 3 \\ 0 & 0 \\ 4 & 3 \end{bmatrix}$; (f) $\begin{bmatrix} 5 & 3 \\ 0 & 0 \\ 4 & 3 \end{bmatrix}$.

5. $A + E = E + A = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$, $B + G = G + B = \begin{bmatrix} 5 \\ 2 \\ 2 \end{bmatrix}$

   $C + J = J + C = [7 \quad 5 \quad 4]$, $D + F = F + D = \begin{bmatrix} 2 & 3 \\ 5 & 0 \end{bmatrix}$

   $G + H = H + G = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$, $B + H = H + B = \begin{bmatrix} 4 \\ 2 \\ 1 \end{bmatrix}$.

   That is, twelve possible additions but only six distinct results.