[Academic Script]

Vector Spaces and Matrices

Subject:	Business Economics
Course:	B. A. (Hons.), 2nd Semester, Undergraduate
Paper No. & Title:	Paper – 202 Mathematics for Business Economics
Unit No. & Title:	Unit – 4 Linear Algebra
Lecture No. & Title:	Lecture – 1 Vector Spaces and Matrices

Academic Script

1. Introduction

We shall begin by looking at $\mathbb{R}^n$ from a view point wherein we look at elements of $\mathbb{R}^n$ as vectors. Our approach shall be not that of Physicists who start looking at vectors as entities which have direction and magnitude. We shall also not adopt the pure axiomatic approach either. Our approach will be quite informal in the sense we only want to appreciate that some of the important properties of $\mathbb{R}^n$ are taken as the defining axioms in the definition of Vector space.

2. Vector space operations

Note that two elements $x = (x_1, x_2, \dots x_n)\ and\ y = (y_1, y_2, \dots y_n) \in \mathbb{R}^n$ are equal if all the corresponding coordinates are equal i.e.

$$x = (x_1, x_2, \dots x_n) = (y_1, y_2, \dots y_n) = y \text{ iff } x_i = y_i\ for\ all\ i = 1,2, \dots n$$

Suppose $x, y, z \in \mathbb{R}^n\ and\ \alpha, \beta \in \mathbb{R}$ and if the operations are defined as

$$x + y = (x_1, x_2, \dots x_n) + (y_1, y_2, \dots y_n) = (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n)$$

$$\alpha x = (\alpha x_1, \alpha x_2, \dots \alpha x_n)$$

i.e. operations defined this way are called point wise defined operations. They are called the addition and scalar multiplication. These operations have the properties:

(i) For each $x, y \in \mathbb{R}^n\ and\ \alpha \in \mathbb{R}, x + y \in \mathbb{R}^n$ and $\alpha x \in \mathbb{R}^n$.

(ii) For each $x, y, z \in \mathbb{R}^n$, $x + (y + z) = (x + y) + z$. This property is called associativity.

(iii) There exists element $0 = (0,0, \dots 0) \in \mathbb{R}^n$ with the property that $x + 0 = 0 + x = x$ for all $x \in \mathbb{R}^n$.

(iv) For each element $x = (x_1, x_2, \dots x_n) \in \mathbb{R}^n$ there exists element which is denoted by $-x = (-x_1, -x_2, \dots - x_n) \in \mathbb{R}^n$ such that $x + (-x) = 0$.

(v)For all $x, y \in \mathbb{R}^n$, $x + y = y + x$

(vi) For all $x \in \mathbb{R}^n$ and $\alpha, \beta \in \mathbb{R}$ $(a\beta)(x) = a(\beta x)$

(vii) For all $x, y \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$ $a(x + y) = ax + ay$

(viii) For all $x \in \mathbb{R}^n$ and $\alpha, \beta \in \mathbb{R}$ $(a + \beta)x = ax + \beta x$

(ix) $1x = x$ $for\ all$ $x \in \mathbb{R}^n$

With this algebraic structure we say that $\mathbb{R}^n$ is real vector space or real linear space or simply vector space or linear space. The elements of $\mathbb{R}^n$ are called vectors and the elements of $\mathbb{R}$ in this context are called scalars. The two operations are respectively called vector addition and scalar multiplication. Any system wherein there are operations similar to the operations just considered and which satisfies all the properties just mentioned is also called vector space. i.e. Any set V on which there are operations similar to those in $\mathbb{R}^n$ that satisfies the properties (i) to (ix) is called a vector space. And elements of V in this case are called vectors.

3. Geometric Representation

Given a point $x \in \mathbb{R}^n$, often we look at x as a directed line segment from origin towards the point. When two vectors x and $y \in \mathbb{R}^n$ are given, one can show that the sum defined earlier is the same as finding $x + y$, by completing the parallelogram. That means that at the tail of x put a directed line segment which is parallel to y and is of same length as the length of y; this is called a parallel translate of y. All these are shown in the

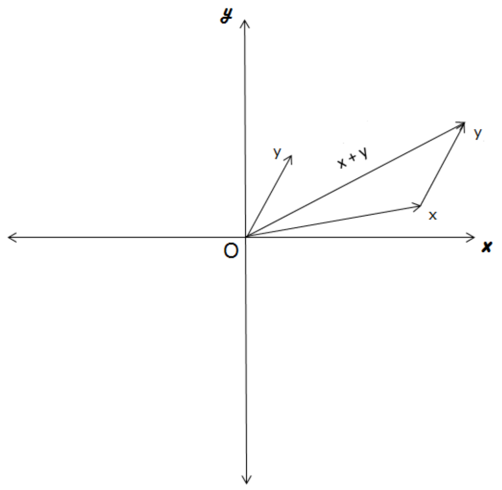
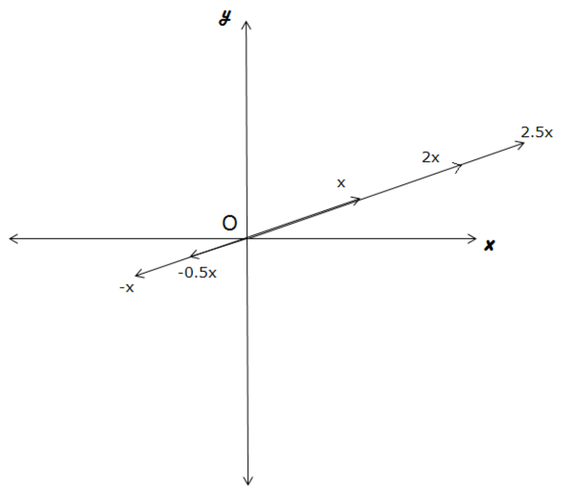


figure.

The scalar multiplication αx is interpreted as the vector which is in the same direction or the reverse direction depending on α being positive or negative. That is if α is 2 then $2x$ is the vector twice the size of vector x, which is in the same direction as of x. This is illustrated in the figure.



4. Linear Combination, Span and Basis

Almost all the notions that we define makes sense in a more general setting of Vector space. However we will not go into this generality but restrict ourselves to only the most concrete vector space $\mathbb{R}^n$.

If $v^1, v^2, \ldots, v^m$ are elements of a vector space $\mathbb{R}^n$, a vector

$$v = \alpha_1 v^1 + \alpha_2 v^2 + \cdots + \alpha_m v^m$$

where $\alpha_1, \alpha_2, \dots \alpha_m$ *are scalars* that is they are real numbers, is called a linear combination of vectors $v^1, v^2, \dots, v^m$.

If we take all α_i , $i = 1, \dots m$, to be zero; we can express zero vector as a linear combination of any collection of vectors. This is called a trivial linear combination.

If $S \subseteq \mathbb{R}^n$ is a given set, its linear span or simply span denoted by $[S]$ is defined as the set of all the linear combinations of all the elements of S. i.e.

$[S] = \{\alpha_1 v^1 + \alpha_2 v^2 + \cdots + \alpha_m v^m \mid v^1, v^2, \dots, v^m \in S \; and \; \alpha_1, \alpha_2, \dots \alpha_m \in \mathbb{R}\}$

Examples:

1. If $S = \{e^1 = (1,0)\}$ is the singleton set, $[S] = x - axis$. This should be clear from: $(x, 0) = x(1,0) = xe^1$
2. If $T = \{e^2 = (0,1)\}$ is the singleton set, $[T] = y - axis$. Which follows from: $(0, y) = y(0,1) = ye^2$
3. If $U = \{e^1 = (1,0), e^2 = (0,1)\}$ is the doubleton set, $[U] = \mathbb{R}^2$. This follows from the simple observation:

$$(x, y) = x(1,0) + y(0,1) = xe^1 + ye^2$$

Linearly Independent set:

A set $\{v^1, v^2, \dots, v^m\}$ is called linearly independent set if zero can be expressed only as a trivial linear combination of these vectors. In other words no non-trivial linear combination of these vectors is the zero vector. That is,

If $\alpha_1 v^1 + \alpha_2 v^2 + \cdots + \alpha_m v^m = 0$ then $\alpha_1 = \alpha_2 = \dots = \alpha_m = 0$

A set which is not linearly independent is called linearly dependent. That is in this case some non-trivial linear combination of the given vectors is equal to a zero vector.

Examples:

1. $S = \{e^1 = (1,0), e^2 = (0,1)\}$ is a linearly independent set, as $\alpha_1 e^1 + \alpha_2 e^2 = (0,0)$ implies $(\alpha_1, \alpha_2) = (0,0)$ which implies $\alpha_1 = 0 \; and \; \alpha_2 = 0$.
2. $S = \{v^1 = (1,1), v^2 = (2,2)\}$ is linearly dependent as $2v^1 - 1\,v^2 = (0,0)$. That is a non-trivial linear combination of vectors from S is a zero vector.

Basis:

A subset of $\mathbb{R}^n$ is said to a basis for $\mathbb{R}^n$ if it is a linearly independent set and it also spans the whole space $\mathbb{R}^n$.

Example:

1. $U = \{e^1 = (1,0), e^2 = (0,1)\}$ is a basis of $\mathbb{R}^2$.
2. $S = \{e^1 = (1,0,0), e^2 = (0,1,0)\}$ is not a basis of $\mathbb{R}^3$, as $[S]$ will only be the xy-plane in $\mathbb{R}^3$ and not $\mathbb{R}^3$.

Scalar Product or Inner Product and Orthogonality:

Suppose$x = (x_1, x_2, \ldots x_n) \; and \; y = (y_1, y_2, \ldots y_n) \in \mathbb{R}^n$. Then scalar product of $x \; and \; y$ which is denoted by $x \cdot y$ or $<x, y>$ is defined as:

$$x \cdot y = < x, y > = x_1 y_1 + x_2 y_2 + \ldots + x_n y_n$$

Two vectors $x, y \in \mathbb{R}^n$ are said to be orthogonal if $x \cdot y = 0$.

Example:

$e^1 = (1,0,0) \; and \; e^2 = (0,1,0)$ are orthogonal vectors. As

$$e^1 \cdot e^2 = < e^1, e^2 > = 1 \times 0 + 0 \times 1 + 0 \times 0 = 0$$

We note that this coincides with our usual notions.

5. Matrices

We have considered earlier, demand and supply functions of single commodity. In practice, generally demand and supply of any commodity, depends on demand and supply of several other commodities.

Example: Suppose we have two commodities whose demand and supply are given by:

For the first commodity:

Suppose the demand of the first commodity as a linear function of prices of two commodities is described by:

$Q_{d1} = a_2 P_2 + a_1 P_1 + a_0$

Suppose the supply of the first commodity as a linear function of prices of two commodities is described by:

$Q_{s1} = b_2 P_2 + b_1 P_1 + b_0$

Similarly the demand and supply of the second commodity as linear functions of prices of two commodities are described by:

$Q_{d2} = \alpha_2 P_2 + \alpha_1 P_1 + \alpha_0$

$Q_{s2} = \beta_2 P_2 + \beta_1 P_1 + \beta_0$

As we know absence of Excess demands for the two commodities provides market equilibrium.

i.e. $E_i \equiv Q_{di} - Q_{si} = 0 \quad i = 1,2.$

Hence algebraic simplification leads to equations:

$$(a_2 - b_2)P_2 + (a_1 - b_1)P_1 + (a_0 - b_0) = 0$$

$(\alpha_2 - \beta_2)P_2 + (\alpha_1 - \beta_1)P_1 + (\alpha_0 - \beta_0) = 0$

To get the elegant form we introduce new constants

$$c_i \equiv (a_i - b_i) \quad i = 0,1,2.$$

$$\gamma_i \equiv (\alpha_i - \beta_i) \quad i = 0,1,2.$$

On substitution we get:

$$c_2 P_2 + c_1 P_1 = -c_0$$

$$\gamma_2 P_2 + \gamma_1 P_1 = -\gamma_0$$

Which are called two simultaneous equations in two variables P_1, P_2.

In the standard variables, two simultaneous equations in two variables are written as:

$ax + by = c$

$\alpha x + \beta y = \gamma$

When there are m equations in n variables they are written as:

$\boldsymbol{a_{11}x_1 + a_{12}x_2 + \cdots a_{1n}x_n = b_1}$

$\boldsymbol{a_{21}x_1 + a_{22}x_2 + \cdots a_{2n}x_n = b_2}$

$\vdots \quad \vdots \qquad \vdots \quad \vdots \qquad\quad \vdots \quad \vdots$

$\boldsymbol{a_{m1}x_1 + a_{m2}x_2 + \cdots a_{mn}x_n = b_m}$

Now we are in a position to define a Matrix with proper motivation.

If we look at these equations closely we immediately understand the three types of entities are involved. First there are coefficients a_{ij} which are mn in numbers. Secondly there are n variables $x_1, x_2, \ldots x_n$ and m numbers $b_1, b_2, \ldots b_m$. To manipulate these entities efficiently Matrices are at our service.

So we formally define matrix as simply an array of entities usually of numbers and variables. A general matrix of m rows and n columns is written as:

$$A = [a_{ij}] = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ldots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix} \quad i = 1,2 \ldots m \text{ and } j = 1,2, \ldots n$$

Often instead of square brackets, round brackets are also used.

We also note as in the case of vectors that two matrices are equal iff their corresponding entries are same. That is if $A = [a_{ij}]$ $and\, B = [b_{ij}]$, are two $m \times n$ matrices, which is read as "m by n" matrices, i.e. they are of the same type then,

$$A = B \text{ if and only if } a_{ij} = b_{ij} \text{ for all } i = 1,2 \ldots m \text{ and } j = 1,2, \ldots n$$

When A and B are two matrices with real numbers or complex numbers as entries and of the same type they have m rows and n columns that is they are $m \times n$, then they are added as:

$$A + B = [a_{ij}] + [b_{ij}] = [a_{ij} + b_{ij}]$$

That is i^th^ row, j^th^ column element of the first matrix is added to the i^th^ row, j^th^ column of the second matrix to form the i^th^ row, j^th^ column entry of the matrix $A + B$.

Scalar Multiplication:

If a matrix $A=[a_{ij}]$ and $\alpha \in \mathbb{R}$ are given, we define scalar multiplication as:

$$\alpha A = [\alpha a_{ij}]$$

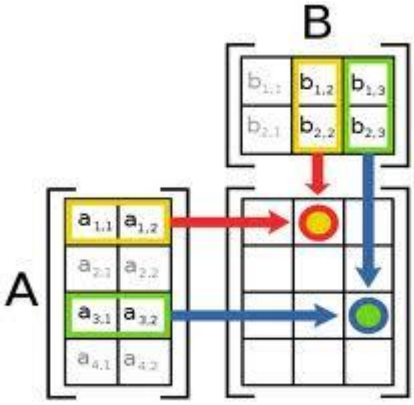
which means that α is multiplied to every entry of the matrix A.

Multiplication of Two matrices:

Suppose we are given two matrices A and B. The first matrix A is $m \times k$ and the second matrix B is $k \times n$. Then the product AB is the $m \times n$ matrix whose general term c_{ij} is formed by multiplying the corresponding terms of the i^th^ row of A and j^th^ column of B both of which are having k entries and then summing up these multiplied term. That is,

$$\boldsymbol{C = AB = [c_{ij}] \quad where\, c_{ij} = \sum_{l=1}^{k} a_{il} b_{lj}}$$

A special case can be visualized as:



Zero Matrix And Identity Matrix:

A matrix $A=[a_{ij}]$ of the type $m \times n$, all of whose entries are zero, is called a zero matrix.

A matrix $A=[a_{ij}]$ of the type $n \times n$, that is a matrix with equal number of rows and columns is called a square matrix.

For a square matrix $A=[a_{ij}]$ of the type $n \times n$, the terms a_{ii}, $i = 1,2,\dots n$; are called main diagonal terms. If in a square matrix all terms on the main diagonal are identity i.e. 1 and off the main diagonal all the terms are zero, then the matrix is called the Identity matrix, this matrix is usually denoted by I. That is

$$I = [a_{ij}] = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix} \quad i = 1,2\dots n \; and \; j = 1,2,\dots n$$

which can also be written as:

$I = [a_{ij}]$ where $a_{ii} = 1$, $i = 1,\dots n$ and $a_{ij} = 0 \quad i \neq j, \; i,j = 1,\dots n$.

This matrix has interesting property viz.

$$AI = IA = A$$

This follows as, the ith row, jth column term in AI is formed by summing the corresponding multiplied terms from the ith row of A and jth column of I and the fact that jth column of I has all terms zero except the jth term. Thus the only surviving term in the sum will be a_{ij}. Hence $AI = A$. Similarly $IA = A$.

Transpose of a Matrix:

Transpose of a $m \times n$ matrix $A = [a_{ij}]$ is a $n \times m$ matrix $A^T = [a_{ji}]$ that means jth row, ith column entry in A^T is the ith row, jth column entry of A. In other words the rows of A are made columns of A^T

6. Elementary Row/Column operations

The matrix operations of

1. Interchanging two rows or columns
2. Adding a multiple of one row or column to another row or column
3. Multiplying any row or column by a nonzero element.

are known as elementary row/ column operations.

These operations are important. They are used to transform the original Matrix into a matrix having definite simple form. Also certain properties of the original Matrix and the transformed matrix remain invariant or are related. If the transformed matrix has simple desired form and if the definite entity can be found out for it, we have got the entity for the original matrix also. This is the main use of these operations.

7. Row and Column Vectors

Given a $m \times n$ matrix $A = [a_{ij}]$, it can be looked upon as made up of m many row vectors $(a_{i1}, a_{i2}, \ldots, a_{in})$, $i = 1, \ldots, m$ which are elements of $\mathbb{R}^n$, these row vectors may also be called n-vectors. Similarly A can also be looked upon as if it is made up of n many column vectors $(a_{1j}, a_{2j}, \ldots, a_{mj})$, $j = 1, \ldots, n$ which are elements of $\mathbb{R}^m$. These column vectors may be called m-vectors. It is convenient and very natural to visualize these vectors as:

$$\begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ \\ \vdots \\ \\ a_{mj} \end{bmatrix}$$

These vectors can also be regarded as appropriate matrices. Row vectors mentioned above are $1 \times n$ matrices. And column vectors are $m \times 1$ matrices.

Now we go back to simultaneous equations:

Note that now we can write the equations:

$$\begin{aligned}
&a_{11}x_1 + a_{12}x_2 + \cdots a_{1n}x_n = b_1\\
&a_{21}x_1 + a_{22}x_2 + \cdots a_{2n}x_n = b_2\\
&\vdots \quad \vdots \quad\quad \vdots \quad \vdots \quad\quad\quad \vdots \quad \vdots\\
&a_{m1}x_1 + a_{m2}x_2 + \cdots a_{mn}x_n = b_m
\end{aligned}$$

as

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \dots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ \vdots \\ b_m \end{bmatrix}$$

which is same as:

$$Ax = b \text{ where } A = [a_{ij}],\ x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ \vdots \\ x_n \end{bmatrix} \text{ and } b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ \vdots \\ b_m \end{bmatrix}.$$

Soon we will use techniques of matrix algebra and discuss about the solution of simultaneous equations.

8. Symmetric, orthogonal and idempotent matrices

A $n \times n$ square matrix $A = [a_{ij}]$ is called symmetric if $a_{ij} = a_{ji}$ for all $i, j = 1,2 \dots n$. That is A is symmetric if $A = A^T$.

Examples:

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

are symmetric matrices.

A $n \times n$ square matrix $A = [a_{ij}]$ is called an orthogonal matrix if $AA^T = I$.

Example:

$$A=\begin{bmatrix}\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\end{bmatrix}$$

is an orthogonal matrix as:

$$\boldsymbol{AA^T}=\begin{bmatrix}\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\end{bmatrix}\begin{bmatrix}\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\end{bmatrix}$$

$$=\begin{bmatrix}\frac{1}{\sqrt{2}}\frac{1}{\sqrt{2}}+\left(-\frac{1}{\sqrt{2}}\right)\left(-\frac{1}{\sqrt{2}}\right) & \frac{1}{\sqrt{2}}\frac{1}{\sqrt{2}}+\left(-\frac{1}{\sqrt{2}}\right)\left(\frac{1}{\sqrt{2}}\right)\\ \frac{1}{\sqrt{2}}\frac{1}{\sqrt{2}}+\left(\frac{1}{\sqrt{2}}\right)\left(-\frac{1}{\sqrt{2}}\right) & \frac{1}{\sqrt{2}}\frac{1}{\sqrt{2}}+\left(\frac{1}{\sqrt{2}}\right)\left(\frac{1}{\sqrt{2}}\right)\end{bmatrix}$$

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

The length, often called norm, of a vector $x=(x_1,x_2,\dots x_n)\in\mathbb{R}^n$ is defined as:

$$|x|=|(x_1,x_2,\dots x_n)|=\sqrt{{x_1}^2+{x_2}^2\dots+{x_n}^2}$$

Which is the distance of the point x from origin of $\mathbb{R}^n$.

We note that matrix A is an orthogonal matrix iff its column vectors are unit vectors i.e. their lengths are 1 and any two distinct column vectors are orthogonal. Similarly matrix A is an orthogonal matrix iff its row vectors are unit vectors i.e. their lengths are 1 and any two distinct row vectors are orthogonal.

A $n\times n$ square matrix $\boldsymbol{A}=[\boldsymbol{a_{ij}}]$ is called idempotent if $\boldsymbol{A^2}=\boldsymbol{A}$**.**

Example:

$$A=\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}$$

$$A^2=\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}$$

$$= \begin{bmatrix} 1+0 & 0+0 \\ 0+0 & 0+0 \end{bmatrix}$$

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

Thus A is an idempotent matrix. Of course zero matrices and identity matrices are idempotent matrices. We shall look at some more aspects of matrices and their applications in the next module.

9. Summary

We here started with the definitions of addition of two elements of $\mathbb{R}^n$ and also of scalar multiplication. We then listed down important properties of these operations, which actually form the basis of an abstract i.e. axiomatic definition of Vector Space or Linear Space. We have also conveyed geometric representation of vectors and the interpretation of vector addition and scalar multiplication. Then after, we covered with examples, the notions of linear combination, linear independence, linear dependence, linear span and basis. We also discussed Scalar product and orthogonality. Lastly we introduced Matrices with motivation from simultaneous equations. Operations of addition, multiplication were discussed with care. We have also discussed symmetric, orthogonal and idempotent matrices.