1 Preliminary Concepts

1.1 INTRODUCTION

This chapter provides a brief review of all basic concepts, definitions, theorems, etc. which will be used in the subsequent chapters. Various theorems and results have been stated only without giving their proofs as the same may be looked up in any standard textbook on Algebra. It has been assumed that the reader is familiar with the elementary set theory.

1.2 OUTLINE

Here is a listing and brief description of the material in this set of notes.

Systems of Equation and Matrices

Systems of equations In this section, we will introduce most of the basic topics that we will need in order to solve systems of equations including augmented matrices and row operations.

Solving systems of equations Here we will look at the Gaussian elimination and Gauss-Jordan method of solving systems of equations.

Matrices We will introduce many of the basic ideas and properties involved in the study of matrices.

Matrix arithmetic and operations In this section, we will take a look at matrix addition, subtraction and multiplication and also take a quick look at the transpose and trace of a matrix.

Properties of matrix arithmetic We will take a more in-depth look at many of the properties of matrix arithmetic and the transpose.

Inverse and elementary matrices Here we will define the inverse and take a look at some of its properties and also introduce the idea of elementary matrices.

Finding inverse matrices In this section, we will develop a method for finding inverse matrices.

Special matrices We will introduce diagonal, triangular and symmetric matrices in this section.

LU-decompositions In this section, we will introduce the LU-decomposition, a way of "factoring" certain kinds of matrices.

Systems revisited Here we will revisit solving systems of equations. We will take a look at how inverse matrices and LU-decompositions can help with the solution process. We will also take a look at a couple of other ideas in the solution of systems of equations.

Determinants

The determinant function We will give the formal definition of the determinant in this section. We will also give formulae for computing determinants of 2×2 and 3×3 matrices.

Properties of determinants Here we will take a look at quite a few properties of the determinant function including formulae for determinants of triangular matrices.

The method of cofactors In this section, we will take a look at the first of two methods for computing determinants of general matrices.

Using row reduction to find determinants Here we will take a look at the second method for computing determinants in general.

Cramer's rule We will take a look at yet another method for solving systems. This method will involve the use of determinants.

Euclidean *n*-Space

Vectors In this section, we will introduce vectors in 2-space and 3-space as well as some of the important ideas about them.

Dot product and cross product Here we will look at the dot product and the cross product, two important products of vectors. We will also take a look at an application of the dot product.

Euclidean n-space We will introduce the idea of Euclidean n-space in this section and extend many of the ideas of the previous two sections.

Linear transformations In this section, we will introduce the topic of linear transformations and look at many of their properties.

Examples of linear transformations We will take a look at quite a few examples of linear transformations in this section.

Vector Spaces

Vector spaces In this section, we will formally define vectors and vector spaces.

Subspaces Here we will be looking at vector spaces that live inside of other vector spaces.

Span The concept of the span of a set of vectors will be investigated in this section.

Linear independence Here we will take a look at what it means for a set of vectors to be linearly independent or linearly dependent.

Basis and dimension We will be looking at the idea of a set of basis vectors and the dimension of a vector space.

Change of basis In this section, we will see how to change the set of basis vectors for a vector space.

Fundamental subspaces Here we will take a look at some of the fundamental subspaces of a matrix, including the row space, column space and null space.

Inner product spaces We will be looking at a special kind of vector spaces as well as define the inner product.

Orthonormal basis In this section, we will develop and use the Gram-Schmidt process for constructing an orthogonal/orthonormal basis for an inner product space.

Least squares In this section, we will take a look at an application of some of the ideas that we will be discussing in this chapter.

R-decomposition Here we will take a look at the *QR*-decomposition for a matrix and how it can be used in the least squares process.

Orthogonal matrices In this section, will take a look at a special kind of matrix, i.e. the orthogonal matrix.

Eigenvalues and Eigenvectors

Review of determinants In this section, we will do a quick review of determinants.

Eigenvalues and eigenvectors Here, we will take a look at the main section of this chapter. We will be looking at the concept of eigenvalues and eigenvectors.

Diagonalization We will be looking at diagonalizable matrices in this section.

1.3 CARTESIAN PRODUCT OF SETS AND RELATIONS

Cartesian Product of Sets

Let A and B be any two non-empty sets. The set of all ordered pairs (a, b) such that $a \in A$ and $b \in B$ is called the cartesian product of the sets A and B and is denoted by $A \times B$, that is,

$$A \times B = \{(a, b) : a \in A \text{ and } b \in B\}$$

If $A = \phi$ or $B = \phi$, then we define $A \times B = \phi$

Relation

Let A and B be two sets. A relation R from set A to set B is a subset of $A \times B$.

If R is a relation from a non-void set A to a non-void set B and if $(a, b) \in R$, then we write $a\ R\ b$ which is read as 'a is related to b by the relation R'. If $(a, b) \notin R$, then $(a, b) \in R^1$.

A relation from a set A to itself is called a relation on set A.

Inverse of a relation Let A and B be two sets and let R be a relation from a set A to a set B. Then the inverse of R denoted by R^{-1} is a relation from B to A and is defined by

$$R^{-1} = \{(b, a) : b \in B,\ a \in A\}$$

Clearly, $(a,b) \in R \Leftrightarrow (b,a) \in R^{-1}$

Identity relation The relation I_A on a set is identity relation if every element of A is related to itself only, that is

$$I_A = \{(a,a) : a \in A\}$$

Reflexive relation A relation R on set A is said to be reflexive if every element of A is related to itself.

Thus, R is reflexive on set A iff $(a, a) \in R$ for all $a \in A$.

Symmetric relation A relation R on set A is said to be a symmetric relation iff $(a,b) \in R \Rightarrow (b,a) \in R$ for all $a,b \in A$.

Transitive relation A relation R on a set A is said to be a transitive relation iff

$$(a,b) \in R \text{ and } (b,c) \in R \Rightarrow (a,c) \in R \text{ for all } a, b, c \in A.$$

Equivalence relation A relation R on a set A is said to be an equivalence relation iff R is :

(i) Reflexive, i.e. $(a, a) \in R$ for all $a \in A$.

(ii) Symmetric, i.e. $(a, b) \in R \Rightarrow (b, a) \in R$ for all $a, b \in A$.

If R is an equivalence relation on a non-empty set A and $a \in A$, then the set of all those elements of A which are related to a by the relation R is called the equivalence class determined by a and is denoted by $[a]$.

Thus, $[a] = \{x \in A : (x,a) \in R\}$

For any $a, b \in A$

(i) if $b \in [a]$, then $[b] = [a]$

(ii) $[a] = [b]$ iff $(a, b) \in$ R

(iii) either $[a] = [b]$ or $[a] \cap [b] = \phi$

1.4 FUNCTIONS

Function as a Set of Ordered Pairs

Let A and B be two non-empty sets. A relation f from A to B, i.e. a subset of $A \times B$ is called a function (or a mapping or a map) from A to B if

(i) for each $a \in A$ there exists $b \in B$ such that $(a, b) \in f$

(ii) $(a, b) \in f$, then b is called the image of a under f

Function as a Correspondence

Let A and B be two non-empty sets. A function 'f' from set A to set B is a rule relating elements of set A to elements of set B such that

(i) all elements of set A are associated to elements in set B.

(ii) an element of set A is associated to a unique element in set B.

Terms such as 'map' (for "mapping"), "correspondence" are used as synonyms for "function".

If f is a function from a set A to a set B, then we write $f : A \to B$ or $A \xrightarrow{f} B$, which is read as 'f' is a function from A to B or f maps A to B. If an element $a \in A$ is associated to an element $b \in B$, then b is called "the f image of a" or "image of a under function" or "the value of the function f at a". Also, a is called the pre-image of b under the function f and we write $b = f(a)$ or $f^{-1}(b) = a$.

The set A is known as the domain of f and the set B is known as the co-domain of f. The set of all f-images of elements of A is known as the range of f or image set of A under f and is denoted by $f(A)$.

Thus, $f(A) = \{f(x) : x \in A\}$ = range of f.

The set of all functions or mappings from a set X to a set A is denoted by A^X.

Remark: A rule relating all elements of set A to elements of set B is a function or a well-defined rule iff

$$x = y \Rightarrow f(x) = f(y) \text{ for all } x, y \in A$$

Injective Map

A function $f : A \to B$ is said to be an injective map or a one-one function if distinct elements in A have distinct images in B.

Thus, f is injective if and only if

$$x \neq y \Rightarrow f(x) \neq f(y) \text{ for all } x, y \in A$$

$$\Leftrightarrow \quad f(x) = f(y) \Rightarrow x = y \text{ for all } x, y \in A$$

If $f : A \to B$ is not a one-one function, then it is said to be a many-one

function, that is, a function $f : A \to B$ is said to be a many-one function if two or more elements of set A have the same image in B.

Surjective Map

A function $f : A \to B$ is said to be an onto function or a surjective map if every element of B is the f image of some element of A, i.e. if $f(A) = B$ or range of f is the co-domain of f or $l_m(f) = B$.

$f : A \to B$ is said to be a bijective map if it is injective as well as surjective.

Bijective Map

A function $f : A \to B$ is said to be a bijective map if it is injective as well as surjective.

In other words, a function $f : A \to B$ is a bijective map or a bijection if

(i) it is injective, i.e. $f(x) = f(y) \Rightarrow x = y$ for all $x, y \in A$

(ii) it is surjective, i.e. for all $y \in A$ there exists $x \in A$ such that $f(x) = y$

Permutation

Let A be a non-empty set. A bijective map from A to itself is called a permutation on A.

If A is a finite set equal to $\{a_1, a_2, ..., a_n\}$, then a permutation f on A is written in two row notation as follows.

$$f = \begin{pmatrix} a_1 & a_2 & a_3 & \dots & a_n \\ f(a_1) & f(a_2) & f(a_3) & \dots & f(a_n) \end{pmatrix}$$

Inverse of a Function

Let $f : A \to B$ be a bijection. Then a function $g : B \to A$ which associate each element $y \in B$ to a unique element $x \in A$ such that $f(x) = y$ is called the inverse of f and is denoted by f^{-1}.

Thus, if $f : A \to B$ is bijection, then $f^{-1} : B \to A$ is such that

$$f(x) = y \Leftrightarrow f^{-1}(y) = x$$

If f has an inverse, then f is said to be invertible. The inverse of an invertible function is unique.

Clearly f is invertible iff f is a bijection.

Composition of Functions

Let $f : A \to B$ and $g : C \to D$ be two functions such that $B \subset C$ or range $f \subset C$, then the composite h of f and g, denoted by gof, is the

mapping $h : A \to D$ defined by

$$h(x) = g(f(x)) \text{ for each } x \in A.$$

The composition of functions is not necessarily commutative, i.e. $gof \neq fog$. But it is always associative, i.e. for any three functions f, g, h, we have

$(fog)oh = fo(goh)$ provided that fog and goh are defined.

Also, if $f : A \to B$, then

$$foI_B = f = I_A of$$

where I_A and I_B are identity functions on A and B respectively.

If f and g are both invertible functions such that gof is defined, then gof is also invertible and $(gof)^{-1} = f^{-1}og^{-1}$.

Following are some useful results on composition of functions:

(i) If $f : A \to B$ and $g : B \to A$ are two functions such that $gof = I_A$, then f is an injection and g is a surjection.

(ii) If $f : A \to B$ and $g : B \to A$ are two functions such that $fog = I_B$, then f is a surjection and g is an injection.

(iii) If $f : A \to B$ and $g : B \to C$ be two functions, then

(a) g of $A \to C$ is onto $\Rightarrow g : B \to C$ is onto

(b) g of $A \to C$ is one-one $\Rightarrow f : A \to B$ is one-one

(c) g of $A \to C$ is onto and $g : B \to C$ is one-one $\Rightarrow f : A \to B$ is onto

(d) g of $A \to C$ is one and $f : A \to B$ is onto $\Rightarrow g : B \to C$ is one-one

List: Let n denote the set of first n natural numbers, i.e. $N = \{1, 2,..., n\}$. Then a function $f : n \to A$ is called a list of elements in A and is written $(f_1, f_2, f_3,..., f_n)$.

We use the notation n to denote the list $(1, 2, 3,..., n)$.

1.5 BINARY OPERATIONS

Binary Operations: Let S be a non-empty set. A function $f : S \times S \to S$ is called a binary operation (or a binary composition) on this set S.

Thus, a binary operation f on a set S associates each ordered pair (a, b) in S. We shall use the notation of $f\,b$ instead of $f(a, b)$ for a binary operation f on a set S. Generally, binary operations are denoted by the symbols like $*$, $\circ$, $\odot$, $\oplus$, etc. Thus, if $*$ is a binary operation on a set S, then image of an element $(a, b) \in S \times S$ is written as $a * b$ (instead of the usual notation $*(a, b)$).

Addition $(+)$ and multiplication $(\cdot)$ are binary operations on N but subtraction

and division are not binary operations on N. Subtraction is a binary operation on each of sets Z, Q, R and C.

Commutative binary operation A binary operation $*$ on a set S is said to be commutative if

$$a * b = b * a \text{ for all } a, b \in S$$

Addition on R is commutative but subtraction is not commutative.

Associative binary operation A binary operation $*$ on a set S is said to be associative if

$$(a * b) * c = a * (b * c) \text{ for all } a, b, c \in S$$

Distributivity Let S be a non-void set and $*$ and $\odot$ be two binary operations on S. The binary operation $*$ is said to be

(i) left distributive over $\odot$ $(b \odot c) * a = (b * a) \odot (c * a)$

(ii) Right distributive over $\odot$ $a * (b \odot c) = (a * b) \odot (a * c)$ for all $a, b, c \in S$

The binary operation $*$ is said to be distributive over $\odot$ if it is both left as well as right distributive.

Closure property Let $*$ be a binary operation on a set S. A subset T of S is said to be closed under $*$ if $a * b \in T$ for all $a, b \in T$.

Clearly, S is closed under $*$ by the definition.

Restriction of a binary operation Let S and T be two sets such that $T \subset S$. A binary operation $*$ on T is said to be the restriction of a binary operation $\odot$ on S if $a * b \in T$ for all $a, b \in T$, that is, $*$ and $\odot$ are equal on T.

If $*$ is restriction of $\odot$ on S, then we also say that the binary operation $*$ is induced by $\odot$ on S.

Usually, we use the same symbol for the binary operation $\odot$ and its restriction $*$ on T.

Addition on Z is restriction of addition on R. Similarly, multiplication on R is restriction of multiplication on C.

Left identity Let $*$ be a binary operation on a set S. An element $e_1 \in S$ is called a left identity if

$$e_1 * a = a \text{ for all } a \in S$$

Right identity Let $*$ be a binary operation on a set S. An element $e_2 \in S$ is called a right identity if

$$a * e_2 = a \text{ for all } a \in S$$

Identity element Let $*$ be a binary operation on a set S. An element $e \in S$ is called identity element if it is both a left identity and a right identity, i.e. $e * a = a = a * e$ for all $a \in S$

The identity element for a binary operation $*$ on a set S, if it exists, is unique.

Left inverse Let $*$ be a binary operation on a set S and $e \in S$ be the identity element for $*$ on S. An element b is a left inverse of $a \in S$ if

$$b * a = e$$

Right inverse Let $*$ be a binary operation on a set S and $e \in S$ be the identity element for $*$ on S. An element c is a right inverse of $a \in S$ if

$$a * c = e$$

Inverse of an element Let $*$ be a binary operation on a set S and $e \in S$ be the identity element for $*$ on S. An element x is an inverse of an element $a \in S$ if x is both left inverse as well as right inverse of a, i.e.

$$x * a = e = a * x$$

The inverse of a is usually denoted by a^{-1}. For additive binary operation on a set S, the inverse of a is denoted by $-a$.

An element $a \in S$ is said to be invertible, if it possesses its inverse. The inverse of an invertible element is unique. The identity element is always invertible and is inverse of itself.

Algebraic Structure

A non-empty set S equipped with one or more binary operations on it is called an algebraic structure.

1.6 GROUPS

Semigroup An algebraic structure $(G, *)$ consisting of a non-void set G and a binary operation $*$ defined on G is called a semigroup if it satisfies the following axiom.

SG-1 Associativity The binary operation $*$ is associative on G

i.e. $(a * b) * c = a * (b * c)$ for all $a, b, c \in G$

The algebraic structures $(N, +)$, $(Q, +)$ $(R, +)$, $(C, +)$ $(Z, +)$, $(Q, +)$, etc. are semigroups.

Let $P(S)$ be the power set of a set S. Then, $(P(S), \cup)$ and $(P(S), \cap)$ are semigroups.

Monoid: An algebraic structure $(G, *)$ consisting of a non-void set G and a binary operation $*$ defined on G is called a monoid if it satisfies the following axioms.

M-1 Associativity The binary operation $*$ is associative on G

i.e. $(a * b) * c = a * (b * c)$ for all $a, b, c \in G$.

M-2 Existence of identity There exists a unique element $e \in G$ such that

$$a * e = a = e * a \text{ for all } a \in G.$$

The algebraic structures $(N, \times)$, $(Z, +)$, $(Q, \times)$ are monoids but $(N, +)$ is not a monoid.

Group: An algebraic structure $(G, *)$ consisting of a non-void set G and a binary operation $*$ defined on G is called a group if it satisfies the following axioms:

G-1 Associativity The binary operation $*$ is associative on G,

i.e. $(a * b) * c = a * (b * c)$ for all $a, b, c \in G$

G-2 Existence of Identity There exists an element $e \in G$ such that

$$a * e = a = e * a \text{ for all } a \in G$$

G-3 Existence of Inverse For each $a \in G$ there exists an element $a' \in G$ such that

$$a * a' = e = a' * a$$

The element a' is called the inverse of a and is denoted by a^{-1}.

The algebraic structures $(Z, +)$, $(Q, +)$, $(R, +)$, $(C, +)$ and $(Q, \times)$ are groups.

When it is not necessary to indicate the binary operation $*$, the group $(G, *)$ is simply referred to as the group G. If the binary operation on a group G is addition, it is usually called an additive group, its identity element is called zero, written as 0, and the inverse of a is called the negative of a, written as $-a$. When the binary operation on a group G is multiplication, then G is called a multiplicative group and $a \times b$ is written as ab. The identity element is usually denoted by 1 and the inverse of an element a is written as $\frac{1}{a}$.

Abelian Group

A group $(G, *)$ is called an abelian group if $*$ is commutative on G,

i.e. $a * b = b * a$ for all $a, b \in G$.

$(Z, +), (Q, \times), (R, +), (C, +)$, etc. are abelian groups.

Following are some useful properties of groups.

(i) The identity element in a group $(G, *)$ is unique.

(ii) The inverse of every element of a group $(G, *)$ is unique.

(iii) The inverse of identity element in a group is the identity element itself.

(iv) In a group $(G, *)$, $(a * b)^{-1} = b^{-1} * a^{-1}$ for all $a, b \in G$.

(v) Let $(G,*)$ be a group. Then for all $a,b,c \in G$

$a * b = a * c \Rightarrow b = c$ (Left cancellation law)

and $b * a = c * a \Rightarrow b = c$ (Right cancellation law)

(vi) In a group $(G,*)$, $(a^{-1})^{-1} = a$, $a \in G$

(vii) Let $(G,*)$ be a group. Then for any $a,b \in G$, the equation $a * x = b$ and $y * a = b$ have unique solutions in G.

Cyclic Group

A group $(G,*)$ is said to be cyclic, if there exists an element $a \in G$ such that every element of G is expressible as some integral power of a.

The element $a \in G$ is called the generator of G and we write $G = |a|$.

For example, $(Z,+)$ is a cyclic group generated by 1.

Following are some useful properties of a cyclic group:

(i) Every cyclic group is abelian but an abelian group need not be cyclic.

(ii) If a is generator of a cyclic group, then a^{-1} is also a generator of G.

(iii) The order of a cyclic group is the same as the order of its generator.

(iv) Every infinite cyclic group has two and only two generators.

Subgroup

Let $(G,*)$ be a group and H be a non-void subset of G such that

(i) H is closed for the binary operation $*$ on G

(ii) H itself is group for the composition induced by that of G, i.e. H itself is a group under the restriction of $*$ on H. Then, we say that $(H,*)$ is the subgroup of $(G,*)$.

Trivially $\{e\}$ and G itself are subgroups of G.

For the sake of convenience, we simply say that H is a subgroup of G if $(H,*)$ is a subgroup of $(G,*)$.

Cosets: Let H be a subgroup of a group G and let $a \in G$. Then the sets

$$aH = \{ah : h \in H\} \text{ and } Ha = \{ha : h \in H\}$$

are known as left and right cosets respectively of H in G.

Obviously, $aH \subset G$ and $Ha \subset G$ for all $a \in G$.

If the binary operation on G is addition, then

$$a + H = \{a + h : h \in H\} \text{ and } \{h + a : h \in H\}$$

are respectively the left and right cosets of H in G.

Any two right (left) cosets of a subgroup H of group G are identical or disjoint.

If H is a subgroup of a group G and $a, b \in G$, then

$$Ha = Hb \Leftrightarrow ab^{-1} \in H$$

If H is a subgroup of a group G and $a, b \in G$, then

$$H + a = H + b \Leftrightarrow a - b \in H$$

Normal Subgroup

A subgroup N of a group G is said to be a normal subgroup of G if

$xax^{-1} \in N$ for all $x \in G$ and all $a \in N$

$\Rightarrow$ $xNx^{-1} \subset N$ for all $x \in G$

A subgroup N of a group G is a normal subgroup of G iff $xN = Nx$ for all $x \in G$.

This means that there is no distinction between left and right cosets of a normal subgroup of a group.

If N is a normal subgroup of a group G, then $NaNb = Nab$ for all $a, b \in G$

Quotient Group

Let N be a normal subgroup of a group G. Then the set $\frac{G}{N} = \{Nx : x \in G\}$ of all cosets of N in G is a group under the multiplication of cosets as a binary operation, i.e.

$$Na\,Nb = Nab \text{ for all } a, b \in G$$

This group is known as the quotient group or factor group of G by N.

1.7 RINGS AND FIELDS

Ring An algebraic structure $(R, +)$ consisting of a non-empty set R and two binary operations '+' and '×' on R is called a ring if the following axioms are satisfied.

Axiom-1: $(R, +)$ is an abelian group

Axiom-2: $(R, \cdot)$ is a semigroup

Axiom-3: '$\cdot$' is distributive over '+', i.e. for all $a, b \in G$

(i) $a \cdot (b + c) = a \cdot b + a \cdot c$

(ii) $(b + c) \cdot a = b \cdot a + c \cdot a$

Clearly, $(Z, +, \times)$, $(Q, +, \times)$ and $(R, +, \times)$ are rings.

Ring with unity A ring $(R, +, \cdot)$ is said to be a ring with unity if R has the identity element for multiplicative binary operation.

The identity element for multiplicative binary operation is denoted by 1.

Commutative ring A ring $(R, +, \cdot)$ is said to be a commutative ring if its multiactive binary operation is commutative,

i.e. $a \cdot b = b \cdot a$ for all $a, b \in R$

Following are some useful results in a ring $(R, +, \cdot)$:

(i) $a + 0 = 0 + a = 0$ for all $a \in R$, where 0 is the zero element, i.e. additive identity in R.

(ii) $a(-b) = -(ab) = (-a)b$ for all $a, b \in R$

(iii) $(-a)(-b) = ab$ for all $a, b \in R$

(iv) $a(b - c) = ab - ac$ for all $a, b, c \in R$

(v) $(b - c)a = ba - ca$ for all $a, b, c \in R$

Let $(R, +, \cdot)$ be a ring and n be a positive integer, then we define

$$na = a + a + a + ... + a \text{ (up to } n\text{-terms)}$$

Also, we define $0a = 0$, where 0 on the left-hand side is integer 0 and 0 on the right-hand side is the zero element (additive identity) of the ring.

Characteristic of a ring Let $(R, +, \cdot)$ be a ring with zero element 0. If there exists a positive integer n such that $na = 0$,

i.e. $a + a + a + ... + a$ (n times) $= 0$ (zero of the ring) for all $a \in R$. Then, we say that the ring is of finite characteristic. If n is the smallest positive integer such that $na = 0$ for all $a \in R$, then n is called the characteristic of ring R.

If there exists no positive integer n, such that $na = 0$ for all $a \in R$, then R is said to be of characteristic zero or infinite.

The ring $(Z, +, \times)$ is of characteristic zero, whereas $(Z_6, +_6, \times_6)$ is of characteristic 6.

Subring A non-void subset S of a ring $(R, +, \times)$ is a subring of R iff

(i) S is closed with respect to the binary operations of addition and multiplication on R

(ii) S itself is a ring for the induced binary operations.

The necessary and sufficient conditions for a non-void subset S of a ring R to be a subring of R are (i) $a - b \in S$ and (ii) $ab \in S$.

Field An algebraic structure $(F, +, \cdot)$ consisting of a non-void set F and two binary operations '+' and '·' on F is called a field if the following axioms are satisfied.

Axiom-1: $(F, +)$ is an abelian group

Axiom-2: $(F, \cdot)$ is an abelian group

Axiom-3: '·' is distributive over '+', i.e. for all $a, b, c \in G$

(i) $a \cdot (b + c) = a \cdot b + a \cdot c$

(ii) $(b + c) \cdot a = b \cdot a + c \cdot a$

A commutative ring with unity is a field, if its every non-zero element

has multiplicative inverse.

$(Q,+,\times),(R,+,\times)$ and $(E,+,\times)$ are fields.

Subfield A non-void subset K of a field $(F,+,\cdot)$ is a subfield of F iff

(i) K is closed under the binary operations on F.

(ii) K itself is a field for the induced binary operations.

The necessary and sufficient conditions for a non-void subset K of a field F to be a subfield of F are

(i) $ab = ba$ for all $a,b \in K$

(ii) $ab^{-1} \in K$ for all $a \neq b \in K$

1.8 MATRICES

Matrix A matrix over a field F or simply a matrix A (when F is implicit) is a rectangular arrangement of scalars.

If there are mn scalars $a_{ij} \in F$, where $i \in m$ and $j \in n$, then the following arrangement of these mn scalars is a matrix.

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{in} & \cdots & a_{mn} \end{bmatrix}$$

The element a_{ij} is called the ij-element or ij-entry which appears in ith row and jth column. Such a matrix is usually denoted by $A = [a_{ij}]_{m \times n}$ or simply $A = [a_{ij}]$.

A matrix with m rows and n columns is called an m by n matrix, written as $m \times n$.

The rows of matrix A (given above) are m horizontal lists of scalars as

$$(a_{11},a_{12},a_{13},\ldots,a_{1n}),\ (a_{21},a_{22},a_{23},\ldots,a_{2n}),\ (a_{m1},a_{m2},a_{m3},\ldots,a_{mn})$$

If we denote these rows by $A_1, A_2, A_3, \ldots, A_m$ respectively, then matrix A can also be written as a vertical list as given below.

$$A = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_m \end{bmatrix}$$

The columns of matrix A are n vertical lists of scalars $\begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{bmatrix}, \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{bmatrix} \cdots \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}$

These columns are generally denoted by $A^1, A^2, \ldots, A^n$ and the matrix A can be written as list of n columns as given below.

$$A = (A^1, A^2, \ldots, A^n)$$

Matrix as a mapping Let F be a field and m, n be positive integers. A mapping $A : m \times n \to F$ associating each ordered pair $(i, j) \in m \times n$ to the scalar $a_{ij} \in F$ is called an $m \times n$ matrix.

Clearly, a_{ij} is the image of $(i, j) \in m \times n$ under mapping A, i.e. $A(i, j) = a$ and is called the ij-entry or ij-element of matrix A. In such a case, the matrix A is also written as $A = [a_{ij}]$.

Column matrix A matrix with only one column is called a column matrix or a column vector.

As discussed above, a matrix A can be written as a row vector of its columns $A^1, A^2, \ldots, A^n$ and a column vector of its rows $A_1, A_2, \ldots, A_m$.

Square matrix An $n \times n$ matrix A is called a square matrix of order n.

If $A = [a_{ij}]$ is a square matrix of order n, then the elements $a_{11}, a_{22}, \ldots, a_{nn}$ are called the diagonal elements and the line along which they lie is called the principal diagonal or leading diagonal of the matrix. The diagonal elements are also written as a_{ii}, $i \in n$.

Diagonal matrix A square matrix $A = [a_{ij}]$ is called a diagonal matrix if all the elements except those in the leading diagonal are zero, i.e. $a_{ij} = 0$ for all $i \neq j$.

A diagonal matrix of order $n \times n$, having $d_1, d_2, \ldots, d_n$ as diagonal elements is denoted by diag $(d_1, d_2, \ldots, d_n)$.

For example, the diagonal matrix

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

is written as diag (1, 2, 3).

Scalar matrix A square matrix $A = [a_{ij}]$ is called a scalar matrix if

(i) $a_{ij} = 0$ for all $i \neq j$

(ii) $a_{ii} = c$ for all i, where $c \neq 0$.

Identity matrix A square matrix $A = [a_{ij}]_{n \times n}$ is called an identity matrix if

(i) $a_{ij} = 0$ for all $i \neq j$ and (ii) $a_{ii} = 1$ for all $i \in n$

An identity matrix of order $n \times n$ is generally denoted by I_n.

Null matrix A matrix whose all elements are zero is called a null matrix or a zero matrix.

Upper triangular matrix A square matrix $A = [a_{ij}]_{n \times n}$ is called an upper triangular matrix if $a_{ij} = 0$ for all $i \neq j$.

All elements below the leading diagonal of an upper triangular matrix are zero.

Lower triangular matrix A square matrix $A = [a_{ij}]_{n \times n}$ is a lower triangular matrix if $a_{ij} = 0$ for all $i < j$.

All elements above the leading diagonal of a lower triangular matrix are zero.

A triangular matrix $A = [a_{ij}]$ is called strictly triangular iff $a_n = 0$ for all $i \in n$.

Echelon matrix A matrix is called an echelon matrix or is said to be in echelon form if A is either the null matrix or satisfies the following conditions.

(i) All zero rows, if any, are at the bottom of the matrix. Here zero row means a row whose all entries are zeros.

(ii) The number of zeros before the first non-zero element in a row is less than the number of such zeros in the next row.

That is, $A = [a_{ij}]$ is an echelon matrix if there exist non-zero entries.

$$a_{j_1}, a_{j_2}, \ldots, a_{jr} \text{ where } j_1 < j_2 < \ldots < j_r$$

with the property that

$$a_{jj} = 0 \text{ for } \begin{cases} i \leq r \text{ and } j < j_i \\ i > r \text{ and } j < j_i \end{cases}$$

The elements a_{1j}, a_{2j}, a_{3j}, which are the leading non-zero ele-

ments in their respective rows are called the pivots of the echelon matrix.

The matrix A given by

$$A = \begin{bmatrix} 0 & \textcircled{3} & 2 & 1 \\ 0 & 0 & \textcircled{2} & 5 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

is an echelon matrix whose pivots have been encircled.

Clearly, each pivot is to the right of the one above.

The following matrix is also an echelon matrix whose pivots have been encircled.

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

Matrix in Row Canonical Form

A matrix A is said to be in row canonical form if it is an echelon matrix satisfying the following additional conditions.

(i) Each pivot is equal to 1.

(ii) Each pivot is the only non-zero entry of its column.

It should be noted that in echelon matrix there must be zeros below the pivots but in a matrix in row canonical form each pivot must be equal to 1 and there must also be zeros above the pivots.

The null matrix O and the identity matrix I (of any order) are examples of matrices in row canonical form.

The matrix A given by

$$A = \begin{bmatrix} 0 & \textcircled{1} & 2 & 0 & 0 & 4 \\ 0 & 0 & 0 & \textcircled{1} & 0 & 2 \\ 0 & 0 & 0 & 0 & \textcircled{1} & 3 \end{bmatrix}$$

is in row canonical form but the matrix B given on next page

$$B = \begin{bmatrix} \textcircled{-1} & 2 & 3 & 0 & 5 & -46 \\ 0 & 0 & \textcircled{1} & -2 & 3 & 0 \\ 0 & 0 & 0 & 0 & \textcircled{6} & 5 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

is not in row canonical form.

Equality of Matrices

Two matrices $A = \left[a_{ij}\right]_{m\times n}$ and $B = \left[b_{ij}\right]_{r\times s}$ are equal if (i) $m = r$ (ii) $n = s$ and (iii) $a_{ij} = b_{ij}$ for all i, j.

Scalar Multiplication

Let $A = \left[a_{ij}\right]$ be an $m \times n$ matrix and k be a scalar. Then the matrix obtained by multiplying every element of A by k is called the scalar multiple of A by k and it is denoted by kA, that is,

$$kA = \left[ka_{ij}\right]_{m\times n}$$

The negative of an $m \times n$ matrix $A = [a_{ij}]$, written as $-A$ is defined to be the $m \times n$ matrix given by $-A = [-a_{ij}]$ for all $i \in m, j \in n$.

Addition of Matrices

Let $A = [a_{ij}]$ and $B = [b_{ij}]$ be two $m \times n$ matrices over a field F. Then their sum $A + B$ is also an $m \times n$ matrix over F such that

$$(A + B)_{ij} = a_{ij} + b_{ij} \text{ for all } i \in m,\ j \in n$$

Let F be a field and m, n be positive integers. Then $F^{m\times n}$ denotes the set of all $m \times n$ matrices over field F. It is evident from the above definition that addition of matrices is a binary operation which possesses the following properties.

(i) Matrix addition on $F^{m\times n}$ is commutative,
i.e. $A + B = B + A$ for all $A, B \in F^{m\times n}$

(ii) Matrix addition on $F^{m\times n}$ is associative,
i.e. $(A + B) + C = A + (B + C)$ for all $A, B, C \in F^{m\times n}$

(iii) Null matrix O is the additive identity,
i.e. $A + O = A = O + A$ for all $A \in F^{m\times n}$

(iv) For every matrix $A \in F^{m\times n}$, there exists $-A \in F^{m\times n}$ such that
$A + (-A) = O = (-A) + A$

It follows from the above properties that $(F^{m \times n}, +)$ is an abelian group. Let $A, B, C \in F$ and λ, μ are scalars. Then we also have the following results:

(i) $\lambda(A + B) = \lambda A + \lambda B$

(ii) $(\lambda + \mu)A = \lambda A + \mu A$

(iii) $(\lambda\mu)A = \lambda(\mu A) = \mu(\lambda A)$

(iv) $1A = A$.

Multiplication of Matrices

If $A = [a_1, a_2, \ldots, a_n]$ is a row matrix and $B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$ is a column matrix, then

their product AB is defined to be the scalar (or 1×1 matrix) obtained by multiplying corresponding entries and adding, that is,

$$AB = [a_1, a_2, \ldots, a_n] \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = a_1 b_1 + a_2 b_2 + \ldots + a_n b_n = \sum_{r=1}^{n} a_r b_r$$

Note that the product AB is not defined when A and B have different number of elements. Let us now generalize the above definition for arbitrary matrices.

Let $A = [a_{ij}]$ and $B = [b_{ij}]$ be the two matrices over a field F such that the number of columns of A is equal to the number of rows of B; say A is $m \times p$ matrix and B is $p \times n$ matrix. Then, the product AB is $m \times n$ matrix whose ij-entry is obtained by multiplying ith row of A by the jth column of B, that is,

$$(AB)_{ij} = [a_1, a_2, \ldots, a_n] \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = \sum_{r=1}^{n} a_{ir} b_{rj}$$

The multiplication of matrices is not commutative. However, matrix multiplication satisfies the following properties.

Theorem: Let A, B and C be the three matrices over a field F such that various products and sums are defined. Then

(i) $(AB)C = A(BC)$
(ii) $A(B+C) = AB + AC$
(iii) $(B+C)A = BA + CA$
(iv) $k(AB) = (kA)B = A(kB)$, where $k \in F$
(v) $A_{m \times n} O_{n \times p} = O_{m \times p}$ and $O_{p \times m}\ A_{m \times n} = O_{p \times n}$

Here A is an $m \times n$ matrix.

Positive Integral Powers of a Square Matrix

For any square matrix A, we define (i) $A^1 = A$ and (ii) $A^{n+1} = A^n A$, where $n \in N$.

It is evident from the above definition that:

$$A^2 = AA, A^3 = A^2 A = AAA \text{ etc.}$$

Also

(i) $A^m A^n = A^{m+n}$ and (ii) $(A^m)^n = A^{mn}$ for all $m, n \in N$.

Matrix Polynomial

Let $f(x) = a_0 x^n + a_1 x^{n-1} + ... + a_{n-1} x + a_n$ be a polynomial over a field F and A be a square matrix over F. Then

$$f(A) = a_0 A^n + a_1 A^{n-1} + ... + a_{n-1} A + a_n I$$

is called a matrix polynomial.

Let F be a field and n be a positive integer. Then the product of two $n \times n$ matrices over F is an $n \times n$ matrix over F. So the set $F^{n \times n}$ of all $n \times n$ matrices over F is closed under multiplication of matrices. The foregoing discussion suggests that $(F^{n \times n}, +, \times)$ is a non-commutative ring with unity if $n > 1$.

Transpose of a Matrix

Let $A = [a_{ij}]$ be an $m \times n$ matrix over a field F. Then the transpose of A denoted by A^T or A' is an $n \times m$ matrix such that

$$(A^T)_{ij} = a_{ji} \text{ for all } i \in m, j \in n.$$

Clearly, A^T is obtained from A by interchanging rows and columns of A.

Following are the basic properties of the transpose operation.

Let A and B be the two matrices over a field F and λ be a scalar in F. Then whenever the sum and product are defined

(i) $(A^T)^T = A$ (ii) $(A+B)^T = A^T + B^T$ (iii) $(\lambda A)^T = \lambda A^T$
(iv) $(AB)^T = B^T A^T$

Symmetric Matrix

A square matrix $A = [a_{ij}]$ over a field F is said to be a symmetric matrix if

$$a_{ij} = a_{ji} \text{ for all } i, j \in n$$

$$\Leftrightarrow \quad A = A^T$$

Skew-Hermitian Matrix

A matrix A over the field C of all complex numbers is a skew-hermitian matrix if its conjugate transpose is equal to $-A$, i.e. $A^{-T} = -A$ or $A^* = -A$.

The diagonal elements of a skew-hermitian matrix are purely imaginary.

Unitary matrix A square matrix A over C is a unity matrix if $A^*A = 1 = AA^*$.

If A is the square matrix over C, then A is a normal matrix if $AA^* = A^*A$.

Clearly, this definition reduces to the definition of a normal matrix over R if C is replaced by R.

The conjugate transpose of a square matrix satisfies the following properties.

(i) $(A^*)^* = A$ (ii) $(\lambda A)^* = \bar{\lambda} A^*, \lambda \in C$ (iii) $(A+B)^* = A^* + B^*$

(iv) $(AB)^* = B^*A^*$ (v) $(A^*)^{-1} = (A^{-1})^*$

Inverse of a Matrix

A square matrix B is said to be the inverse of a square matrix A if $AB = BA = 1$.

If the inverse of a square matrix A exists, then it is unique and we say that A is invertible. The inverse of A is denoted by A^{-1}.

Following are the properties of inverse of a matrix.

(i) If A is an invertible matrix, then $(A^{-1})^{-1} = A$.

(ii) A square matrix is invertible if it is non-singular.

(iii) Let A and B be invertible matrices, then AB is invertible and $(AB)^{-1} = B^{-1}A^{-1}$

(iv) Let A, B, C be square matrices of the same order and if A is an invertible matrix, then

$$AB = AC \Rightarrow B = C \text{ and } BA = CA \Rightarrow B = C$$

(v) If A is an invertible matrix, then A^T is also invertible and $(A^T)^{-1} = (A^{-1})^T$

(vi) The inverse of an invertible symmetric matrix is symmetric matrix.

(vii) The set $F^{n\times n}$ of all invertible matrices over a field F is a non-abelian group under multiplication of matrices.

1.9 DETERMINANTS

Determinants: Every square matrix over a field F can be associated to a scalar in F which is known as its determinant.

The determinant of A is denoted by $|A|$.

$$\text{If } A=\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \text{ is a square matrix over a field } F, \text{ then}$$

$$|A| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

$$\text{If } A=\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \text{ is a square matrix over a field } F, \text{ then}$$

$$|A| = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

$$= a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

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

$$= a_{11}a_{22}a_{33} - a_{12}a_{23}a_{31} + a_{13}a_{32}a_{21} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}$$

Singular Matrix

A matrix A over a field F is called a singular matrix if $|A| = 0$. Otherwise, it is a non-singular matrix.

Minor

Let $A = [a_{ij}]$ be a square matrix of order n. Then the minor M_{ij} of a_{ij} in A is the determinant of the square submatrix of order $(n-1)$ obtained by leaving ith row and jth column of A.

Cofactor: Let $A = [a_{ij}]$ be a square matrix of order n. Then the cofactor C_{ij} of a_{ij} in A is equal to $(-1)^{i+j}$ times the determinant of the submatrix of

order $(n-1)$ obtained by leaving ith row and jth column of A.

Also,

(i) $\sum_{i=1}^{n} a_{ij}C_{ij} = |A|$ and $\sum_{j=1}^{n} a_{ij}C_{ij} = |A|$

(ii) $\sum_{i=1}^{n} a_{ij}C_{ij} = 0$ and $\sum_{j=1}^{n} a_{ij}C_{ij} = 0$

Adjoint of a Square Matrix

Let $A = \left[a_{ij}\right]$ be a square matrix of order n and let C_{ij} be a cofactor of a_{ij} in A. Then the transpose of the matrix of cofactor of elements of A is called the adjoint of A and is denoted by adj A, that is,

$$\text{adj}A = \left[C_{ij}\right]^T \text{ for all } i,\ j \in n$$

Following are some useful properties of adjoint of a matrix.

(i) For any square matrix A of order n

$A(\text{adj}A) = |A| I_n = (\text{adj}A)A$

(ii) For any square matrices A and B of order n

$\text{adj}AB = \text{adj}B\ \text{adj}A$

(iii) If A is an invertible matrix, then $\text{adj}A^T = (\text{adj}A)^T$

(iv) If A is an invertible matrix of order n, then

$\text{adj}(\text{adj}A) = |A|^{n-2} A$

(v) If A is a non-singular matrix, then $A^{-1} = \frac{1}{|A|}(\text{adj}A)$

1.10 SYSTEMS OF LINEAR EQUATIONS

A system of m linear equations in n unknowns $x_1, x_2, \ldots, x_n$ can be put in the standard form as follows:

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

where a_{ij} and b_{ij} are constants.

This system of equations is known as $m \times n$ (read as m by n) system and can be written in the matrix form as follows :

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

Or $AX = B$, where $A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}_{m \times n}$, $X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}_{n \times 1}$ and

$$B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}_{m \times 1}$$

The matrix $A = \left[a_{ij} \right]_{m \times n}$ is called the coefficient matrix and the matrix

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

is called the **augmented matrix** and is generally denoted by $[A : B]$.

A system of equations $AX = B$ is called a homogeneous system if $B = 0$. Otherwise, the system is said to be non-homogeneous.

A solution of the system of equations $AX = B$ is a list of the values for the unknowns which satisfy each equation of the system. Equivalently, a vector $U \in F^n$ is a solution of the system of equations $AX = B$.

If $AX = B$ is a system of n equations with n unknowns such that $|A| = 0$, then the system has unique solution given by $X = A^{-1}B$.

If $|A| \neq 0$, then the system of equations is either inconsistent or it has infinitely many solutions.

A homogeneous system of equations $AX = 0$ is always consistent and has trivial solution only if $|A| \neq 0$. If $|A| = 0$, then $AX = 0$ has non-trivial solutions also.

1.10.1 Systems of Equations in Triangular Form

Consider the following system of linear equations:

$$\begin{aligned} 3x_1 - 2x_2 + 4x_3 - 3x_4 &= 8 \\ 5x_2 - 2x_3 + 3x_4 &= 7 \\ 7x_3 - 2x_4 &= 3 \\ 3x_4 &= 6 \end{aligned}$$

We observe that the system is square and the first unknown x_1 is the leading unknown in the first equation, the second unkonwn x_2 is the leading unknown in the second equation and so on. Such a system is said to be in triangular form.

Thus, a square system of linear equations is said to be in triangular form if each leading unknown is directly to the right of the leading unknown in the preceding equation.

Clearly, a triangular system always has a unique solution which may be obtained by back substitution.

1.10.2 Systems of Equations in Echelon Form

A system of simultaneous linear equations is said to be in echelon form if the leading unknown in each equation other than the first is to the right of the leading unknown in the preceding equation.

Consider the following system of equations in echelon form

$$\begin{aligned} 3x_1 - 5x_2 + 3x_3 + 3x_4 - x_5 &= 13 \\ x_3 - 5x_4 + 2x_5 &= 7 \\ 2x_4 - 7x_5 &= 9 \end{aligned}$$

Clearly, x_1, x_3 and x_4 are the leading unknowns in this system.

These unknowns are called **pivot** variables and the other unknowns x_2 and x_5 are called **free variables.**

The solution set of a system of m simultaneous linear equations in n unknowns in echelon form is described into the following theorem.

Theorem: Let there be a system of simultaneous linear equations in n unknowns in echelon form, then

(i) The system has a unique solution if $m = n$, i.e. the system is in triangular form.

(ii) The system has an infinite number of solutions, if $m < n$, i.e. there are more variables than the number of equations.

Remark: If the echelon system of simultaneous linear equations contains more variables than equations, then each of the remaining $n - m$ free variables may take any value. So, the system has infinitely many solutions. The general solution of such a system may be obtained in either of the following two equivalent ways.

(i) Arbitrarily assign values to the $n - m$ free variables and solve uniquely for the m pivot variables to obtain a solution of the system.

(ii) Find the values of m pivot variables in terms of $(n - m)$ free variables to obtain the general solution of the system.

1.11 RANK OF A MATRIX

The rank of a matrix is defined in many different ways. But all the definitions lead to the same number.

Rank of a Matrix

The rank of a matrix is the order of the highest order non-singular square submatrix.

It is evident from the above definition that a positive integer r is rank of an $m \times n$ matrix if

(i) Every square submatrix of order $(r + 1)$ or more is singular.

(ii) There exists at least one square submatrix of order r which is non-singular.

The rank of a matrix A is written as rank (A).

Clearly, the rank of the identity matrix I_n is n.

If A is an $m \times n$ matrix, then rank $(A) \leq \min(m, n)$.

The rank of a matrix in echelon form is equal to the number of non-zero rows of the matrix or the number of pivots.

Theorem 1: The system of linear equations $AX = B$ is consistent iff the rank of the augmented matrix $[A : B]$ is equal to the rank of the coefficient matrix A.

Theorem 2: Let $AX = B$ be a system of m simultaneous linear equations in n unknowns such that $m \geq n$.

(i) If $r(A) = r([A : B]) = n$, the system has a unique solution.

(ii) If $r(A) = r([A : B]) = r > n$, the system is consistent and has infinite number of solutions.

In fact, in this case $(n - r)$ variables are free variables.

(iii) If $r(A) \neq r([A : B])$, the system is inconsistent, i.e. it has no solution.