NETAJI SUBHAS OPEN UNIVERSITY

STUDY MATERIAL

MATHEMATICS
POST GRADUATE

PG (MT) 01
GROUPS A & B

Abstract Algebra

●

Linear Algebra

PREFACE

In the curricular structure introduced by this University for students of Post-Graduate Degree Programme, the opportunity to pursue Post-Graduate course in any subject introduced by this University is equally available to all learners. Instead of being guided by any presumption about ability level, it would perhaps stand to reason if receptivity of a learner is judged in the course of the learning process. That would be entirely in keeping with the objectives of open education which does not believe in artificial differentiation.

Keeping this in view, study materials of the Post-Graduate level in different subjects are being prepared on the basis of a well laid-out syllabus. The course structure combines the best elements in the approved syllabi of Central and State Universities in respective subjects. It has been so designed as to be upgradable with the addition of new information as well as results of fresh thinking and analysis.

The accepted methodology of distance education has been followed in the preparation of these study materials. Co-operation in every form of experienced scholars is indispensable for a work of this kind. We, therefore, owe an enormous debt of gratitude to everyone whose tireless efforts went into the writing, editing and devising of proper lay-out of the materials. Practically speaking, their role amounts to an involvement in 'invisible teaching'. For, whoever makes use of these study materials would virtually derive the benefit of learning under their collective care without each being seen by the other.

The more a learner would seriously pursue these study materials, the easier it will be for him or her to reach out to larger horizons of a subject. Care has also been taken to make the language lucid and presentation attractive so that they may be rated as quality self-learning materials. If anything remains still obscure or difficult to follow, arrangements are there to come to terms with them through the counselling sessions regularly available at the network of study centres set up by the University.

Needless to add, a great deal of these efforts is still experimental–in fact, pioneering in certain areas. Naturally, there is every possibility of some lapse or deficiency here and there. However, these do admit of rectification and further improvement in due course. On the whole, therefore, these study materials are expected to evoke wider appreciation the more they receive serious attention of all concerned.

Professor (Dr.) Subha Sankar Sarkar
Vice-Chancellor

Seventh Reprint : October, 2017

Printed in accordance with the regulations of the Distance Education Bureau
of the University Grants Commission.

Subject : Mathematics

Post Graduate

Paper : PG (MT) 01 : Group A

Writer

Prof. Pranay Kumar Chaudhuri

Editor

Prof. Mithil Ranjan Gupta

Paper : PG (MT) 01 : Group B

Writer

Prof. Pranay Kumar Chaudhuri

Editor

Prof. Manjusha Majumdar

Notification

Mohan Kumar Chattopadhyay
Registrar

Netaji Subhas
Open University

PG (MT)—01
Abstract Algebra,
Linear Algebra

Group A
Abstract Algebra

Group B
Linear Algebra

Unit : 1 ❑ Preliminaries and basic concepts

The present section aims at refreshing the basic concepts that we have already encountered in our earlier studies. The study of algebraic systems will require clear concepts on sets, relations, mappings, operations etc. and a quick review will help better understanding of the materials presented in this course. Here we shall introduce the ideas in short and only state the relevant theorems but we shall not go into the details of these concepts.

1.1 Sets, relations, mappings :

Set : A set is a well defined collection of objects. By well defined collection of objects we understand that if 5 is a set and a is some object, then either a is definitely in S, denoted by $a \in S$, or, a is definitely not in 5, denoted by $a \notin S$.

The union of two sets A and B, written as $A \cup B$, is the set $\{x \mid x \in A$ or $x \in B\}$

The intersection of two sets A and B, written as $A \cap B$, is the set $\{x \mid x \in A$ and $x \in B\}$.

A null set is the set having no element. A null set is denoted by ϕ. If A and B are two sets such that every element of A is an element of B, then A is called a subset of B and is written as $A \subseteq B$.

Thus the null set is a subset of every set.

Two sets A and B are said to be disjoint if $A \cap B = \phi$.

Given two sets A and B, the difference $A - B$ is the set $\{x \in A \mid x \notin B\}$.

Cartesian product of two sets : If A and B are two sets, then the Cartesian product of the sets A and B, denoted by $A \times B$, is the set

$$A \times B = \{(x, y) \mid x \in A, y \in B\}.$$

Thus, if $A = \{x, y\}$ and $B = \{y, z, t\}$, then $A \times B$ is the set of distinct ordered pairs

$$\{(x, y), (x, z), (x, t), (y, y), (y, z), (y, t)\}$$

Relations : Let A and B be two sets and let ρ be a subset of $A \times B$. Then ρ is called a relation from A to B. If $(x, y) \in \rho$, then x is said to be in relation ρ to y, written $x\rho y$. A relation from A to A is called a relation on A (or in A).

Let ρ be a relation in the set A. ρ is said to be

(a) reflexive if $x\rho x$ for all $x \in A$

(b) symmetric if $x\rho y$ implies $y\rho x$; $x, y \in A$

(c) antisymmetric if $x\rho y$ and $y\rho x$ imply $x = y$; $x, y \in A$

(d) transitive if $x\rho y$ and $y\rho z$ imply $x\rho z$; $x, y, z \in A$

If the relation ρ is reflexive, symmetric and transitive then ρ is called an equivalence relation on A. If ρ is reflexive, antisymmetric and transitive then ρ is called a partial ordering relation on A.

An equivalence relation ρ defined on a set A partitions the set A into a number of disjoint classes, called the equivalence classes. Thus if $a \in A$, the equivalence class of a is denoted by cl(a) and it is the set $\{x|a\rho x\}$.

Mappings : If S and T are non empty sets, then a mapping from S to T is a subset M of $S \times T$ such that for every $s \in S$, there is a unique $t \in T$ such that the ordered pair (s, t) is in M. *Let* σ be a mapping from S to T; we often denote this by writing $\sigma : S \to T$ *or* $S \xrightarrow{\sigma} T$. If t is the image of s under σ, then we shall write $t = \sigma(s)$.

Let S be any set. Let us define the mapping $I : S \to S$ by $I(s) = s$ for any $s \in S$. This mapping I is called the identity mapping of S.

The mapping σ of S to T is said to be onto (or surjective) mapping if given $t \in T$ there exists an element $s \in S$ such that $\sigma(s) = t$.

The mapping σ of S to T is said to be one-to-one (or injective) mapping if whenever $s_1 \neq s_2$, ($s_1, s_2 \in$ S), then $\sigma(s_1)$ $\sigma(s_2)$.

A one-to-one and onto mapping is called a bijective mapping.

The two mapping σ and τ of S into T are said to be equal if $\sigma(s) = \tau(s)$ for every $s \in S$.

If $\sigma : S \to T$ and $\tau : T \to U$, then the composition of σ and τ (also called their product) is the mapping $\tau_0\sigma : S \to U$ defined by means of $(\tau_0\sigma)$ $(s) = \tau(\sigma(s))$ for every $s \in S$.

Thus for the composition $\tau_0\sigma$ of σ and τ, we shall always mean : first apply σ and then τ.

Lemma 1.1 : (Associative law). If $\sigma : S \to T$, $\tau : T \to U$ and $\mu : U \to V$ are three mappings, then the associative law $(\mu\ o\ \tau)$ os $= \mu o(\tau o \sigma)$ holds.

1.2 Binary operation :

A binary operation o on a set S is a rule that assigns to each ordered pair of elements of the set S some element of the set. Thus, for an arbitrary set S, we call a mapping of $S \times S$ into S, a binary operation on S.

Example 1.1 : On Z^+ (the set of positive integers), define a binary operation o by aob equals the smaller of a and b or the common value if $a = b$. $a, b \in Z^+$. Thus $2o11 = 2$; $15o10 = 10$ and $4o4 = 4$.

Example 1.2 : On Z^+ define (he operation o by $aob = a|b$, $a, b \in Z^+$. Clearly the operation o is not a binary operation as $1o3 = \frac{1}{3} \in Z^+$.

A binary operation o on a set S is commutative if and only it $aob = boa$ for all $a, b \in S$. The operation o is associative if and only if $(aob)\ oc = ao\ (boc)$ for all $a, b, c \in S$.

Remark : In defining a binary operation on a set S it is necessary that (i) exactly one element is assigned to each ordered pair $(a., b)$, $(a, b \in S)$ and (ii) for each ordered pair of elements of S, the element assigned to it is again in S. If the second condition is not satisfied then we say that S is not closed under the operation.

1.3 Algebraic structure or algebraic system :

An algebraic structure or an algebraic system is a non empty set together with one or more binary operations on that set. Algebraic structures whose binary operations satisfy particularly important properties are groupoids, monoids, semi groups, groups, rings, fields, modules and so on.

Groupoid : A groupaid is an algebraic system consisting of a non empty set G and a binary operation o on G. The pair $[G, o]$ is called a groupoid.

The set of real numbers with the binary operation of addition is a groupoid.

Semigroup : If $[G, o]$ is a groupoid and if the associative rule $ao(boc) = (aob)oc$ holds for all $a, b, c \in G$, then $[G, o]$ is called a semigroup.

An element e of a groupoid [G, o] is called an identity element if $eoa = aoe = a$ for all $a \in G$. If there is an identity element in a groupoid then it is unique.

Monoid : A semigroup with identity element is called a monoid.

Example 1.3 : The set of all $n \times n$ matrices under the operation of matrix multiplication is a monoid. Here the identity element is the unit matrix of order n.

Let $[G, o]$ be a monoid. An element $a' \in G$ is culled an inverse of the element $a \in G$ if $a' \, o \, a = a \, o \, a' = e$ (the identity element of G). The inverse of the element $a \in G$ is denoted by a^{-1}.

Group : A monoid in which every element has an inverse is called a group.

Example 1.4 : The set of all $n \times n$ matrices under the operation of matrix multiplication is not a group since not every $n \times n$ matrix has its multiplicative inverse, but if G is the set of all $n \times n$ nonsingular matrices, then G forms a group under the operation of matrix multiplication.

1.4 Some elementary properties of groups, important theorems on groups :

(i) The identity element e is unique in a group $[G, o]$.

(ii) The inverse a^{-1} of the element $a \in G$ is unique.

(iii) For every $a \in G$, $(a^{-1})^{-1} = a$.

(iv) For all $a, b \in G$, $(aob)^{-1} = b^{-1}oa^{-1}$.

(v) The cancellation laws

$aob = aoc \Rightarrow b = c$ (left cancellation law)

$boa = coa \Rightarrow b = c$ (right cancellation law)

hold in G; $a, b, c \in G$.

(vi) For a, $b \in G$, the linear equations $aox = b$ and $yoa = b$ have unique solutions in G.

A group G is said to be abelian (or commutative) if for every $a, b \in G$, $aob = boa$.

The order of an element a in a group G is the smallest positive integer n such that $a^n = e$, the identity element in G. Here, by a^n we understant $aoa \, o \, \ldots \, oa$ (n factors). If there is no finite integer n such that $a^n = e$, we say that the element a has order zero. If an element $a \in G$ has order n, we write $o(a) = n$.

The order of a group is the number of elements in the group. Thus, if a group G has n elements, then G is said to be a finite group of order n. and we write $O(G) = n$.

Example 1.5 : If G is a group of even order, prove that it has an element $a \neq e$ satisfying $a^2 = e$.

Solution : Suppose that G is a group of order $2n$, n being a positive integer. Let $a_1, a_2 \ldots a_{2n-1}, e$ be the elements of G. Since in a group every element has its unique inverse and since there are odd number of elements $a_1, a_2 \ldots a_{2n-1}$, none of which is the identity element of G, if follows that there is one element a (say) among $a_1, a_2 \ldots a_{2n-1}$ whose inverse is a itself. Then it follows that $aoa = e$ i.e. $a^2 = e$, $a \in G$ and $a \neq e$.

Example 1.6 : Show that if every element of the group G is its own inverse, then G is abelian.

Solution : Let $a, b \in$ G, then $aob \in G$. From the given condition $a^{-1} = a$, $b^{-1} = b$, $(aob)^{-1} = aob$

But $(aob)^{-1} = b^{-1}oa^{-1}$, so that $aob = b^{-1}oa^{-1}$

or, $aob = boa$. Hence G is abelian.

Note : We have already pointed out that group is an algebraic system with a set G and a binary operation o defined on G. So far we have written aob as the composition of two elements a and b in G, but henceforth we shall drop the symbol o and write simply as ab. But we must keep in mind that there is a suitably defined operation for the composition ab of two elements a and b in G.

Subgroup : A non empty subset H of a group G is said to be a subgroup of G if, under the composition rule in G, H itself forms a group.

Example 1.7 : Let G be the group of integers under addition, H is the subset of G consisting of all the multiples of 5. Then it may be checked that H is a subgroup of G.

Example 1.8 : Let G be a group of all non zero complex numbers $a + ib$ (a, b real and not both are 0) under multiplication, and let

$$H = \{a + ib \in G \mid a^2 + b^2 = 1\}$$

Then H is a subgroup of G.

Theorem 1.1 : A non-void subset H of a group G is a subgroup of G if and only if $a, b \in H \Rightarrow ab^{-1} \in H$.

If H is a subgroup of a group G and $a \in G$ then the set

$Ha = \{ha \mid h \in H\}$

is called a right coset of H in G.

Similarly, the set $aH = \{ha \mid h \in H\}$

is called a left coset of H in G.

It may easily be shown that any two right (or left) cosets of H in G are either identical or have no element in common.

Theorem 1.2 : (Lagrange) If G is a finite group and H is a subgroup of G, then $o(H)$ is a divisor of $o(G)$,

As a corollary to Lagranges theorem we may state that if a is an element of a finite group G, then $o(a) \mid o(G)$ [It is read as order of a divides order of G].

Cyclic group : A group G is said to be a cyclic group if there exists an element $a \in G$ such that every element in G can be expressed as a power of a. The element a is then called a generator of G and we denote G by $< a >$.

When the operation is addition we write $G = \{n\ a \mid n \in Z\} = < a >$.

Example 1.9 : Let n be a positive integer. The equation $x^n = 1$ has roots $e^{\frac{2\pi ki}{n}}$, $k = 0, 1, 2 \ldots n - 1$. These are called the n-th roots of unity, if we consider the set G of n-th roots of unity and define on it the usual multiplication as the binary operation, then we get a group. Clearly, any element of the group can be expressed as a^k, $k = 0, 1, 2 \ldots n - 1$, where $a = e^{\frac{2\pi i}{n}}$. Hence the group G is a cyclic group and the element a is a generator of this group.

Example 1.10 : The additive group Z (the set of integers) is generated by the element $1 \in Z$.

We note here that since the operation is addition, by a^n we mean $a + a + \ldots + a(n$ terms). Thus, any element n of the additive group Z is expressed as $n = 1 + 1 + \ldots + 1(n$ terms). Hence Z is a cyclic group with generator 1.

Theorem 1.3 : Order of a cyclic group is equal to the order of its generator.

We note that a cyclic group may have more then one generating element. For example, the group Z_4 (integer modulo 4) with addition as operation has 1 and 3 as generators. Thus, we may write

$Z_4 = < 1 > = < 3 >$

Theorem 1.4 : A subgroup of a cyclic group is cyclic.

Normal subgroup : A subgroup N of a group G is said to be a normal subgroup of G if for every $g \in G$ and $n \in N$, $g^{-1}ng \in N$.

Equivalently, if by $g^{-1}Ng$ we mean the set of all $g^{-1}ng$, $n \in N$, then N is a normal subgroup of G if and only if $g^{-1}Ng \subseteq N$ for every $g \in G$.

Note : We have defined N to be a normal subgroup of G if $g^{-1}ng \in N$ for $g \in G$, $n \in N$.

Since G is a group, for every $g \in G$, $g^{-1} \in G$, and it follows that N is a normal subgroup of G if for every $g \in G$ and $n \in N$ we have

$$gng^{-1} \in N.$$

Theorem 1.5 : The subgroup N of G is a normal subgroup of G if and only if every left coset of N in G is a right coset of N in G.

Example 1.11 : For any two subgroups H and K of a group G the following assertions hold :

(a) $H \cap K$ is a subgroup of G.

(b) If H is normal in G, then $H \cap K$ is normal in K.

(c) If H and K are both normal in G, then $H \cap K$ is normal in G.

Proof (a) : Since H and K are both subgroups of G, and $H \in K$ and $e \in K$, so that $e \in H \cap K$ and $H \cap K$ is non-empty. Let x and $y \in H \cap K$, then x and y belong to H and also to K. Since H and K are subgroups of G, $xy^{-1} \in H$ and K.

$\therefore$ xy^{-1} and $H \cap K$ is a subgroup of G.

(b) Let H be normal in G. Let $k \in K$ and $a \in H \cap K$. Then $k^{-1}ak \in K$, since $k \in K$ and $a \in K$.

Again $k^{-1}ak \in H$, since H is a normal subgroup of G and $k \in K \subset G$, $a \in H$.

Hence $k^{-1}ak \in H \cap K$ for $a \in H$ and $k \in K$. So $H \cap K$ is a normal subgroup of K.

(c) Let H and K are both normal subgroups of G. Then for $x \in H \cap K \Rightarrow x \in H$ and $x \in K$. But H and K are normal subgroups of G. So for every $g \in G$, $g^{-1}xg \in H$ and $g^{-1}xg \in K$

$\therefore$ For every $g \in G$ and $x \in H \cap K$, $g^{-1}xg \in H \cap K$

Hence $H \cap K$ is a normal subgroup of G.

1.5 Rings, integral domains and fields :

Ring : A non empty set R is said to be a ring if in R there are denned two binary operations + and, which we call addition and multiplication respectively, such that for a, b, c in R :

1. The algebraic system $< R, + >$ is an abetian group
2. The associative law with respect to multiplication holds; that is, $a \cdot (b \cdot c) = (a \cdot b) \cdot c$.
3. The two distributive laws $a \cdot (b + c) = a \cdot b + a \cdot c$ and $(b + c) \cdot a = b \cdot a + c \cdot a$ hold.

We note here that in our future discussion, as in the case of group, we shall drop the multiplication operation dot (·) and simply write ab instead of $a \cdot b$.

In our definition of a ring, if there is an element 1 in R such that $a \cdot 1 = 1 \cdot a = a$ for every $a \in R$, then we shall call R as a ring with unity element.

If the multiplication of R is such that $ab = ba$ for every a, b in R, then R will be called a commutative ring.

Example 1.12 : The set of integers with usual addition and multiplication forms a commutative ring with unity element, while the set of even integers under usual operations of addition and multiplication forms a commutative ring but the ring has no unity element.

Note : In case of a ring, with unity element, the algebraic system $< R, + >$ is an ahelian group and the algebraic system $< R, \cdot >$ is a monoid.

Some elementary properties :

(i) If o is the additive identity of the ring $< R, +, \cdot >$, then for any $a \in R$, $a \cdot o = o = o \cdot a$.

(ii) For all $a, b \in R$, $(-a).b = -(a.b) = a.(-b)$.

(iii) For all elements $a_1, a_2 \dots a_m$ and $b_1, b_2, \dots b_n$ in R, $(a_1 + a_2 + \dots + a_m)$. $(b_1 + b_2 + \dots + b_n) = \sum_{i=1}^{m}\sum_{j=1}^{n} a_i.b_j$.

(iv) For all integers n and all $a, b \in R$
$n(a.b) = (na).b = a.(nb)$.

(v) In any commutative ring R the binomial formula

$$(a+b)^n = \sum_{i+j=n} \binom{n}{i} a^i b^j, \; a, b \in R$$

holds for natural numbers n and non-negative integers i and j.

Example 1.13 : If $< R, +, \cdot >$ is an algebraic system satisfying all the conditions for a ring with unity with possible exception of $a + b = b + a$, $a, b \in R$, prove that the axiom $a + b = b + a$ must hold in R and that R is a ring.

Solution : Since $1 \in R$, we have

$$(a + b).(1 + 1) = (a + b).1 + (a + b).1$$
$$= (a + b) + (a + b)$$

Also, $(a + b).(1 + 1) = a.(1 + 1) + b.(1 + 1)$

$$= (a.1 + a.1) + (b.1 + b.1)$$
$$= (a + a) + (b + b)$$

Thus $(a + b) + (a + b) = (a + a) + (b + b)$

Since associativity with respect to addition holds in R, we have, by left and right cancellation rules

$$b + a = a + b$$

Hence $< R, + >$ is an abelian group and $< R, +, . >$ is a ring with unity.

Some special classes of rings :

Example 1.12 considers rings of integers where for two elements a and b in the ring $ab = 0$ will imply either $a = 0$ or $b = 0$ and also $ab = ba$ holds. But in general rings there is a possibility of $ab = 0$ with neither a nor b being zero. Also there are rings where $ab = ba$ does not hold. For example, in the ring of integers with addition and multiplication modulo b, we have $2 \times 6^3 = 0$, but 2 and 3 are non zero elements. Also it may be easily checked that in the ring $M_2(R)$ or 2×2 matrices with elements as real numbers the commutative property does not hold.

Accordingly, we have the following definitions.

Def. : If R is a commutative ring, then $a \neq o \in R$ is said to be a **zero divisor** if there exists $b \in R$, $b \neq o$ such that $ab = o$.

In the example just considered, the integer 2 is a zero divisor in the ring of integers $< Z, +_6, \times_6 >$.

Def. : An **integral domain** is a commutative ring with unity element and without a zero divisor.

The ling of integers is an example of an integral domain.

Del. : A ring is said to be a **division ring** or a skew Held if its non zero elements form a group under multiplication.

Def. : A field is a commutative ring with unity element such that the non zero elements form a group under multiplication. Thus, a field is a commutative division ring. Since the cancellâtion law holds in any group, any fiéld is an integral domain.

The rational numbers with usual addition and multiplication forms a field. The set R of integers mod 7 under the addition and multiplication mod 7 form a field. Evidently, the elements of R are the seven symbols $\overline{0}, \overline{1}, \overline{2}, \overline{3}, \overline{4}, \overline{5}, \overline{6}$ where, for example, the element $\overline{2}$ represents the class of integers of the form $7k + 2$, k being au integer.

We have just remarked that any field is an integral domain, but for the converse, we have,

Theorem 1.6 : A finite integral domain is a field.

Corollary : If p is a prime number, then Z_p, the ring of integers mod p, is a field.

Def. : A subset S of a ring $R = < R, +, \cdot >$ is said to be a **subring** of R if and only if S forms a ring under the operations $+$ and $\cdot$ of R.

Theorem. 1.7 : A non empty subset S of a ring $R = < R, +, \cdot >$ is a subring of R if and only if with two elements x and y of S, $x - y$ and $xy \in S$.

Example 1.14 : $\{o\}$ and R are always sub rings of R. These are known as improper subrings of R.

$< Z, +,, \cdot >$ is a sub ring of $< Q, +, \cdot >$

The set of all $n \times n$ matrices over the ring of rational numbers is a sub ring of $n \times n$ matrices over the ring of real numbers.

EXERCISES

1. Let G be a group such that $a^2 = e$ for all $a \in G$. Show that G is abelian.

2. Prove that the matrices

$$\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}, \begin{bmatrix}1 & 0\\0 & -1\end{bmatrix}, \begin{bmatrix}-1 & 0\\0 & 1\end{bmatrix} \text{and} \begin{bmatrix}-1 & 0\\0 & -1\end{bmatrix}$$

form a multiplicative group, (This group is known as the Klein four-group).

3. Prove : For any group G, and any $a \in G$, $a^m o\ a^n = a^{m+n}$ where m, n belong to the set of integers.

4. Prove : Every proper subgroup of an abelian group is abelian. State the converse and show by an example that it is false.

5. Let S be a subgroup of a group G. Define $T = \{x : x \in G, xS - Sx\}$. Prove that T is a subgroup of G.

6. Show that the set $\{x : x \in Z, 5 \mid x\}$ is a subgroup of the additive group Z. (The symbol $5 \mid x$ represents the element x which is divisible by 5).

7. Let G be the set of all 2×2 matrices $\begin{pmatrix}a & b\\o & d\end{pmatrix}$ where $ad \neq o$, under matrix multiplication.

   Let $N = \left\{\begin{pmatrix}1 & b\\o & 1\end{pmatrix}\right\}$. Prove that N is a normal subgroup of G.

8. If a cyclic subgroup T of G is normal in G, then show that every subgroup of T is normal in G.

9. Let G be the set of all real 2×2 matrices $\begin{pmatrix}a & b\\o & d\end{pmatrix}$ with $ad \neq o$. Prove that G forms a group under matrix multiplication. Is G abelian?

10. Suppose that H is the only subgroup of $o(H)$ in the finite group G. Prove that H is a normal subgroup of G.

11. Show that the set of integers with addition and multiplication modulo 6 is a ring with zero divisors.

12. An element x in an integral domain D is called an idempotent element if $x^2 = x$. Show that only idempotents in D are 0 and 1.

Unit : 2 ❑ Permutation groups and quotient groups

Here we shall consider groups whose elements are entities called permutations or whose elements are the classes of elements of another group. Permutation groups are non-abelian and it may be shown that any group is structurally the same as some group of permutations. We shall show that a law of composition can be defined on the cosets of a normal subgroup N of any group G and will describe how to make the set of cosets into a group, called a quotient group. We start with the concept of permutation groups.

2.1 Permutations

Def 2.1 : A permutation of a set S is a function from S into S that is both one-to-one and onto. In otherwords, a permutation of S is a one-to-one function from S onto S.

Let $S = \{1, 2, \dots n\}$ and consider the set S_n of n ! permutations of these n symbols. Let i_1, i_2 ... i_n be some arrangement of the elements of S. We introduce a two-line notation for the permutation

$$\alpha = \begin{pmatrix} 1 & 2 & 3 \dots n \\ i_1 & i_2 & i_3 \dots i_n \end{pmatrix},$$

where α takes the element 1 to i_1, the element 2 to i_2, ... and the element n to i_n. Cleary i_1, i_2, ... i_n are the elements 1, 2, ... n, may be in some different order.

Similarly, if j_1, j_2, ... j_n is another arrangement of the elements of S, we write

$$\beta = \begin{pmatrix} 1 & 2 & 3 \dots n \\ j_1 & j_2 & j_3 \dots j_n \end{pmatrix}.$$

We define an operation o on the set S_n of all permutations of S such that $\alpha \, o \, \beta$ means that the permutations α and β are to be performed in that order namely, first α and then β.

For example, if $\alpha = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 2 & 4 \end{pmatrix}$ and $\beta = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 2 & 4 \end{pmatrix}$, then $\alpha \, o \, \beta = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 3 & 4 \end{pmatrix}$.

Another notation for permutation is also used. For example, the permutation

$$\alpha = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 4 & 5 & 1 \end{pmatrix}$$

is written as (1 2 3 4 5). This notation is known as cyclic notation. The cycle (1 2 3 4 5) means 1 is replaced by 2, 2 is replaced by 3, 3 by 4, 4 by 5 and 5 by 1.

The permutation $\beta = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 4 & 5 & 3 \end{pmatrix}$ can be written as (3 4 5). In this cyclic notation 1 and 2 are missing. This means that 1 and 2 are unchanged in position, while 3 is replaced by 4, 4 by 5 and 5 by 3.

The permutation $\gamma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 3 & 2 & 5 & 4 \end{pmatrix}$ is written as, (2 3) (4 5), called the product of cycles.

Now we write formally.

Def. 2.2. : A permutation σ of a set S is a cycle of length n if there exist a_1, a_2 ... $a_n \in S$ such that $\sigma(a_1) = a_2$, $\sigma(a_2) = a_3$, ... $\sigma(\alpha_{n-1}) = a_n$, $\sigma(a_n) = a_1$ and $\sigma(x) = x$ for $x \in S$ but $x \notin \{a_1, a_2, ... a_n\}$. We write $\sigma = (a_1, a_2, ... a_n)$.

In using the cyclic notation for permutation of the set S, the elements of the set S must be known completely.

The permutation $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 2 & 5 & 1 & 4 \end{pmatrix}$, represented in cyclic notation as (1 3 5 4), is a cycle of length 4. We also note that (1 3 5 4) = (3 5 4 1) = (5 4 1 3) = (4 1 3 5). Of course, since cycles are special types of permutations, they can be multiplied just as any two permutations. The product of two cycle need not again be a cycle.

Example 2.1 : Consider the set S_6 of all permutation of $S = \{1, 2, 3, 4, 5, 6\}$.

Then (1 4 5 6) (2 1 5) $= \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 4 & 2 & 3 & 5 & 6 & 1 \end{pmatrix} o \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 5 & 1 & 3 & 4 & 2 & 6 \end{pmatrix}$

$$= \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 4 & 1 & 3 & 2 & 6 & 5 \end{pmatrix}$$

and (2 1 5)(1 4 5 6) $= \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 6 & 4 & 3 & 5 & 2 & 1 \end{pmatrix}$

Neither of these permutations is a cycle.

Def. 2.3 : Two cycles of a set S are disjoint if there is no element of S common to two cycles.

For example the cycles (1 6) and (2 5 3) for the set $S = \{1 \ \ 2\ 3\ 4\ 5\ 6\}$ are disjoint. The permutation $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 6 & 5 & 2 & 4 & 3 & 1 \end{pmatrix}$ can be expressed as $\sigma = (1\ 6)\ (2\ 5\ 3)$, the product of two disjoint cycles.

Multiplication of disjoint cycles is clearly commutative, so the order of the factors (1 6) and (2 5 3) is not important.

Theorem 2.1 : Every permutation σ of a finite set S is a product of disjoint cycles.

Proof : Without loss of generality we may assume that $S = \{1, 2, 3, \ldots n\}$. Now we consider the elements 1, $\sigma(1)$, $\sigma^2(1)$, $\sigma^3(1)$, ...

Since the set S is finite, these elements cannot all be distinct. Let $\sigma^r(1)$ be the first term in the sequence that has appeared previously. Then $\sigma^r(1) = 1$, for if $\sigma^r(1) = \sigma^r(1)$ with $0 < s < r$, we would have $\sigma^{r-s}(1) = 1$, with $r - s < r$, which is a contradiction to the assumption for r.

Let $\tau_1 = (1, \sigma(1), \sigma^2(1), \ldots \sigma^{r-1}(1))$.

We see that τ_1 has the same effect as σ on all elements of S appearing in this cyclic notation for τ_1. Let i be the first element of S not appearing in this cyclic notation for τ_1. Repeating the above argument with the sequence i, $\sigma(i)$, $\sigma^2(i)$ ... , we arrive at a cycle τ_2. Now, τ_2 and τ_1 are disjoint, for, if they had any element j of S in common, they would be identical, since each cycle could be constructed by repeated application of permutation astarting at j.

Continuing, we pick the first element in S not appearing in the cyclic notation of either τ_1 or τ_2 and construct τ_3, and so on. Since S is finite, this process must terminate with some τ_m. The product

$$\tau_1\ \tau_2 \ldots\ldots\ldots\ldots \tau_m$$

ihen clearly has the same effect on each element of S as σ does, so

$$\sigma = \tau_1\ \tau_2 \ldots\ldots\ldots\ldots \tau_m.$$

Def. 2.4 : A cycle of length 2 is a transposition. Thus, a transposition leaves all elements except two fixed and maps each of these onto the other. It may be easily shown that $(a_1, a_2)\ (a_1, a_3) \ldots (a_1, a_n) = (a_1, a_2, \ldots a_n)$

Therefore, any cycle is a product of transpositions. So, any permutation of a finite set of at leal two elements is a product of transpositions.

Example 2.2 : Express the following permutations as product of disjoint cycles :

(i) $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 5 & 1 & 6 & 4 & 3 \end{pmatrix}$

(ii) (4 2 1 5) (3 4 2 6) (5 6 7 1)

Solution : (i) Starting with 1 we see that in the permutation 1 goes to 2, 2 goes to 5, 5 to 4, 4 to 6, 6 to 3 and 3 to 1. Hence the permutation is just a cyclic one of length 6.

(ii) Computing the product of the cycles, the given permutation can be written as

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 4 & 6 & 2 & 1 & 3 & 7 & 5 \end{pmatrix}$$

Stalling with 1 we get a cycle (1 4). Then starting with another element, not in (1 4), say 2, we get the cycle (2 6 7 5 3). These two cycles exhaust the list {1, 2, 3, 4, 5, 6, 7}. Hence σ = (1 4) (2 6 7 5 3) = (2 6 7 5 3) (1 4).

Def. 2.5 : A permutation of a finite set is even or odd according as whether it can be expressed as the product of an even number of transpositions or the product of an odd number of transpositions respectively.

In this context we have the following theorem.

Theorem 2.2 : No permutation of a finite set can be expressed both as a product of an even number of transpositions and as a product of an odd number of transpositions.

Proof : Let us consider the polynomial in n-variables :

$$p(x_1, x_2, \dots x_n) = \prod_{i<j} (x_i - x_j)$$

Let S_n be the set of all permutations of the n-variables $x_1, x_2, \dots, x_n$. Let σ be a permutation in S_n. Let us apply σ on the polynomial p $(x_1, x_2, \dots, x_n)$ [Here applying σ on the polynomial p $(x_1, x_2, \dots, x_n)$ means that if σ takes x_i to x_j, then we change x_i in p to x_j].

Thus σ changes polynomial

It is clear that polynomial p (x_1, x_2, ... x_n) will become polynomial $\pm p(x_1, x_2, ..., x_n)$.

For instance, in S_5, σ = (1 3 4) (2 5) takes

$$p(x_1, x_2, ..., x_5) = (x_1 - x_2)(x_1 - x_3)(x_1 - x_4)(x_1 - x_5)(x_2 - x_3)$$
$$\times (x_2 - x_4)(x_2 - x_5)(x_3 - x_4)(x_3 - x_5)(x_4 - x_5)$$

into $(x_3 - x_5)(x_5 - x_4)(x_3 - x_1)(x_3 - x_2)(x_5 - x_4)(x_5 - x_1)(x_5 - x_2)$
$\times (x_4 - x_1)(x_4 - x_2)(x_1 - x_2)$

which can easily be verified to be $- p$ (x_1, x_2, ..., x_5).

If, in particular, σ is a transposition, σ makes p (x_1, x_2, ..., x_5) to $- p$ (x_1, x_2, ..., x_5).

Thus, if a permuurion σ can be represented as a product of an even number of transpositions in one representation σ must leave p (x_1, x_2, ..., x_n) fixed, so that any representation of σ as a product of transpositions must be such that it leaves p (x_1, x_2, ..., x_n) fixed; that is, in any representation it is a product of an even number of transpositions.

Hence,

1. The product of two even permutations is an even permutation.
2. The product of an even permutation and an odd permutation is an odd permutation.
3. The product of two odd permutations is an even permutation.

Example 2.3 : Determine whether the permutation σ on the set {1,2, 3, 4, 5, 6} is odd or even, where

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 3 & 5 & 6 & 1 & 4 \end{pmatrix}.$$

Solution : The permutation σ can be expressed as the product of two disjoint cycles (1 2 3 5) (4 6). The cycle (1 2 3 5) can be expressed as (1 5), (I 3), (1 2). Therefore,

$$\sigma = (1\ \ 5)\ (1\ \ 3)(1\ \ 2)\ (4\ \ 6)$$

Hence σ is an even permutation.

2.2 Groups of permutations :

We have already defined that a permutation of a set S is a one-to-one function from S onto S. We have also defined the permutation multiptitation on the set of permucations of a set S. Let σ and τ be the permutations of S, so that σ and τ are both one-to-one functions from S onto S. Now the composite function $\sigma\,\tau$ will be a permutation if it is one-to-one and onto from S to S.

Let a_1, $a_2 \in S$ such that $\sigma\tau(a_1) = \sigma\tau(a_2)$

Then $\sigma(\tau(a_1)) = \sigma(\tau(a_2))$

But σ is a permutation, so it is one-to-one mapping from S to S, hance $\tau(a_1) = \tau(a_2)$. Again, τ is also a permutation, so $a_1 = a_2$.

Hence $\sigma\tau(a_1) = \sigma\tau(a_2) \Rightarrow a_1 = a_2$ and so $\sigma\tau$ is one-to-one mapping from S to S.

Now, we shall show that $\sigma\tau$ is onto.

Let $a \in S$. Since σ is a permutation, σ onto and so there is $a' \in S$ such that $\sigma(a') = a$. Now, τ is also onto, there is $a'' \in S$ such that $\tau(a'') = a'$.

Then we have $a = \sigma(a') = \sigma(\tau(a'')) = \sigma\tau(a'')$. This shows that the mapping $\sigma\tau$ is onto.

Hence $\sigma\tau$ is also a permutation.

Theorem 2.3 : Let S be a non-empty set and let P be the collection of permutations of S. Then P is a group under permutation multiplication.

Proof : For any two permutations σ and $\tau \in P$ we have proved that $\sigma\tau \in P$.

To establish the associativity criterion. Let σ, τ, μ, $\in P$ and $a \in S$. Then

$$[(\sigma\tau(\,\mu]\ (a) = (\sigma\tau)\ (\mu\,(a)) = \sigma(\tau(\mu\,(a)] = [\sigma(\tau\mu)]\ (a)$$

Thus, $(\sigma\tau)\mu$ and $\sigma(\tau\mu)$ map each $a \in S$ onto the same element and hence they are same permutations. Hence the function composition is associative.

The permutation $\eta \in P$ such that $\eta\,(a) = a$ for all $a \in S$, obviously acts as identity element in P.

For a permutation $a \in P$ such that $\sigma(a) = a'$; a, $a' \in S$, the unique inverse mapping σ^{-1} exists such that $\sigma^{-1}(a') = a$, this is due to the fact that σ is one-to-one and onto.

Thus for every $\sigma \in P$, σ^{-1} exists in P, such that $\sigma\sigma^{-1} = \sigma^{-1}\sigma = \eta$ = the identity mapping. Hence the set P of all permutations of S forms a group under permutation multiplication.

Note : There was nothing in our definition of a permutation to require that the set S be finite. However, we shall consider permutation groups of finite sets.

Def. 2.6 : If S is the finite set $\{1, 2, \ldots n\}$, then the group of ail permutations of S is the symmetric group on n letters, and is denoted by S_n. Clearly, S_n has $n!$ elements.

Let, A_n be the subset of S_n, consisting of all the even permutations of n distinct elements. Since the product of two even permutations is even, A_n must be a subgroup of S_n. We shall show later that A_n is a normal subgroup of S_n. The group A_n is called the alternating group of degree n. We may show that A_n contains $\frac{1}{2}n!$ elements.

2.3 Quotient Groups :

We recall that a subgroup N of a group G is a normal subgroup of G if for every $g \in G$ and $n \in N$, $gng^{-1} \in N$. It follows that the subgroup N of G is a normal subgroup of G if and only if every left coset of N in G is a right coset of N in G.

Let G be a group and N be a normal subgroup of G. Let a, $b \in G$. We consider the right cosets Na and Nb (which are also the left cosets of N as N is normal) of N. We consider the set G/N of all right cosets of N in G and define a composition on this set by the formula

$$(Na)\ (Nb) = Nab \qquad \ldots\ldots (1)$$

Na		Nab
	Nb	

The formula (1) does not represent the simple composition rule of two elements of a group G, rather it represents the rule of composition of two classes Na and Nb of elements of G. Before proceeding further we must show that the rule of composition (1) is well defined. In otherwords, we have to show that if $a_1 \in Na$ and $b_1 \in Nb$ then $a_1b_1 \in Nab$.

Since $a \in Na$, $b \in Nb$ [we recall that the elements of the set Na are obtained by right multiplication of the elements of N by a. Since N is a subgroup, $e \in N$ and hence $ea = a \in Na$], we have $a_1 \in Na \Rightarrow \exists n_1 \in N$ such that $a_1 = n_1a$. Similarly, $b_1 \in Nb \Rightarrow \exists n_2 \in N$ such that $b_1 = n_2b$.

$$\therefore\ a_1b_1 = n_1an_2b = n_1an_2a^{-1}ab$$
$$= n_1(an_2a^{-1})ab$$

Since N is a normal subgroup of G, $an_2a^{-1} \in N$.

Let us write $an_2a^{-1} = n_3 \in N$.

Then $a_1b_1 = n_1n_3ab = n_4ab$, $n_4 = n_1n_3 \in N$

This shows that $a_1b_1 \in Nab$

Thus the rule of composition (1) is well defined.

Let $Na, Nb, Nc \in G/N$; $a, b, c \in G$.

$$\text{Then } [(Na)(Nb)](Nc) = (Nab)(Nc) = N(ab)c = Na(bc)$$
$$= (Na)(Nbc)$$
$$= (Na)\ [(Nb)(Nc)]$$

So the composition rule (1) is associative on G/N. For the identity element e of the group G, we have $N = Ne \in G/N$ and $(Ne)(Na) = Nea = Na$, $\forall a \in G$. This shows that $N = Ne$ is the identity element of G/N.

Finally, $a \in G \Rightarrow a^{-1} \in G$ and we have $(Na)(Na^{-1}) = Naa^{-1} = Ne = N =$ the identity element of G/N. Similarly, $(Na^{-1})(Na) = N$. This shows that every element $Na \in G/N$ has its inverse. So G/N is a group under the composition rule (1).

The group G/N generated from the group G and a normal subgroup N of G through the composition rule (1) is called the **quotient group of G by N** or a **factor group of G.** Clearly, the elements of G are the cosets of N in G. The cosets are called the residue classes of G by N.

In particular, if the composition rule in the group G is addition, then the (right) cosets of ihe normal subgroup N of G are $N + a$, $N + b$ etc. and so the composition rule (1) in G/N will look like

$$(N + a) + (N + b) = N + (a + b)$$

Theorem 2.4 : If G is finite and N is a normal subgroup of G, then $o(G/N) = \dfrac{o(G)}{o(N)}$.

Proof : Since G is finite and N is a subgroup of G, by Lagrange's theorem, $o(N)$ divides $o(G)$. Thus there exists a positive integer K such that

$$o(G) = K.o(N) \qquad \text{........ (2)}$$

Clearly, K is the number of distinct cosets of N in G. Again, since the elements of G/N are the distinct cosets of N in G, it follows that

$$o(G/N) = K$$

Hence $\quad o(G/N) = K = \dfrac{o(G)}{o(N)}$, by (2).

Example 2.4 : The set Z of integers is a group under addition (+). The set 3Z = { ..., –6, –3, 0, 3, 6, ...} is a subgroup of Z. For, if we consider any two elements a and b of 3Z, then as both a and b are multiples of 3, $a - b$ is a multiple of 3. So 3Z is a subgroup of Z. Again, if we take any element $x \in Z$, then $x + a + (-x) = a \in 3Z$. This shows that 3Z is a normal subgroup of Z. There are three cosets of 3Z. These are 3Z + 0 (= 3Z itself), 3Z + 1 and 3Z + 2. The elements of 3Z + 1 are {..., –6 +1, 0 + 1, 3 + 1, 6 + 1, ...} we may easily verify that Z/3Z is a group with three residue classes 3Z + 0, 3Z + 1 and 3Z + 2. The group Z/3Z is a quotient group of Z by 3Z.

2.4 Generators of a subgroup :

We have already encountered the idea of a generator in the case of a cyclic group. Now, we shall try to generalize the concept of generators in the case of arbitray groups.

Def. 2.7 : Let 5 be a subset of a group G. A subgroup //of G is said to be generated by S if the following conditions are satisfied :

(i) $S \subseteq H$

(ii) If K is any subgroup of G such that $S \subseteq K$, then $H \subseteq K$

Subgroup generated by S will be denoted by $< S >$.

Theorem 2.5 : If S is any non-void subset of a group G, then ihe subgroup $<S>$ of G generated by S is the set of all **finite** products of the form $a_1a_2 \dots a_n$ where for each i, either $a_i \in S$ or $a_i^{-1} \in S$.

Proof : Let H be the set of all finite products of the form $a_1a_2 \dots a_n$ such that either a_i or a_i^{-1} belongs to S and n is any positive integer. Let $x, y \in H$, so that $x = a_1a_2 \dots a_n$, $y = b_1b_2 \dots b_m$, where either b_j or b_j^{-1}, $(j = 1, 2, \dots, m)$ belongs to S. Then

$$\begin{aligned} xy^{-1} &= a_1a_2 \dots a_n (b_1b_2 \dots b_m)^{-1} \\ &= a_2b_2 \dots a_nb_m^{-1}b_{m-1}^{-1} \dots b_2^{-1}b_1^{-1}. \end{aligned}$$

Since either a_i or $a_i^{-1} \in S$ and b_j or $b_j^{-1} \in S$, if follows that $xy^{-1} \in H$. This shows that H is a subgroup of G. Let K be any other subgroup of G such that $S \subseteq K$. Then $a \in K$ and hence $a^{-1} \in K$ as K is a subgroup of G. Thus, if $x = a_1a_2 \dots a_n$, where either $a_i \in S$ or $a_i^{-1} \in S$, is any element of H, then since $a_i \in K$, we have $x \in K$. Hence $H \subseteq K$. This proves that H is the subgroup of G generated by S.

Def. 2.8 : Let G be a group. For any $a, b \in G$, the element $aba^{-1}b^{-1}$ is called a commutator in G. The subgroup of G generated by the set of all commutators in G is called the commutator subgroup of G or the derived group of G and is denoted by G'.

Theorem 2.6 : Let G be a group and let G' be the derived group (commutator subgroup) of G. Then

(i) $G' \Delta G$, (i.e., G' is a normal subgroup of G)

(ii) G/G' is abelian

(iii) If $H \Delta G$ then G/H is abelian if an only if $G' \subset H$.

Proof : (i) Let $x = aba^{-1}b^{-1}$ be any commutator in G $(a, b \in G)$. Then $x^{-1} = bab^{-1}a^{-1}$ is also a commutator. Also, for any $g \in G$.

$$\begin{aligned} gxg^{-1} &= gaba^{-1}b^{-1}g^{-1} \\ &= (gag^{-1})(gbg^{-1})(ga^{-1}g^{-1})(gb^{-1}g^{-1}) \\ &= (gag^{-1})(gbg^{-1})(gag^{-1})(gbg^{-1})^{-1} \\ &= pqp^{-1}q^{-1};\ p, q \in G \end{aligned}$$

This shows that gxg^{-1} is a commutator and $gxg^{-1} \in G'$.

Now, since G' is the subgroup of G generated by the set of commutators, any element $y \in G'$ is the product of finite number of commutators, say $x_1, x_2, \dots, x_n$.

So, $y = x_1x_2 \ldots x_n, x_i$ is a commutator. Then for any $g \in G$

$$gyg^{-1} = gx_1x_2 \ldots x_ng^{-1}$$
$$= (gx_1g^{-1})(gx_2g^{-1}) \ldots (gx_ng^{-1})$$

As we have already proved that $gx_ig^{-1} \in G'$, it follows thai $gyg^{-1} \in G'$ for any $y \in G'$. Hence G' is a normal subgroup of G i.e., $G'\Delta G$.

(ii) For all $a, b \in G$, we have

$$(aG')(bG')(aG')^{-1}(bG')^{-1} = (aba^{-1}b^{-1})G'$$

Since $aba^{-1}b^{-1} \in G'$, we have $(aba^{-1}b^{-1})G' = G'$.

$$\therefore (aG')(bG')(aG')^{-1}(bG')^{-1} = G' \qquad \ldots\ldots\ldots (1)$$

We recall that G' is the identity element of the quotient grovp G/G', whose elements are the cosets aG', bG' etc. Right multiplication of (1) by $(bG)(aG')$ gives

$$(aG')(bG') = G'(bG')(aG')$$
$$\text{or, } (aG')(bG') = (bG')(aG')$$

This shows that the quotient group G/G' is abelian.

(iii) Suppose that G/H is abelian for a subgroup H of G. Then for $a, b \in G$, we have

$$(aba^{-1}b^{-1})H = (aH)(bH)(aH)^{-1}(bH)^{-1}$$
$$= (aH)(aH)^{-1}(bH)(bH)^{-1} \quad (\because G/H \text{ is abelian})$$
$$= H, \text{ the identity element of } G/H.$$

This shows that the commutator $aba^{-1}b^{-1} \in H$. But $aba^{-1}b^{-1}$ is an element of the derived group G'; hence $G' \subseteq H$.

Conversely, suppose that $G' \subseteq H$. Now, for any $a, b \in G$, $aba^{-1}b^{-1} \in G'$. Since $G' \subseteq H$ by assumption, $aba^{-1}b^{-1} \in H$. This implies that

$$(aba^{-1}b^{-1})H = H$$
$$\Rightarrow (aH)(bH)(aH)^{-1}(bH)^{-1} = H$$
$$\Rightarrow (aH)(bH) = (bH)(aH)$$

Hence G/H is abelian.

Example 2.5 : Let A be set of all transpositions in the symmetric group S_n. Since every $\sigma \in S_n$ is a product of transpositions, we find that A geneates S_n.

Def. 2.9 : The centre of a group G is the set of those elements in G that commute with every element in G. It is usually denoted by $Z(G)$. Thus,

$$Z(G) = \{a \in G \mid ax = xa \ \ \forall x \in G\}$$

Ex. 2.6 : Prove that the centre $Z(G)$ of a group G is a normal subgroup ol G.

Proof : Since $ex = xe \ \forall x \in G$, $e \in Z(G)$, i.e., $Z(G)$ is nonempty. Let a, $b \in Z(G)$. Then for all $x \in G$, $ab^{-1}x = ab^{-1}xe = ab^{-1}xbb^{-1} = ab^{-1}bxb^{-1}$ (since $b \in Z(G)$, $xb = bx$) $= axb^{-1} = xab^{-1}$

or, $ab^{-1}x = xab^{-1} \ \ \forall x \in G$

Hence $ab^{-1} \in Z(G)$. Therefore $Z(G)$ is a subgroup of G.

To prove that $Z(G)$ is a normal subgroup of G, we have to prove that for any $g \in G$ and any $a \in Z(G)$, $gag^{-1} \in Z(G)$.

Since $a \in Z(G)$, a commutes with every element of G.

$\therefore \ ag = ga$

Then $gag^{-1} = agg^{-1} = a$

But $a \in Z(G)$, so $gag^{-1} \in Z(G)$ for any $g \in G$

Hence $Z(G)$ is a normal subgroup of G.

Ex. 2.7 : Le H and K be normal subgroups of a group G and let $HK = \{hk \mid h \in H$ and $k \in K\}$. Show that HK is a normal subgroup of G.

Soln. : First of all we shall show that HK is a subgroup of G. We consider any two elements h_1k_1 and h_2k_2 of the set HK. Then

$$h_1k_1(h_2k_2)^{-1} = h_1k_1k_2^{-1}h_2^{-1} = h_1h_2^{-1}h_2(k_1k_2^{-1})h_2^{-1}$$

Since $k_1k_2^{-1} \in K$ and K is a normal subgroup of G, $h_2(k_1k_2^{-1})h_2^{-1} \in K$. Let us denote this element by k_3 and also the element $h_1h_2^{-1} \in H$. Denoting this element by h_3 we see that

$$h_1k_1(h_2k_2)^{-1} = h_3k_3 \in HK$$

Hence HK is a subgroup of G.

Now, to show that HK is a normal subgroup of G, we take any element $hk \in HK$ and any element $g \in G$. Then

$$ghkg^{-1} = (ghg^{-1}) \ (gkg^{-1})$$

Since H is a normal subgroup of G, $ghg^{-1} \in H$ and since K is a normal subgroup of G, $gkg^{-1} \in K$, so $ghKg^{-1} \in HK$. Hence HK is a normal subgroup of G.

Ex. 2.8 : If a group G has only one element a of order n, then $a \in Z(G)$ and $n = 2$.

Soln. : Given that a is the only element for which $o(a) = n$. To show that $a \in Z(G)$, we have to show that the element a commutes with every eiemeni of G. Let $x \in G$. Then

$$(xax^{-1})^n = (xax^{-1})\ (xax^{-1})\ \ldots\ (xax^{-1})$$

$$= xa^nx^{-1} = xex^{-1} = e$$

But $(xax^{-1})^m \neq e$ for $m < n$

Hence $o(xax^{-1}) = n$

But a is the only element of order n.

Hence it follows that

$$xax^{-1} = a$$

or, $xa = ax \quad \forall x \in G$

$\Rightarrow \quad a \in Z(G)$

To prove the second part, assume that $n > 2$. There exists a positive integer m (>1) relatively prime to n. Then $o(a^m) = o(a) = n$. But this contradicts the fact that a is the only element having order n. Hence $n = 2$.

EXERCISES

1. Prove that two right cosets Ha and Hb of a subgroup H of a group G are identical if and only it $ab^{-1} \in H$.
2. Let H be a subgroup of a group G. Prove that $Ha = H$ if and only if $a \in H$.
3. If H and K are subgroups whose orders are relatively prime, prove that $H \in K = \{e\}$.
4. Le G be a group in which, for some integer $n > 1$, $(ab)^n = a^nb^n$ for all $a, b \in G$. Show that $G^{(n)} = \{x^n | x \in G\}$ is a normal subgroup of G.

5. Show that if G is a group of even order then there are exactly an odd number of elements of order 2. [Hint : $x^2 \neq e \Rightarrow x \neq x^{-1} \Rightarrow$ there are even number of elements satisfying $x^2 \neq e \Rightarrow$ there are even number of elements satisf .ing $x^2 = e$].

6. Let a be an element of a group G such that $o(a) = r$. Let m be a positive integer. Prove $o(a^m) = r/(m, r)$ where $(m, r) = ged$ of m and r.

7. Which of the following permutations are even ?

(a) (4, 5, 6)

(b) (1, 2) (7, 8, 9, 10)

(c) (1, 2) (1, 3) (1, 4) (1, 5) (1, 6)

(d) (1, 2, 3) (4, 5, 6) (7, 8, 9)

Unit : 3 □ Morphisms of groups

Morphism is an abbreviation of homomorphism. Homomorphism is basically a mapping from one algebraic system lo a like algebraic system which preserves structure. In the case of groups we define homomorphism below.

Def. 3.1 : A mapping ϕ from a group $< G, o >$ into a group $< G', o' >$ is said to be a homomorphism if for all $a, b \varepsilon G$, $\phi(aob) = \phi(a)\ o'\phi(b)$ ...... (1)

Keeping the rules of composition in groups G and G' in mind, the condition (1) is usually expressed in the form

$$\phi(ab) = \phi(a)\phi(b) \qquad \text{...... (2)}$$

Examples 3.1 :

(1) Let $\phi : G \to G'$ be defined by

$$\phi(x) = e' \forall x \varepsilon G.$$

Then dearly, for $a, b \varepsilon G$, $\phi(a) = \phi(b) = e'$

Also $\phi(ab) = e'$

$\therefore \quad \phi(ab) = e' = e'e' = \phi(a)\phi(b)$

Thus, ϕ is a homomorphism from G to G'.

(2) Let G be the group of all real numbers under addition and G' be the group of non-zero real numbers under multiplication.

Let $\phi : G \to G'$ be such that

$$\phi(a) = 2^a,\ a \varepsilon G.$$

Then we have for $a, b \varepsilon G$, $a + b \varepsilon G$ and so $\phi(a + b) = 2^{a+b} = 2^a.2^b = \phi(a)\phi(b)$

Thus, ϕ is a homomorphism from G to G'.

(3) Let G be the group of integers under addition and let G'_n be the group of integers under addition modulo n. Let $\phi : G \to G'_n$ be such that $\phi(x)$ = remainder of x on division by n. We can easily verify that ϕ is a homomorphism.

The homomorphism condition (1) or (2) may be represented diagrammatically as shown below :

$$\begin{tikzcd}
G \times G \arrow[r, "o"] \arrow[d, "\phi \times \phi"'] & G \arrow[d] \\
G' \times G' \arrow[r, "o'"] & G'
\end{tikzcd}$$

If the mapping $\phi : G \rightarrow G'$ is one-one (i.e., injective) the morphism is called monomorphism; if ϕ is onto (i.e., surjective), the morphism is called epimorphism and if the mapping ϕ is one-one and onto (i.e., bijective) the morphism is termed as isomorphism.

If ϕ is a mapping from a group G to G itself such that ϕ is a homomorphism and ϕ is bijeciive, then ϕ is called an automorphism. Thus, automorphism is an isomorphism fiom the group $<G, o>$ onto itself.

Example 3.2 : Let G be the group of integers under addition and T be the mapping from to G such that $T(x) = -x$, $x \varepsilon G$. Examine whether T is an automorphism of G.

We have for $a, b \varepsilon G$, $a + b \varepsilon G$ and

$$\begin{aligned} T(a+b) &= -(a+b) = -a-b = (-a) + (-b) \\ &= T(a) + T(b) \end{aligned}$$

Thus T is a homomorphism.

To examine T for automorphism, we have to test whether T is bijective.

$$\begin{aligned} \text{Clearly, } T(a) + T(b) &\Rightarrow T(a) - T(b) = 0 \\ &\Rightarrow T(a-b) = 0 \\ &\Rightarrow -a + b = 0 \qquad \text{or, } a = b \end{aligned}$$

Thus T is one-one

Again, for every $a \varepsilon G$, we have $-a \varepsilon G$ such $T(-a) = -(-a) = a$. So T is onto. Hence T is an automorphism of G.

Example 3.3 : Let C be the group of complex numbers under addition. Consider the mapping $\phi : C \rightarrow C$ defined as $\phi(a + ib) = a - bi$, a and b are real numbers. Show that ϕ is an automorphism on C.

Solution : Let $x = a + ib$ and $y = c + id$ be two elements of C; a, b, c, d being real numbers. Then

$$\begin{aligned}\phi(x + y) &= \phi((a + ib) + (c + id)) = \phi(a + c + i(b + d)) \\ &= a + c - i(b + d) = (a - ib) + (c - id) \\ &= \phi(x) + \phi(y)\end{aligned}$$

So ϕ is a homomorphism.

$$\begin{aligned}\text{Now, } \phi(x) = \phi(y) &\Rightarrow a - ib = c - id \\ &\Rightarrow a = c \text{ and } b = d \\ &\Rightarrow a + ib = c + id \\ &\Rightarrow x = y\end{aligned}$$

$\therefore$ ϕ is one-one.

Again, for every $z = p + iq \varepsilon C$, there exist $t = p - iq \ \varepsilon C$ such that $\phi(t) = \phi(p - iq) = p + iq = z$. Hence ϕ is onto.

So ϕ is a one-one and onto homomorphism from C to C. In other words ϕ is an automorphism on C.

Theorem 3.1 : Suppose G is a group and N a normal subgroup of G. Define the mapping ϕ from G to G/N by $\phi(x) = Nx$, $\forall x \varepsilon G$. Then ϕ is a homomorphism of G onto G/N.

Proof. : We have already proved that G/N is a group. Clearly the mapping ϕ : $G \rightarrow G/N$ is onto. For, any element $y \varepsilon G/N$ is of the form $y = Nx$, for some $x \varepsilon G$. Hence for each $y \varepsilon G/N$, $\exists x \ \varepsilon G$ such that $\varphi(x) = y$. So ϕ is onto. Now for a, $b \varepsilon G$

$$\phi(ab) = N(ab) = (Na)\ (Nb) = \varphi(a)\ \phi(b)$$

Hence ϕ is a homomorphism.

The homomorphism $G \rightarrow G/N$ is called the natural (or canonical) homomorphism of G onto G/N.

Def. 3.2 : (Kernel of homomorphism). If ϕ is a homomorphism of a group G into a group G' then the kernel of ϕ, denoted by K_ϕ is defined by

$$K_\phi = \{x \varepsilon G / \varphi(x) = e', \text{ the identify element of } G'\}$$

In the natural honomorphism $G \rightarrow G/N$, since for any $n\varepsilon N$ we have $\phi(n) = Nn = N =$ the identity element of G/N, it follows that N is the kernel of honomorphism.

Theorem 3.2. If ϕ is a honomerphism of a group G into a group G', then

(i) $\phi(e) = e'$, e and e' are the identity elements of G and G' respectively

(ii) $\phi(x^{-1}) = [\phi(x)]^{-1} \forall x \varepsilon G$.

Proof. : We have $\phi(x)e' = \phi(x) = \phi(xe) = \phi(x)\varphi(e)$

By cancellation rule, $\phi(e) = e'$. This proves (i)

To prove part (ii), we have $e' = \phi(e) = \phi(xx^{-1})$

or, $e' = \phi(x)\ \phi(x^{-1})$

$\Rightarrow \phi(x^{-1}) = [\phi(x)]^{-1}$

Theorem 3.3 : A honomorphism $\phi : G \rightarrow G'$ is injective if and only if $K_\phi = \{e\}$.

Proof. : Suppose ϕ is injective and let $x\varepsilon K_\phi$.

Then $\phi(x) = e' = \phi(e)$

Since ϕ is injective, we must have $x = e$

Hence $K_\phi = \{e\}$

Conversely, suppose that $K_\phi = \{e\}$. Then for $x.y\varepsilon G$, $\phi(x) = \phi(y)$

$\Rightarrow \phi(x)\ [\phi(y)]^{-1} = e'$

$\Rightarrow \phi(x)\ [\phi(y^{-1}) = e'$

$\Rightarrow \phi(xy^{-1}) = e'$

$\Rightarrow xy^{-1}\varepsilon K_\phi$

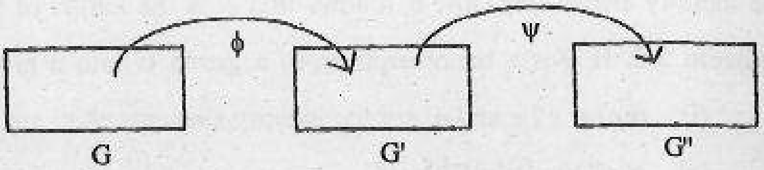
But $K_\phi = \{e\}$

Hence $xy^{-1} = e \Rightarrow x = y$

So ϕ is injective.

Example 3.4 : Prove that if $\phi : G \rightarrow G'$ and $\psi : G' \rightarrow G''$ are homomorphisms of groups, then $\psi\phi : G \rightarrow G''$ is also a homomorphism.

Solution :



Let x, $y \varepsilon G$, then since φ is a homomorphism $\phi(xy) = \phi(x)\ \varphi(y)$

Now, $\psi\phi(xy) = \psi(\phi(xy))$

$$= \psi(\phi(x)\ \phi(y))$$

$$= \psi(\phi(x)\ \psi(\varphi(y)) \qquad [\because\ \psi \text{ is a homomorphism}]$$

$$= \psi\phi(x)\ \psi\varphi(y)$$

This shows that $\psi\phi$ is a homomorphism from G to G''.

Theorem 3.4 : Let $\phi : G \rightarrow G'$ be a homomorphism of groups. Then K_ϕ is a normal subgroup of G and the image of G under ϕ (denoted by $I_{m\phi}$) is a subgroup of G'.

Proof. : Since G and G' are groups and $\varphi(e) = e'$, it follows that $e\, \varepsilon\, K_\phi$ and so K_ϕ and $I_m\phi$ are both non empty. Let a, $b \varepsilon\ K_\phi$, then $\phi(a) = \phi(b) = e'$.

Then $\varphi(ab^{-1}) = \varphi(a)\ \varphi(b^{-1}) = \varphi(a)\ [\phi(b)]^{-1} = e'.e'^{-1} = e'$

This implies that $ab^{-1}\ \varepsilon\ K_\phi$ and so K_ϕ is a subgroup of G.

Again let $a \varepsilon\ K_\phi$ and $g \varepsilon G$.

Then we have $\phi(g^{-1}ag) = \phi(g^{-1})\ \phi(a)\ \phi(g) = [\varphi(g)]^{-1}e'\phi(g) = e'$

$\therefore\ g^{-1}\ ag\ \varepsilon\ K_\phi$ for every $g \varepsilon G$ and $a \varepsilon K_\phi$

Hence K_ϕ is a normal subgroup of G.

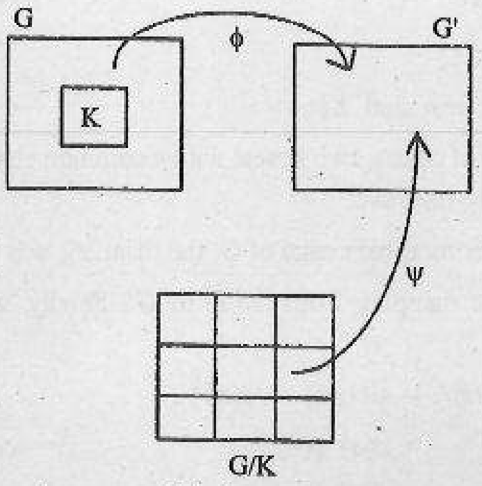
To prove the second part, let x, $y \varepsilon\, I_m\phi$. Il follows from the definition of the mapping ϕ that there exist elements a, $b \varepsilon G$ such that

$$\varphi(a) = x \text{ and } \phi(b) = y$$

Then $xy^{-1} = \varphi(a)\ [\varphi(b)]^{-1} = \phi(a)\ \phi(b^{-1}) = \phi(ab^{-1})$

But since $ab^{-1} \varepsilon G$, $\phi(ab^{-1})\ \varepsilon\, I_m\phi \Rightarrow xy^{-1}\ \varepsilon I_m\phi$ for x, $y \varepsilon I_m\phi$. Hence $I_m\phi$ is a subgroup of G'.

Theorem 3.5 : (Fundamental Theorem of homomorphism) Every homomorphic image of a group G is isomorphic to the quotient group G / K of G by K, where K is the kernel of homomorphism.



Proof. : Let G' be a homomorphic image of a group G under the mapping ϕ : $G \to G'$. If K is the kernel of ϕ, we have to prove that G / K is isomorphic to G'.

In symbol $G / K \cong G'$.

Consider the mapping $\psi : G / k \to G'$ defined by $\psi(ka) = \phi(a)$, $a \varepsilon G$.

Since the (ψ-image of a coset depends upon the φ-image of a member of the coset, we have to show that the mapping ψ is well-defined, For this we require that the image of each member of Ka is the same as the ϕ-image of a. In other words, ψ will map the whole class Ka to the ϕ-image of a.

Let $k \varepsilon K$, then ka is any member of Ka

Now, $\phi(ka) = \phi(k)\ \phi(a)$ [$\because$ ϕ is a homomorphism]

or, $\phi(ka) = e'\phi(a) = \phi(a)$

Since ka is any element of Ka, it follows that all elements of Ka will map on the image of a. So the mapping is well-defined.

We shall now show that ψ is one-one (injective).

Assume that $\psi(Ka) = \psi(Kb)$; a, $b \varepsilon G$

$\Rightarrow \phi(a) = \phi(b)$

$$\Rightarrow \phi(a)\,[\phi(b)]^{-1} = e'$$

$$\Rightarrow \phi(ab^{-1}) = e'$$

$$\Rightarrow ab^{-1} \varepsilon K$$

$$\Rightarrow a\ \varepsilon Kb$$

But $a \varepsilon Ka$. So a εKa and Kb.

From the property of cosets, two cosets with a common element are identical. Hence $Ka = Kb$ and so ψ is injective.

Since G' is the homomorphic image of G, the mapping ψ is clearly onto (subjective).

$\therefore$ ψ is a bijective mapping from G/K to G'. Finally, we shall show that ψ is a homomorphism.

We have, $\psi(ka)\ (kb) = \psi(kab) = \phi(ab)$

$$= \phi(a)\ \phi(b)$$

$$= \psi(ka)\ \psi(kb)$$

Thus ψ is an isomorphism from G/K to G' and hence $G/K \cong G'$. Proved.

Theorem 3.6 : (Cayley) Every group is isomorphic to a permutation group.

Proof. : Let G be a group and let S_G denote the group of all permutations of G called the symmetric group on the set G. For each $a \varepsilon G$ we define mapping.

$$f_a : G \rightarrow G \text{ by}$$

$$f_a(x) = ax, \ \ \forall x \varepsilon G.$$

Thus for $x, y \varepsilon G$, $f_a(x) = f_a(y) \Rightarrow ax = ay$

$$\Rightarrow x = y.$$

So the mapping f_a is one-one.

Now, for any $x \varepsilon G$, we have $f_a(a^{-1}x) = a(a^{-1}x) = x$

That is, for any $x \varepsilon G$, there exists $a^{-1}x$ in G such that $f_a(a^{-1}x) = x$. Hence the mapping f_a is onto. Hence f_a is a permutation of G. So $f_a\ \varepsilon S_G$.

Now, for any a, b, $x \varepsilon G$

$$(f_a f_b)\ (x) = f_a(f_b(x)) = f_a(bx) = abx$$

$$= (ab)x = f_{ab}(x)$$

Hence $f_a f_b = f_{ab}$

Let us define a mapping $\varphi : G \to S_G$ by $\varphi(a) = f_a$, $a \varepsilon G$

Then for all $a', b \varepsilon G$,

$$\varphi(ab) = f_{ab} = f_a f_b = \varphi(a)\ \varphi(b)$$

$\therefore\ \phi$ is a homomorphism from $G \to S_G$

Moreover, $\varphi(a) = \varphi(b)$

$$\Rightarrow f_a = f_b$$

$$\Rightarrow f_a(e) = f_b(e)$$

$$\Rightarrow ae = be$$

$$\Rightarrow a = b$$

Therefore, ϕ is one-one homorphsms of G into S_G. Hence G is isomorphic to the image $\phi(G)$ of G under ϕ and $\phi(G)$ being a subgroup of S_G, is a permutation group. Hence the group G is isomorphic to a permutation group. This proves the theorem.

Example 3.5 : Let R^* be the group of non zero real numbers under multiplication. Show that the mapping ϕ from R^* to R^* defined by $\varphi(x) = |x|$ is a homomorphism. Find K_φ

Solution : Let $x, y \varepsilon R^*$, then $\phi(xy) = |xy| = |x|\ |y| = \varphi(x)\ \phi(y)$

So ϕ is a homomorphism.

Since 1 is the identify element of the group R^* under multiplication, we see that

$$\varphi(1) = 1 \text{ and } \phi(-1) = 1$$

and there is no other element mapping onto 1 Hence $K_\varphi = \{1, -1\}$.

Example 3.6 : Let $R[x]$ denote the group of all polynomials with real coefficients under addition. For any $f(x)\ \varepsilon\ R[x]$ let $f'(x)$ be the derivative of $f(x)$. Let $\phi : R[x] \to R[x]$ be defined by $\phi(f) = f'$.

Prove that ϕ is a homomorphism and find K_ϕ

Solution : Let $f(x)$ and $g(x)\ \varepsilon R[x]$. Since $R[x]$ is a group under addition, we have $f + g \varepsilon R[x]$ and $\phi(f + g) = (f + g)' = f'(x) + g'(x) = \phi(f) + \phi(g)$.

Hence ϕ is a homomorphism.

Since 0 is the identity element of $R[x]$, we have for any constant polynomial c $\varepsilon R[x]$

$$\phi(c) = c' = \frac{d}{dx}(\text{const}) = 0$$

This shows that K_ϕ is the set of all constant polynomials.

Def. 3.3 : Let G be a group and let $a \varepsilon G$. The function φ_a defined by $\varphi_a(x) = axa^{-1}$, $\forall x \varepsilon G$ is called the **inner automorphism of** G induced by a.

We may show that the mapping $\phi_a : G \rightarrow G$ is an automorphism of G.

For this, first we show that ϕ_a is one-one.

We have $\phi_a(x) = \varphi_a(y)$

$$\Rightarrow axa^{-1} = aya^{-1}$$

$$\Rightarrow x = y, \text{ by cancellation law}$$

$\therefore$ φ_a is one-one

Also, $\varphi_a(a^{-1}xa) = a(a^{-1}xa)a^{-1} = x$

This shows that for every $x \varepsilon G$, there is an element $a^{-1}xa$ εG such that $\varphi_a(a^{-1}xa) = x$. Hence φ_a is onto.

Finally, $\varphi_a(xy) = a(xy)a^{-1} = axa^{-1}\ aya^{-1} = \phi(x)\ \phi(y)$

$\therefore$ φ_a is a homomorphism.

Hence φ_a is an automorphism of G.

We shall denote the set of all automorphisms of the group G by Aut (G) and the set of all inner automorphisms of G by $I_{nn}(G)$.

Theorem 3.7 : The set of automorphisms of a group G (i.e., Aut (G)) forms a group under the function composition operation.

Proof. : It is obvious that the identity mapping on G defined by $\varphi(x) = x$ is an automorphism of G. Hence Aut (G) is not empty.

Let f and g ε Aut (G). Then as f and g are both one-one mappings of G onto G, fog is also one-one mapping of G onto G. Now, for a, $b \varepsilon G$,

$$(fog)\ (ab) = f[g(ab)]$$

$= f[g(a)\ g(b)]$, since g is a homomorphism

$= f(g(a))\ f(g(b))$, since f is a homomorphism.

$= fog(a)\ fog(b)$

Thus fog is an automorphism of G and hence $fog \ \varepsilon$ Aut (G).

Since the resultant composition for mapping is in general associative, the composition in Aut (G) is associative. The identity mapping is clearly the identity element in Aut (G).

Since f is one-one mapping from G to itself, f^{-1} is defined and f^{-1} is also an injective mapping from G to itself. We now show that $f^{-1}\ \varepsilon$ Aut (G).

Let a, $b \varepsilon G$, then there exist $x.y \varepsilon G$ such that

$f(x) = a$ and $f(y) = b$

$\Rightarrow x = f^{-1}(a)$ and $y = f^{-1}(b)$

Consequently, $ab = f(x)f(y) = f(xy)$, since f is a homomorphism

$= f(f^{-1}(a)f^{-1}(b))$

$\Rightarrow f^{-1}(ab) = f^{-1}(a)f^{-1}(b)$

Hence f^{-1} is an automorphism of G and so $f^{-1}\ \varepsilon$ Aut (G)

Thus, Aut (G) is a group.

Theorem 3.8 : The inner automorphisms of any group G, $I_{nn}(G)$ form a subgroup of the group of all automorphisms, Aut (G).

Proof. : Clearly the identity mapping $\varphi_e(x) = x = exe^{-1}\ \varepsilon\ I_{nn}(G)$.

Let φ_a, $\phi_b\ \varepsilon\ I_{nn}(G)$.

First of all, we show that $(\varphi_b)^{-1}\ \varepsilon\ I_{nn}(G)$.

Since G is a group, for every $y \varepsilon G$, there exists unique $x \varepsilon G$ such that $y = bxb^{-1}$, $b \varepsilon G$, and x is uniquely determined as $x = b^{-1}yb$. From the definition of inner auto morphism, we have $y = \varphi_b(x)$.

Thus the unique inverse $(\phi_b)^{-1}$ of the mapping φ_b exists and

$(\varphi_b)^{-1}(y) = x = b^{-1}y\ b = \phi_b{}^{-1}(y)$, $\forall y \varepsilon G$.

Hence $(\phi_b)^{-1} = \phi_b{}^{-1}$

Now, $\varphi_a(\varphi_b)^{-1}(x) = \varphi_a\varphi_b{}^{-1}(x)$

$$= \varphi_a(b^{-1}xb) = a(b^{-1}xb)_a^{-1}$$
$$= ab^{-1}x(ab^{-1})^{-1} = \varphi_{ab}{}^{-1}(x)$$

This shows that if φ_a, φ_b ε $I_{nn}(G)$,

Then $\varphi_a(\varphi_b)^{-1}$ ε $I_{nn}(G)$. In other words, $I_{nn}(G)$ is a subgroup of Aut (G).

Example 3.7 : Let G be a group and ϕ be an automorphism of G. If $a \varepsilon G$ is of order greater than zero, then $o(\varphi(a)) = o(a)$.

Solution : Let $n(> o)$ be the order of $a \varepsilon G$. Then $a^n = e$ and $a^m \neq e$ for positive integer $m < n$. We have

$$\{\varphi(a)\}^n = \varphi(a)\varphi(a) \ldots \varphi(a). \text{ [n factors]}$$
$$= \varphi(a^n) = \varphi(e) = e.$$

$\therefore$ $o(\varphi(a)) \leq n$.

If $o(\varphi(a)) = m < n$, then

$$e = (\varphi(a))^m = \varphi(a^m)$$

Since ϕ is one-one, it follows that $a^m = e$ for $m < n$, which is a contradiction to the assumption that n is the smallest positive integer such that $a^n = e$.

Hence $o(\varphi(a)) = n = o(a)$

Theorem 3.9 : Let H and N be two subgroups of a group G and N be a normal subgroup of G, then

$$H/H \cap N \cong HN\backslash N$$

Proof. : We consider the subset $HN = \{hn \mid h \varepsilon H, n \varepsilon N\}$

Let h_1n_1, h_2n_2 εHN, then $h_1n_1(h_2n_2)^{-1} = h_1n_1n_2{}^{-1}h_2{}^{-1}$

$$= h_1h_2{}^{-1}(h_2n_1n_2{}^{-1}h_2{}^{-1})$$

Since N is a normal subgroup of G, $h_2(n_1n_2{}^{-1})h_2{}^{-1} \varepsilon N$,

and hence $h_1n_1(h_2n_2)^{-1} \varepsilon HN$.

Thus, HN is a subgroup of G. Since N is a normal subgroup of G, $HN = NH$. We want to show that N is a normal subgroup of HN.

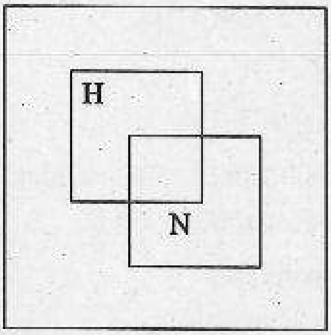
Consider any element h_1n_1 εHN and any element $n \varepsilon N$, then

$$h_1n_1n(h_1n_1)^{-1} = h_1n_1nn_1{}^{-1}h_1{}^{-1}$$

$$= h_1(n_1nn_1^{-1})h_1^{-1} = h_1n_2h_1^{-1},\ n_2 = n_1nn_1^{-1} \varepsilon N$$

Since N is normal subgroup of G, $h_1n_2h_1^{-1} \varepsilon N$ and so N is a normal subgroup of HN.

Let us consider the mapping



$$\varphi : H \to HN \mid N \text{ defined by } \varphi(h) = hN,\ \forall h\ \varepsilon H.$$

Here φ is a restriction of the natural homomorphism $G \to G \mid N$ to the subgroup H. Hence the kernel K_ϕ will be the common elements of H and N, i.e., $K_\phi = H \cap N$.

Moreover, ϕ is clearly surjective. Thus using the fundamental theorem of homomorphism (Th 3.5) and noting that here H replaces G, $H \cap N$ replaces K_φ and $HN \mid N$ replaces the homomorphic image G' of G, we have

$$H \mid H \cap N \cong HN \mid N$$

Theorem 3.10 : Let H and K be normal subgroups of a group G and $K \subset H$. Then

$$(G \mid K) \mid (H \mid K) \cong G \mid H.$$

Proof. : We define the mapping ϕ between the two quotient groups $G \mid K$ and $G \mid H$ of G given by

$$\phi(xK) = xH,\ x \varepsilon G$$

i.e., the left coset of K determined by $x \varepsilon G$ will map onto the left coset of H determined by x.

We note that the mapping ϕ is the mapping of a class of elements of the group G to another class of elements of G, so it is necessary to show that the mapping is well-defined, i.e., we have to show that

$$xK = yK \Rightarrow xH = yH.$$

This means that if two elements x and y belong to the same coset of K, then their images will belong to the same coset of H.

We have, $xK = yK \Rightarrow xk_1 = yk_2,\ k_1,\ k_2 \varepsilon K$

$$\Rightarrow xy^{-1} = k_2k^{-1} \varepsilon K$$

But $K \subset H$. $\quad \therefore\ xy^{-1} \varepsilon H$

$$\Rightarrow xH = yH$$

Hence the mapping ϕ is well-defined. The mapping ϕ is clearly onto, as for each xH in $G \mid H$ we have its pre-image xK in $G \mid K$.

We now test ϕ for homomorphism.

For $\quad x,\ y \varepsilon G$, we have

$$\phi((xK)\ (yK)) = \phi(xyK) = xyH = (xH)\ (yH)$$
$$= \phi(xK)\ \phi(yK)$$

$\therefore\ \phi$ is a homomorphism.

The kernel of ϕ is

$K_\phi = \{xK \mid \varphi(xK) = H\}$, ($H$ is the identity element of the quotient group $G \mid H$)

$$= \{xK \mid xH = H\}$$
$$= \{xK \mid x \varepsilon H\}$$
$$= H \mid K.$$

Using the fundamental theorem of homomorphism for groups, we have,

$$(G \mid K) \mid K_\phi \cong G \mid H$$

i.e., $\quad (G \mid K) | (H \mid K) \cong G \mid H.$

Conjugacy :

Def. 3.4 : Let a, b be the elements of a group G. The element b is said to be a conjugate of the element $a \varepsilon G$ if there exists an element $c \varepsilon G$ such that

$$b = c^{-1}ac$$

If a and b are so related then we ezpress this relation by the symbol $a \sim b$ and we call the relation a conjugate relation.

Theorem 3.11 : The relation of conjugacy is an equivalence relation.

Proof. : Since $a = a^{-1}aa$, we have $a \cdot\cdot a$ $\forall a \varepsilon G$. Hence the relation is reflexive.

Let $a \sim b$. Then there exists $c \varepsilon G$ such that

$$b = c^{-1}ac$$

$$\Rightarrow cbc^{-1} = a \text{ or, } a = (c^{-1})^{-1}b\, c^{-1}$$

Since $c^{-1} \varepsilon G$, we have $b \sim a$

So the relation is symmetric.

Finally, let $a \sim b$ and $b \sim c$, then there exist x and $y \varepsilon G$ such that

$$b = x^{-1}ax,\ c = y^{-1}by$$

Hence $c = y^{-1}(x^{-1}ax)y = (xy)^{-1}a(xy)$

Since $xy \varepsilon G$, $a \sim c$

This means that the relation is transitive. Hence the relation of conjugacy is an equivalence relation.

Since an equivalence relation on a set partitions the set into a number of equivalence classes, it follows that the conjugacy relation partitions the group G into a number of disjoint equivalence classes known as the classes of conjugate elements.

Thus, for $a \varepsilon G$, let $c(a) = \{x \varepsilon G \mid a \sim x\}$.

$c(a)$, the equivalence class of a in G under the relation $\sim$, is usually called the conjugate class of a in G; it consists of the set of all distinct elements of the form $y^{-1}ay$, $y \varepsilon G$.

For a **finite group** G, if $|\, c(a)\,|$ denotes the number of elements in the class $c(a)$, then

$$o(G) = \sum_{a \varepsilon G} |\, c(a)\,|$$

where the summation on the right hand side runs over one element from each conjngacy class.

Def. 3.5 : If $a \varepsilon G$, then the set $N(a) = \{x \varepsilon G \mid xa = ax\}$ is called the normalizer of a in G.

From the definition it follows that $N(a)$ consists of those elements of G which commute with the element $a \varepsilon G$. Since $ea = ae$, then set $N(a)$ is non empty as $e \varepsilon N(a)$.

Lemma : $N(a)$ is a subgroup of G.

Let x, $y \varepsilon N(a)$. Then $xa = ax$, $ya = ay$.

Now, $(xy)a = x(ya) = x(ay) = (xa)y = a(xy) \Rightarrow xy \varepsilon N(a)$

Now, $ax = xa \Rightarrow x^{-1}ax = a$

$$\Rightarrow x^{-1}a = ax^{-1} \quad \therefore \; x^{-1} \varepsilon N(a)$$

Since $x^{-1} \varepsilon N(a)$, $y \varepsilon N(a)$, it follows from the above that $x^{-1}y \varepsilon N(a)$. Hence $N(a)$ is a subgroup of G.

Theorem 3.12 : Two elements x, y of a group G give rise to the same conjugate of an element $a \varepsilon G$ if and only if they belong to the same right coset of the normalizer of a in G.

Proof. : Let $N(a)$ denote the nermalizer of $a \varepsilon G$, then from the Lemma just proved $N(a)$ is a subgroup of G. Let x, y belong to the same right coset of $N(a)$ in G.

Since $y \varepsilon N(a)y$, $x \varepsilon N(a)y$

Now, $x \varepsilon N(a)y \Rightarrow xy^{-1} \varepsilon N(a)$

$$\Rightarrow a(xy^{-1}) = (xy^{-1})a$$

$$\Rightarrow (ax)y^{-1} = x(y^{-1}a)$$

$$\Rightarrow x^{-1}ax = y^{-1}ay$$

This shows that x and y give rise to the same conjugate of an element a.

Conversely, let $x^{-1}ax = y^{-1}ay$

Then $x(x^{-1}ax)y^{-1} = xy^{-1}a$

$$\Rightarrow axy^{-1} = xy^{-1}a$$

$$\Rightarrow xy^{-1} \varepsilon N(a)$$

$$\Rightarrow x \varepsilon N(a)y$$

In other words x belongs to the same right coset as that of y.

Theorem 3.13 : In a finite group G the number of elements in the class $c(a)$ of all elements conjugate to a is the index of the normalizer of a in G, that is,

$$|c(a)| = \frac{o(G)}{o(N(a))} = [G : N(a)]$$

Proof. : By Theorem 3.12, if x, y belong to the same right coset of $N(a)$ in G, then x, y give rise to the same conjugate of an element $a \varepsilon G$ and if xy belong to the different right cosets of $N(a)$ in G then they give rise to different conjugates of a in G. Thus, to each right coset of $N(a)$, there corresponds to a single element conjugate to a. In otherwords, there is one-one correspondence between conjugates of a and the right cosets of $N(a)$. Thus, in a finite group, there are as many conjugates of the element $a \varepsilon G$, as there are distinct right cosets of $N(a)$. But the number of right cosets of $N(a)$ in the finite group G is

$$\frac{o(G)}{o(N(a))} = [G:N(a)]$$

Hence the number of elements conjugate of $a \varepsilon G$ is

$$|c(a)| = \frac{o(G)}{o(N(a))} = [G:N(a)]$$

Since the relation of conjugacy is an equivalence relation, in a finite group G, the equivalence relation of conjugacy will partition G into a number of equivalence classes, viz., a class containing the conjugate elements of a, a class containing the conjugate elements of b and so on.

Thus, $o(G)$ = sum of the elements in different conjugate classes in G.

$$= \sum |c(a)|$$

$$= \sum \frac{o(G)}{o(N(a))} \qquad \text{.........(1)}$$

where the summation runs over one element from each conjugate class of G.

The equation (1) is called the class equation (or class equation formula) of G.

We recall that the center $Z(G)$ of a group G is the set of those elements in G that commute with every element in G, i.e.,

$$Z(G) = \{a \varepsilon G \mid ax = xa, \ \forall x \varepsilon G\}$$

Theorem 3.14 : Let $Z(G)$ be the centre of a group G and let $a \varepsilon G$. Then $a \varepsilon Z(G)$ if and only if $N(a) = G$.

Proof. : We remember that the set $N(a)$ contains those elements of G which commute with $a \varepsilon G$. Now $a \varepsilon Z(G)$ means that the element a commutes will every element of G. In other words every element of G commutes with the element a. Hence $N(a) = G$. conversely, suppose that $N(a) = G$. Then every element of G commutes with a, i.e., $ax = xa \ \ \forall x \varepsilon G$. Hence $a \varepsilon Z(G)$.

Theorem 3.15 : Let $Z(G)$ be the centre of a finite group. Then $o(G) = o(Z(G)) + \sum_{a \notin N(G)} [G.N(a)]$.

Proof. : We have from the class equation formula $o(G) = \sum_{a} [G:N(a)]$ ....(1) where the summation on the right hand side runs over one element from each conjugate class. Since for each element $a \in G$ we have either $a \in Z(G)$ or, $a \notin Z(G)$, the class equation (1) may be written as

$$o(G) = \sum_{a \in N(G)} [G:N(a)] + \sum_{a \notin N(G)} [G:N(a)]$$

By theorem 3.14, $a \in N(G) \Rightarrow N(a) = G$, and so

$$[G:N(a)] = \frac{o(G)}{o(N(a))} = \frac{o(G)}{o(G)} = 1$$

$$\therefore \sum_{a \in Z(G)} [G:N(a)] = \sum_{a \in N(G)} 1 = \text{No. of elements in Z(G)} = \text{o(Z(G))}$$

$$\text{Hence } o(G) = o(Z(G)) + \sum_{a \notin N(G)} [G:N(a)]$$

where the summation runs through a set of representative of the conjugacy classes.

Theorem 3.16 : If $o(G) = p^n$, where p is a prime number, then centre $Z(G) \neq \{e\}$.

Proof. : If $a \in G$, then we have proved that $N(a)$ is a subgroup of G. By Lagrange's theorem on finite group G, we know that $o(N(a))$ divides the order of the group G. But since $o(G) = p^n$, where p is a prime number, $o(n(a))$ must be of the form p^k, $o < k \leq n$.

Now from the class equation formula

$$o(G) = \sum_{a} [G:N(a)]$$

the summation extends over one element 'a' from each conjugacy class.

Let us assume that $o(Z(G)) = m$.

$$\begin{aligned} \text{Now, } a \in Z(G) &\Leftrightarrow N(a) = G \\ &\Leftrightarrow o(N(a)) = o(G) = p^n \\ &\Leftrightarrow p^k = p^n \Leftrightarrow k = n \end{aligned}$$

and $a \notin Z(G) \Leftrightarrow k < n$.

So that from the class equation

$$o(G) = o(Z(G)) + \sum_{a \notin Z(G)} [G:N(a)]$$

we have, $p^n = m + \sum_{k<n} \frac{p^n}{p^k}$

or, $m = p^n - \sum_{k<n} p^{n-k}$.

Since p is a divisor of the right hand side, p must divide m. Since $e \epsilon Z(G)$, $Z(G)$ is non-empty. Now, m is a positive integer divisible by the prime number p and consequently $m > 1$ or $o(Z(G)) > 1$. Hence $Z(G)$ must contain an element besides e.

Example 3.8 : If p is a prime number, any group G of order $2p$ has a normal subgroup of order p.

Solution : Since p is a prime number and $p \mid o(G)$, by Cauchy's theorem G has an element 'a' of order p. Then

$$N = \{a,\ a^2,\ \dots\ a^p = e\}$$

is a subgroup of order p.

Now, $\quad [G : N] = \dfrac{o(G)}{o(N)} = 2$

Hence N is a subgroup of index 2. So N is a normal subgroup of G.

Example 3.9 : Let G be a finite group, and T be an automorphism of G with the property that $T(x) = x$ if and only if $x = e$. Prove that every $g \epsilon G$ can be represented as $g = x^{-1}T(x)$ for some $x \epsilon G$.

Solution : We have $T(x) = x$ if and only if $x = e$. Consider the mapping $y : G \to G$ such that $y(x) = x^{-1}T(x)$, $x \epsilon G$.

Then for x, $y \epsilon G$, $\psi(x) = \psi(y)$

$$\Rightarrow \quad x^{-1}T(x) = y^{-1}T(y)$$

$$\Rightarrow \quad T(x) = xy^{-1}T(y)$$

$$\Rightarrow \quad T(x)[T(y)]^{-1} = xy^{-1}$$

$$\Rightarrow \quad T(x)T(y^{-1}) = xy^{-1}, \text{ (since } T \text{ is an automorphism)}$$

$$\Rightarrow \quad T(xy^{-1}) = xy^{-1}$$

$$\Rightarrow \quad xy^{-1} = e\ a\ x = y$$

$\therefore$ ψ is one-one.

Since G is finite, ψ is onto

Hence any $g\varepsilon G$ can be expressed as $g = x^{-1}T(x)$ for some $x\varepsilon G$.

Example 3.10 : Let G be a finite group, T an automorphism of G with the property that $T(x) = x$ if and only if $x = e$. Suppose further that $T^2 = I$ = identify mapping. Prove that G must be abelian.

Solution : From Example 3.7, we find that any $a\varepsilon G$ can be expressed as $a = x^{-1}T(x)$ for some $x\varepsilon G$.

$$\therefore\ T(a) = T(x^{-1})T^2(x)$$

$$= T(x^{-1})x, \quad \therefore\ T^2(x) = x$$

$$\therefore\ aT(a) = x^{-1}T(x)T(x^{-1})x$$

$$= x^{-1}T(xx^{-1})x$$

$$= x^{-1}T(e)x = x^{-1}ex = e$$

$$\therefore\ T(a) = a^{-1},\ a\varepsilon G$$

Now for $b\varepsilon G$, $(ab)^{-1} = T(ab)$

or, $T[(ab)^{-1}] = T^2(ab) = ab$

or, $T(b^{-1}a^{-1}) = ab$

or, $T(T(b)\ T(a)) = ab$

or, $T^2(b)\ T^2(a) = ab$

or, $ba = ab,\ a,\ b\varepsilon G$

This shows that the group G is abelian.

EXERCISES

1. If A and B are groups, prove that $A \times B$ is isomorphic to $B \times A$.

2. Let G be a group. Show that

$D = \{(g, g)\ \varepsilon G \times G \mid g\varepsilon G\}$ is a group isomorphic to G.

3. Prove that every group of infinite order is isomorphic to the additive group of integers.

[**Hint** : For infinite cyclic group G and $a \in G$, consider the mapping $n \rightarrow a^n$, n is an integer.]

4. If ϕ is an isomorphism of a group G onto to a group G', prove that $\{\varphi(a)\}^n = e'$ if and only if $a^n = e$.

5. Show that a cyclic group of order 8 is homomorphic to a cyclic group of order 4.

6. Let G be a group, and $f : G \rightarrow G$ defined by $f(x) = x^n$ be an automorphism. Show that for each $a \varepsilon G$, $a^{n-1} \varepsilon Z(G)$ (i.e., centre of G).

7. Let $G = < a >$ and $G' = < b >$ be any two cyclic groups of same order. Define $\varphi : G \rightarrow G'$ by $\varphi(a^r) = b^r$ for all integers r. Show that φ is well-defined and it is an isomorphism.

8. In the following, verify it the mapping defined are homomorphisms and in those cases in which they are homomorphisms, determine the kernel.

(a) G is the group of non zero real numbers under multiplication,

$\phi : G \rightarrow G$ such that $\phi(x) = x^2$, $x \in G$

(b) G is the same as in part (a),

$\phi : G \rightarrow G$ such that $\phi(x) = 2^x$, $x \in G$

(c) G is the group of real numbers under addition,

$\phi : G \rightarrow G$ such that $\phi(x) = x + 1$, $x \in G$

(d) G is any abelian group.

$\phi : G \rightarrow G$ such that $\phi(x) = x^5$, $x \in G$.

9. If $\phi : G \rightarrow H$ is a homomorphism and $G = <a>$ is a cyclic group, prove that ImG under ϕ is cyclic also.

Unit : 4 □ Characteristic of a ring, Ideals of a ring and Homomorphisms of rings

4.1. Characteristic of a ring : Let $< R, +, \cdot >$ be a ring. If each element of the additive group $<R, +>$ is of finite order and m (>0) is the maximum for the orders in $<R, +>$, then we say that the ring R has the characteristic m. For example, the ring $<R, +_6, \times_6>$ where $R = \{0, 1, 2, 3, 4, 5\}$ has characteristic 6.

On the other hand if then exists no maximum for the order of the elements of the additive group of a ring, then we say that the ring has the characteristic zero (or infinity). Thus, the usual rings Z, Q, R, C all have the characteristic zero.

Theorem 4.1 : The characteristic of a ring with unity is zero or m (> 0) according as the unity element 1 has the order infinity or it has finite order m.

Proof : Let the unity element 1 of the ring R have the finite order m and let x be any element of R then

$$\begin{aligned} mx = m(1 \cdot x) &= 1 \cdot x + 1 \cdot x + \ldots + 1 \cdot x, \text{ (m terms)} \\ &= (1 + 1 + \ldots + 1)x, \text{ (distributive rule)} \\ &= (m\ 1) \cdot x = 0 \cdot x = 0 \end{aligned}$$

This implies that the order of an arbitray element $x \in R$ is less or equal to m. So, by definition, characteristic of the ring is m.

Theorem 4.2 : The characteristic of an integral domain D is zero or m (> 0) according as the order of any non zero element in D is either infinity or m.

Proof : Let a be any nonzero element of D having finite order m, so that $ma = 0$.

Let d be any other nonzero element in D, then

$$\begin{aligned} 0 &= (ma) \cdot d = (a + a + \ldots + a) \cdot d \\ &= a \cdot d + a \cdot d + \ldots + a \cdot d \\ &= a \cdot (md) \end{aligned}$$

Since an integral domain is without zero divisor and $a \neq 0$, if follows that $md = 0$

or, order of $d \leq m =$ order of a

Let order of $d = p$, so that $pd = 0$

Then $0 = (pd)a = d(pa)$

Since $d \neq 0$, $pa = 0$, and so

$$\text{order of } a \leq p = \text{order of } d$$

By (1) and (2), order of d = order of a.

It follows that every nonzero element of D is of the same finite order m, so that the characteristic of an integral domain is m.

Theorem 4.3 : The characteristic of an integral domain is either zero or a prime number.

Proof : Let p be the characteristic of the integral domain D. We shall prove that p is a prime number. Suppose that p is not a prime number. Then we can write $p = p_1p_2$ where $p_1 \neq 1$, $p_2 \neq 1$ and $p_1, p_2 < p$. Let a be a nonzero element in D. Then $a^2 = a \cdot a \neq 0$. Since both a and a^2 are nonzero, we have

$$\text{order of } a = \text{order of } a^2 = p$$

$$\therefore \quad 0 = pa = pa^2 = (p_1p_2)a^2 = (p_1a)\ (p_2a)$$

Since D is integral domain, it follows that either $p_1a = 0$ or $p_2a = 0$. This contradicts the fact that p is the least positive integer such that $pa = 0$.

Hence p is a prime number.

4.2 Ideals of ring : Normal subgroups have a special role in group theory in the sense that normal subgroups permit us to construct factor groups or quotient groups. In the case of rings we have similar concepts, namely ideals and quotient rings.

Definition 4.1 : A nonempty subset S of a ring $R = < R, +, \cdot >$ is called an ideal of R if,

(i) $a, b \in S$ implies $a - b \in S$

(ii) $a \in S$ and $r \in R$ imply $ar \in S$ and $ra \in S$.

Definition 4.2 : A nonempty subset S of a ring $R = < R, +, \cdot >$ is called a right ideal or R if,

(i) $a, b \in S$ implies $a - b \in S$

(ii) $a \in S$ and $r \in R$ imply $ar \in S$

Similar is the definition for a left ideal.

We recall that a nonempty subset S of a ring R is a subring if and only if for all $a, b \in S$ we have $a - b \in S$ and $ab \in S$.

From the definition of an ideal it is clear that an ideal is a subring of the ring R. Every ideal is both right and left and so an ideal is some times called a *two-sided ideal.*

Consider two special subsets of R, viz, the subset $\{0\}$ and R itself. It is trivially true that these subsets are ideals of R and so are called trivial ideals. Ideals other than $\{0\}$ and R are called *proper ideals.*

Example 4.1 : Let R be a ring and $a \in R$. Consider the subset $aR = \{ax \mid x \in R\}$. We can easily show that aR is a right ideal of R. For, let ax and ay be two elements of aR; $x, y \in R$.

Then $ax - ay = a(x - y)$

Since $(x - y) \in R$, $a(x - y) \in aR$

So $\quad ax - ay \in aR$

Again, $ax \in aR$, and $y \in R$ will imply

$$(ax)y = a(xy)$$

Since $xy \in R$, $a(xy) \in aR$

Hence $(ax)y \in aR$

This shows that aR is a right ideal of R.

We can show that the set $Ra = \{xa \mid x \in R\}$ is a left ideal of R.

If R is commutative, then we say that aR is an ideal of R.

We note here that the element a need not belong to the ideal aR. The sufficient condition for a to be in aR is that R must be a ring with unity, i.e., $1 \in R$ and in this case $a \cdot 1 = a \in aR$.

Example 4.2 : Let R be the ring of all functions from the closed interval $[0, 1]$ to the field of real numbers i.e., the elements of R are real-valued functions $f(x)$ whose domain is the interval $[0, 1]$.

Clearly, R with usual addition and multiplication of functions form a ring. Consider the subset S of R defined as follows :

For $x \in [0, 1]$, $S = \{f \in R \mid f(x) = 0\}$. Then clearly, S is nonempty and for $f, g \in S$,

$$f(x) = 0,\ g(x) = 0$$
$$\Rightarrow f(x) - g(x) = 0,\ x \in [0, 1],$$

Also, $f(x) \in S$ and $h(x) \in R$, $x \in [0, 1]$, will imply, $f(x)\, h(x) = 0h(x) = 0$

$$\therefore \quad f(x)h(x) \in S$$

Hence, by definition, S is an idea! *of* R.

Example 4.3 : Let Z[x] denote the ring of all polynomials with integer coefficients and let S be the subset of $Z[x]$ with even constant coefficients.

Then S is an ideal of $Z[x]$.

Theorem 4.4 : If R is a commutative ring, the set $(a) = \{xa \mid x \in R\}$ of all multiples $ax = xa$ of any fixed $a \in R$ (by a variable element $x \in R$) is an ideal of R.

Proof : Let $x, y \in R$, then $xa, ya \in (a)$

Now, $xa - ya = (x - y)a \in (a)$ and for any $r \in R$, $r\,(xa) = (rx)a \in (a)$

$$(xa)r = (ax)r = a(xr) \in (a)$$

Hence (a) is an ideal of R.

Ideal (a) which consists of all the multiples of some fixed element $a \in R$ is called a *principal ideal of R*.

Theorem 4.5 : A commutative ring with unity is without proper ideals if and only if it is a field.

Proof : Assume that R is a commutative ring with unity but R is not a field. We shall show that R has a proper ideal. Since R is not a field, R will contain some non-invertible $a\ (\neq 0)$. Clearly then multiplying this element a by the elements of R cannot generate the unity element *of* R. Hence the ideal (a) is a proper ideal *of* R.

Now assume that R is a field and H be any ideal *of* R such that $H \neq \{0\}$. Then H must contain some nonzero $h \in R$. Since R is a field, multiplicative inverse h^{-1} of h must exist in R. For any $x \in R$, $xh^{-1} \in R$, and from the definition of an ideal $(xh^{-1})h \in H \Rightarrow x \in H$.

Thus, $x \in R \Rightarrow x \in H$ and so $R \subseteq H$.

But $H \subseteq R$. Hence $H = R$. This means that H is not a proper ideal of R if R is a field.

It follows that $\{0\}$ and $R = (1)$ are the only ideals of the field R.

Note : This result fails to be true in general for noncommutative rings.

Definition 4.3 : Let G be a group and S be a subgroup of G; for a, $b \in G$ we say a is congruent to b mod S, written as $a \equiv b \pmod{S}$ if $ab^{-1} \in S$.

It can be shown that the relation $a \equiv b \pmod{S}$ is an equivalence relation.

Theorem 4.6 : Let H be an additive subgroup of the ring R. Then the partition of R into cosets of H has the substitution property, i.e.

$a \equiv b \pmod{H} \Rightarrow ar \equiv br \pmod{H}$ for all $r \in R$, if and only if H is a right ideal.

Proof : If H is a right ideal, then $a \in H$, $b \in H \Rightarrow a - b \in H$ and $ar \in H$, $br \in H$ for all $r \in R$. Hence $ar - br \in H$, which means that $ar \equiv br \pmod{H}$.

Conversely, if $a - b \in H \Rightarrow ar - br = (\mathrm{a} - b)r \in H$ for all $r \in R$, then by definition H is a right ideal of R.

4.3 Homomorphisms of rings :

Definition 4.4 : Let f be a mapping from a ring R into a ring S such that for a, $b \in R$

(i) $f(a + b) = f(a) + f(b)$,

(ii) $f(ab) = f(a) f(b)$,

Then f is called a homomorphism of R into S.

We recall that if f is one-one, then f is called a monomorphism of R into S. In this case f is also called an embedding of the ring R into the ring S. If f is both one-one and onto then we say that the rings R and S are isomorphic and we write $R \cong S$.

Theorem 4.7 : Let $f : R \to S$ be a homomorphism of a ring R into a ring S, then we have the following :

(i) If 0 is the zero element of R, then $f(0)$ is the zero element of S.

(ii) If $a \in R$, then $f(-a) = -f(a)$

(iii) The set $\{f(a) \mid a \in R\}$ is a subring of S, called the homomorphic image of R by the mapping f and is denoted by $I_m f$ or $f(R)$.

(iv) The set $\{a \in R \mid f(a) = 0\}$ is an idea! in R, called the kernel of f and is denoted by K_f or $Kerf$.

(v) If $1 \in R$, then $f(1)$ is the unity element of the subring $f(R)$.

(vi) If R is commutative, then $f(R)$ is commutative.

Proof : (i) We have, since f is a homomorphism of rings, for $a \in R$, $f(a) = f(a + 0) = f(a) + f(0)$. This shows that $f(0)$ is the zero element of S. Let us denote the zero eiernent $f(0)$ of S by 0_s.

(ii) We have $0_s = f(0) = f(a + (-a)) = f(a) + f(-a)$

This shows that $f(-a) = -f(a)$

(iii) Let $f(a) f(b) \in f(R)$, then

$f(a) - f(b) = f(a) + f(-b) = f(a - b)$

Since $a - b \in R$, $f(a - b) \in f(R)$ and so $f(a) - f(b) \in f(R)$.

Also, $f(a) f(b) = f(ab) \in f(R)$

Hence $f(R)$ is a subring of S.

(iv) Let $a, b \in K_f = Ker\, f$ and let $r \in R$, then

$f(a) = f(b) = 0_s \Rightarrow f(a) - f(b) = 0_s$

$\Rightarrow f(a - b) = 0_s$

$\Rightarrow a - b \in K_f$

and $f(ar) = f(a) f(r) = 0_s f(r) = 0_s$ and hence $ar \in K_f$

$\therefore$ K_f is an ideal of R.

(v) Let $1 \in R$ and $a \in R$, then $1a = a$. Since f is a homomorphism,

$f(a) = f(1) f(a)$

Similarly, $f(a) = f(a) f(1)$

This shows that $f(1)$ is the unity element in $f(R)$.

(vi) If R is commutative, then $a, b \in R \Rightarrow ab = ba$ and since f is a homomorphism,

$f(ab) = f(ba)$

$\Rightarrow$ $f(a) f(b) = f(b) f(a)$

Hence $f(R)$ is commutative.

Theorem 4.8 : Let $f : R \rightarrow S$ be a homomorphism of a ring R into a ring S. Then $K_f = \{0\}$ if and only if f is one-one.

Proof : Suppose that f is one-one. Then, for $a, b \in R, f(a) = f(b) \Rightarrow a = b$. We also know that $f(0) = O_s$, so $O \in K_f$. Let $a \in K_f$, then by definition of a kernel $f(a) = O_s$. So $F(a) = O_s = f(0)$, But since f is one-one, we have $a = 0$.

$\therefore \quad K_f = \{0\}$.

Conversely, let $K_f = \{0\}$. We shall show that f is one-one. Let $a, b \in R$. If $f(a) = f(b)$, then $f(a - b)$, then $f(a - b) = O_s \Rightarrow$ a $- b \in K_f$. But K_f contains only the element O. So $a - b = 0$ or $a = b$. Hence f is one-one.

Theorem 4.9 : (Fundamental Theorem of Homomorphism) Every homornorphic image of a ring R is isomorphic to some quotient ring of R.

Proof : Let S be the homomorphic image of a ring R under a homomorphism $\theta : R \to S$ and let H be the kernel of θ. Then H is an ideal of R. We consider the mapping φ defined as

$$\varphi : R/H \to S$$

such that $\varphi(H + a) = \theta(a), a \in R$. We first show that φ is well-defined.

Let $h \in H$, then $h + a \in H + a$,

so that $\theta(h + a) = \theta(h) + \theta(a) = \theta(a), \quad \because \theta(h) = O_s$.

This shows that all the elements of the additive coset $H + a$ map onto the element $\theta(a)$. Hence φ is well defined.

Next we shall show that rpis one-one. Suppose that

$$\varphi(H + a) = \varphi(H + b), b \in R,$$

then, $\theta(a) = \theta(b)$.

But θ is a homomorphism, so $\theta(a - b) = O_s \Rightarrow a - b \in H$

or, $\quad H + a = H + b$

$\therefore \quad \varphi$ is one-one.

Since S is the homomorphic image of R under θ, for every $x \in S, \exists a \in R$ such that $\theta(a) = x$. This implies that $\varphi(H + a) = x$. So φ is onto.

Finally, we have to show that φ is a homomorphism we have,

$$\begin{aligned}\varphi((H + a) + (H + b)) = \varphi(H + (a + b)) &= \theta(a + b)\\ &= \theta(a) + \theta(b)\end{aligned}$$

$$= \varphi(H + a) + \varphi(H + b)$$

Also $\phi((H + a)\ (H + b)) \ = \phi(H + ab)$

$$= \theta(ab) = \theta(a)\ \theta(b)$$

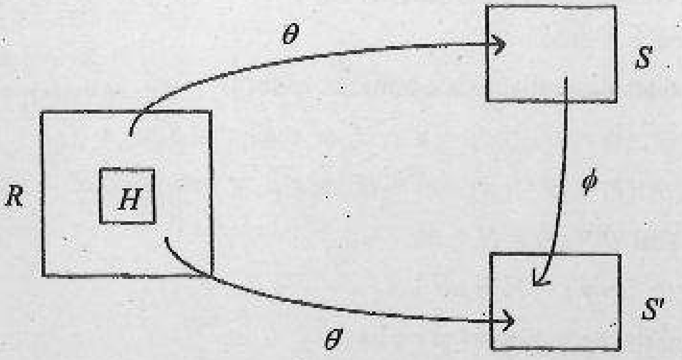
$$= \phi(H + a)\ \phi(H + b)$$

Thus ϕ is a homomorphism.

Since ϕ is one-one and onto homomorphism, ϕ is an isomorhism from R/H to S.

Theorem 4.10 : Let $\theta : R \rightarrow S$ and $\theta' : R \rightarrow S'$ be two epimorphisms of rings with the same domain R and kernel H. Then their images S and S' are isomorphic.

Proof :



Let us define a mapping $f : R/H \rightarrow S$ such that for $a \in R$, $f(a + H) = \theta(a)$. The mapping is well-defined, since $a + H = b + H \Rightarrow a - b \in H \Rightarrow \theta(a - b) = O_s \Rightarrow \theta(a) - \theta(b) = O_s \Rightarrow \theta(a) = \theta(b)$.

Now, $f((a + H) + (b + H)) = f(a + b + H) = \theta(a + b)$

$$= \theta(a) + \theta(b) = f(a + H) + f(b + H)$$

and $\quad f((a + H)\ (b + H)) = f(ab + H) = \theta(ab) = \theta(a)\ \theta(b)$

$$= f(a + H)\ f(b + H)$$

Hence f is a homomoiphism.

Clearly, f is onto.

Again, $f(a + H) = f(b + H) \Rightarrow \theta(a) = \theta(b) \Rightarrow \theta(a) - \theta(b) = O_s \Rightarrow \theta(a - b) = O_s$

$$\Rightarrow a - b \in H \Rightarrow \mathrm{a} + H = \mathrm{b} + H$$

$\therefore$ f is one-one.

Hence f is an isomorphism from $R/H \to S$.

Thus, $R/H \cong S$.

Similar argument shows thai $R/H \cong S'$.

Let us now define the mapping $\varphi : S \to S''$ such that

if $\theta(a) = x \in S$, $\theta'(a) = x' \in S'$, $a \in R$, then $\varphi(x) = x'$

Now $\varphi(x) = \varphi(y) \Rightarrow x' = y' \Rightarrow \theta'(a) = \theta'(b) \Rightarrow \theta'(a - b) = 0' \Rightarrow a - b \in H \Rightarrow a + H = b + H \Rightarrow f(a + H) = f(b + H) \Rightarrow \theta(a) = \theta(b) \Rightarrow x = y$

$\therefore$ φ is one-one.

Since both θ and θ' are epimorphisms and since $\theta^{-1}(x) = \theta'^{-1}(x') = a + H$, it follows that the mapping φ is onto.

Lastly we shal! show that φ is a homoinorphism

We have, $\theta^{-1}(x) = \theta'^{-1}(x') = a + H$, $a \in R$, $x \in S$, $x' \in S'$

$$\theta^{-1}(y) = \theta'^{-1}(y') = b + H,\ b \in R,\ y \in S,\ y' \in S'$$

It follows that $\theta^{-1}(xy) = H + ab$

and $\theta'^{-1}(x'y') = H + ab$

Hence from the definition of φ we have

$$\varphi(xy) = x'y' = \varphi(x)\varphi(y)$$

Also, $\theta^{-1}(x + y) = \theta'^{-1}(x' + y') = (a + b) + H$

$$\therefore \quad \varphi(x + y) = x' + y' = \varphi(x) + \varphi(y)$$

So φ is an isomorphism from S to S' and

$$S \simeq S'.$$

Theorem 4.11 : In any integral domain D of prime characterstic p, the assignment $x \to x^p$ is a monomorphism.

Proof : If a, $b \in D$, we have, by binomial theorem

$$(a \pm b)^p = ap \pm \binom{p}{1} a^{p-1} b + \binom{p}{2} a^{p-2} b^2 + \ldots + \binom{p}{p} (\pm b)^p \qquad \ldots\ldots\ldots (1)$$

Let us consider the term $\binom{p}{K} a^{p-K} b^K$ in (1).

Clearly $a^{p-K}b^K \in D$ and $\binom{p}{K} = \frac{p!}{K!(p-K)!}$

Since p ! is divisible by p and neither K ! nor $(p-K)$! is divisible by p for $0 < K < p$, it follows that $\binom{p}{K}$ is a multiple of p and since p is the characteristic of the integral domain, we have

$$\binom{p}{K}a^{p-K}b^K = 0, \quad 0 < K < p.$$

Hence from (1),

$$(a \pm b)^p = a^p \pm b^p \qquad \text{.......... (2)}$$

Again, since D is a commutative monoid,

$$(ab)^p = a^p b^p \qquad \text{.......... (3)}$$

Thus, if we define a mapping $\mu : D \to D$ by $\mu(x) = x^p$, $x \in D$, then (2) and (3) show that

$$\mu(a+b) = (a+b)^p = a^p + b^p = \mu(a) + \mu(b)$$

and $\mu(ab) = (ab)^p = a^p b^p = \mu(a)\,\mu(b)$

Hence μ is a homomoiphism from D to D.

Finally, $\mu(a) = \mu(b) \Rightarrow a^p = b^p \Rightarrow a^p - b^p = 0 \Rightarrow (a-b)^p = 0$, by (2)

Since D is on integral domain $(a-b)^p = 0 \Rightarrow a = b$

Hence $m(a) = m(b) \Rightarrow a = b$

$\therefore$ μ is a Mionomorphism from D to D.

4.4 Maximal ideal and prime ideal

Def. 4.5 : An ideal M of a ring R is said to be a maximal idea! of R if

(1) $M \neq R$

(2) there exists no other ideal H of R such that $M \subset H \subset R$

Thus from the definition it follows that if we have an ideal H of R such that $M \subseteq H \subseteq R$, then either $H = M$ or $H = R$.

Example 4.4 : In the ring R of integers with usual addition and multiplication the ideal (2) [recall that (2) contains the set of elements obtained by multiplying 2 by the elements of R], is a maximal ideal. This may be established easily. We note that the element $9 \in (2)$ but $9 \notin (2)$. Hence $(2) \subset R$. Now suppose thai H is any other ideal of R such that $(2) \subset H$, this means that there is an element $x \in H$ which is not divisible by 2. Hence $x = 2n + 1$ for some integer n.

Now, $1 = x - 2n$

Since $x \in H$ and $2n \in (2)$, so $2n \in H$ and since H is an ideal of R, $x - 2n - 1 \in H$. Hence from the property of an ideal, we have for any $y \in R$, $1.y \in H$ i.e. $y \in H$. So $R \subseteq H$. But $H \subseteq R$. Hence $H = R$. Hence (2) is a maximal ideal in the ring of integers. It can be easily proved that for any prime number p, the ideal (p) is maximal in the ring of integers.

Theorem 4.12 : In a nonzero commutative ring with unity, an ideal M is maximal if and only if R/M is a field.

Proof : Suppose that R/M is a field and B is any ideal of R such that $M \subset B$. Let $b \in B$ but $b \notin M$. Then $b + M$ is a nonzero element of R/M and since by assumption R/M is a field, there exists an element $c + M$ of R/M such that

$$(b + M)(c + M) = 1 + M \ \text{(= the multiplicative identity of } R/M\text{)}$$

Since $b \in B$ and B is an ideal of R, we have

$$bc \in B$$

Again, $1 + M = (b + M)(c + M) = bc + M$

So, $\quad 1 - bc \in M \subset B$

$\therefore \quad (1 - bc) + bc \in B$

$\therefore \quad 1 \in B$

So for every $x \in R$, $1 \cdot x \in B$ i.e. $x \in B$

Hence $B = R$

$\therefore$ The ideal M is a maximal ideal in R.

Conversely, suppose that M is a maximal ideal in R. We shall show that R/M is a field.

To prove this we have to show that every nonzero element $b + M \in R/M$ has a multiplicative inverse.

Consider the set $B = \{br + a \mid r \in R, a \in M\}$. Clearly B is an ideal in R, and $M \subset B$. But as M is a maximal ideal in R, we must have $B = R$. So $1 \in B$. From the structure of elements of B we must find some $c \in R$ and $a' \in M$ such that

$$1 = bc + a'$$

and hence $1 + M = bc + a' + M = bc + M = (b + M)(c + M)$

This shows that the element $c + M$ is the multiphicative inverse of $b + M$ in R/M. So R/M is a field.

Example 4.5 : If R is the ring of 2×2 matrices of the form $\begin{bmatrix} a & b \\ 0 & 0 \end{bmatrix}$ where a and b are elements of a field F. Show that the set $M = \left\{ \begin{bmatrix} 0 & b \\ 0 & 0 \end{bmatrix} \middle| b \in F \right\}$ is a maximal ideal in R.

Solution : We can easily prove that M is an ideal in R. Let us consider the subset S of R of the form

$$S = \left\{ \begin{bmatrix} a & 0 \\ 0 & 0 \end{bmatrix} \middle| a \in F \right\}$$

Then S is a subring of R and if we define a mapping $f : S \to F$ by

$$\phi\left(\begin{bmatrix} a & 0 \\ 0 & 0 \end{bmatrix}\right) = a,$$

then clearly ϕ is one-one and onto.

Now, $$\phi\left(\begin{bmatrix} a & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} b & 0 \\ 0 & 0 \end{bmatrix}\right) = \phi\left(\begin{bmatrix} a+b & 0 \\ 0 & 0 \end{bmatrix}\right) = a + b$$

$$= \phi\left(\begin{bmatrix} a & 0 \\ 0 & 0 \end{bmatrix}\right) + \phi\left(\begin{bmatrix} b & 0 \\ 0 & 0 \end{bmatrix}\right)$$

and $$\phi\left(\begin{bmatrix} a & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} b & 0 \\ 0 & 0 \end{bmatrix}\right) = \phi\left(\begin{bmatrix} ab & 0 \\ 0 & 0 \end{bmatrix}\right) = ab$$

$$= \phi\left(\begin{bmatrix} a & 0 \\ 0 & 0 \end{bmatrix}\right)\phi\left(\begin{bmatrix} b & 0 \\ 0 & 0 \end{bmatrix}\right)$$

Hence ϕ is an isomorphism.

Now, we consider the mapping $f : R \to S$, where

$$f\left(\begin{bmatrix} a & b \\ 0 & 0 \end{bmatrix}\right) = \begin{bmatrix} a & 0 \\ 0 & 0 \end{bmatrix}$$

we can show that the mapping ϕ is onto nomomorphisrr. whose kernel is M. Thus, by the fundamental theorem of homomorphism $R/M = S$. But S is a field. Hence R/M is a field and so M is maximal idea! in R.

Definition 4.5 Prime ideal : Let R be a commutative ring. An ideal P of R is said So be a prime ideal of R if for $a, b \in R$, $ab \in P$ implies that $a \in P$ or $b \in P$.

Example 4.6 : In any integral domain R, (0) is a prime ideal. For if $ab \in (0)$, $a, b \in R$ then celearly either $a \in (0)$ or $b \in (0)$.

Theorem 4.13 : An ideal F of a commutative ring R is prime if and only if R/P is an integral domain.

Proof : Let R/P be an integral domain.

Then for all $a, b \in R$, $ab \in P \Rightarrow ab + P = P$

$$\Rightarrow (a + P)\ (b + P) = P$$

$$\Rightarrow \text{either } a + P = P \text{ or } b + P = P$$

$$\Rightarrow \text{either } u \in P \text{ or } b \in P.$$

Hence P is a prime ideal.

Conversely, let P be a prime ideal in R. Then

$$(a + P)\ (b + P) = P \Rightarrow ab + P = P$$

$$\Rightarrow ab \in P$$

$$\Rightarrow \text{either } a \in P \text{ or } b \in P$$

$$\Rightarrow \text{either } a + P = P \text{ or } b + P = P$$

Thus R/P is without zero divisor. Also R/P is a commutative ring as R is commutative. Hence R/P is an integral domain.

Theorem 4.14 : If R is a commutative ring with unity, then each maximal ideal is prime, but the converse, in general, is not true.

Proof : Let M be a maxima! ideal in R. Then by Theorem 4.12, R/M is a field. So R/M is necessarily an integral domain. Hence by the Theorem 4.13, M is a prime ideal.

The converse of this therorem is not true, in general. For, the ideal (0) in the ring of integers is prime but not maximal.

EXERCISES

1. Prove that the homomorphism ϕ of a ring R onto a ring R' is an isomorphism of R onto R' if and only if $K_\phi = \{0\}$, where K_ϕ denotes the kernel of ϕ.
2. If R is a ring with unity element 1 and φ is a homomorphism of R onto an integral domain D such that $K_\varphi \neq R$, prove that $\varphi(1)$ is the unity element of D.
3. Let R be the ring of integers and let U be the ideal consisting of all multiples of 17. Prove that if V is an ideal of R and $R \supseteq V \supseteq U$, then either $V = R$ or $V = U$.
4. If U is an ideal of a ring R and $1 \in U$, prove that $U = R$.
5. If U, V are ideals of R and $U + V = \{u + v \mid u \in U, v \in V\}$. Prove that $U + V$ is also an ideal of R.
6. Prove that any homomorphism of a field is either an isomorphism or takes each element onto o.
7. Let R be a ring and L be a left ideal of R. Let $\lambda(L) = \{x \in R \mid xa = 0, \forall a \in L\}$. Prove that $\lambda(L)$ is a two-sided ideal *of* R.
8. If R is a commutative ring and $a \in R$, show that $aR = \{ax : x \in R\}$ is a left and right ideal *of* R.

   What will be your conclusion if R is noncommutative?

Unit : 5 ❑ Edclidean domain, Polynomial rings

Division algorithm for the set of integers Z :

We state here a very fundamental theorem called the Division Algorithm for Z.

Theorem 5.1 : Let $a, b \varepsilon Z, b > o$. Then there exists a unique integers q and r such that $a = hq + r, o \leq r < b$,

where q is called the quotient and r the remainder of a modulo b.

(For proof see, for example, Bhattacharya et.al p 32)

5.1. Euclidean domain :

Definition 5.1 : A Euclidean domain is an integral domain D with a valuation $v : D^* \rightarrow N$ (D^* is the set of non zero elements of D and N is the set of non-negative integers), having the following properties :

For all $x, y \ \varepsilon D^*, v(xy) \geq v(x)$ ...... (1)

Given $a \varepsilon D$ and $b \varepsilon D^*$ there exists a $q \varepsilon D$ such that $a = bq + r$, where either $r = 0$ or, $v(r) < v(b)$ ...... (2)

Example 5.1 : The integral domain Z is a Euclidean domain with valuation v defined by $v(a) = | a |, a \varepsilon D$.

For, $v, y \varepsilon D \Rightarrow v(xy) = | xy | = | x | \ | y |$ and $| y | \geq |$ if $y \neq 0$, so $v(xy) \geq v(x)$ and the condition (1) is satisfied.

Condition (2) is just the division algorithm for integers.

Theorem 5.2 : In any Euclidean domain D, for $xy \varepsilon D^*$, we have $v(xy) = v(x)$ if y is invertible, whereas $v(xy) > v(x)$ if it is not.

Proof. : We have, by the definition of Euclidean domain,

$x, y \varepsilon D^*$ implies $v(xy) \geq v(x)$, ... (1)

Now, if y is invertible. Then

$v(x) = v(xyy^{-1}) \geq v(xy)$ ...... (2)

By (1) and (2), $v(xy) = v(x)$

Moreover, it xy divides x, then there exists some $z \varepsilon D$ such that $xyz = x = x .1$. But since D is an integral domain, by cancellation rule, $yz = 1$. Also, since $x \varepsilon D^*$, y must be invertible.

Now, suppose that xy does not divide x. Then from the second property of Euclidean domain, $x = q(xy) + r$, q, $r \varepsilon D$ ...... (3)

such that either $r = 0$ or, $v(r) < v(xy)$ ...... (4)

Now from (3) $r = x - q(xy) = x(1 - qy)$

$$\therefore \quad v(r) = v\{x(1 - qy)\} \geq v(x) \qquad \text{...... (5)}$$

From (4) and (5), $v(xy) > v(r) > v(x)$

$$\text{i.e., } v(xy) > v(x)$$

Thus, when y is invertible, $v(xy) = v(x)$ and when it is not then $v(xy) > v(x)$.

Def. 5.2 : Let R be a commutative ring with unity. An element $u \varepsilon R$ is said to be unit in R if there exists an element $b \varepsilon R$ such that $ub = 1$.

In other words, u is a unit in R if it has multiplicative inverse in R.

Note : There may be more than one unit in R but the unity element is unique in R.

Example 5.2 : In the integral domain Z the only units are $+ 1$ and $- 1$.

Definition 5.3 : Let R be a commutative ring with unity. We say that two elements a and b in R are associates and write $a \sim b$, of for some unit $u \varepsilon R$, $au = b$.

Theorem 5.3 : For a Euclidean domain with Euclidean valuation v, $v(l)$ is minimal among all $v(a)$, $a \varepsilon D^*$ and $u \varepsilon D$ is a unit if and only if $v(u) = v(l)$.

Proof. : Let $a \varepsilon D^*$. Then since $a.1 = a$, we have $v(a) = v(a.l) \geq v(l)$.

This proves the first part of the theorem.

Let it be a unit in D. Then u has its multiplicative inverse $u^{-1} \varepsilon D$ such that $uu^{-1} = 1$

$$\therefore \quad v(1) = v(u.u^{-1}) \geq v(u)$$

But from the first part, $v(u) \geq v(l)$

Combining these together, we have $v(u) = v(l)$

Conversely, suppose that $u \varepsilon D$ is such that $v(u) = u(1)$, we shall show that u is a unit in D.

Since u and 1 belong to D, it follows from the second property of Euclidean domain that there exist q and $r \varepsilon D$ such that

$$1 = qu + r \qquad \text{...... (1)}$$

where either $r = 0$ or, $v(r) < v(u)$.

But since $v(u) = v(1)$ is minimal over all $v(d)$ with $d\varepsilon D^*$, it follows that $v(r) < v(u)$ is impossible. Hence $r = 0$ and so from (1) we have $1 = qu$ and hence u is a unit in D.

Definition 5.4 : Any integral domain in which every idea! is principal, is said to be a principal ideal domain (*PID*).

Theorem 5.4 : In any Euclidean domain D, every ideal is principal.

Or

A Euclidean domain is a *PID*.

Proof. : Let H be any ideal of the given Euclidean domain D. If $H = \{o\}$, there is nothing to prove. So Suppose that H contains a non zero element b with minimum $v(b)$. We shall show that $H = (b)$, a principal ideal of D.

$$\text{Clearly } H \supseteq (b) \qquad \text{...... (1)}$$

It remains to prove that $H \subseteq (b)$. If $a\varepsilon H$, then from the property of Euclidean domain D, there exist q, $r\varepsilon D$ such that $a = bq + r$, where either $r = 0$ or, $v(r) < v(b)$. But $v(r) < v(b)$ is impossible, since $v(b)$ is minimal among the valuations of elements of H. Hence $r = 0$.

$$\text{So} \quad a = bq$$

This means that any element $a\varepsilon H$ is a multiple of $b\varepsilon H$. So $H \subseteq (b)$ ...... (2)

Combining (1) and (2), we have $H = (b)$ and the ideal H is principal.

Definition 5.5 : Let a, b be two elements of an integral domain D. An element $d\varepsilon D$ is said to be a greatest common divisor (*gcd*) of a and b if:

(i) $d \mid a$ and $d \mid b$, [$d \mid a$ is read as d divides a]

(ii) whenever $c\varepsilon D$ is such that $c \mid a$ and $c \mid b$, then $c \mid d$.

Theorem 5.5 : If D is a Euclidean domain with Euclidean valuation v and a, b are two non zero elements of D, then there exists a *ged* d of a and b.

Moreover $d = \lambda a + \mu b$ for some λ, $\mu\varepsilon D$.

Proof. : Consider the set $H = \{ra + sb \mid r, s\varepsilon D]$

Now, $(r_1a + s_1b) \pm (r_2a + s_2b) = (r_1 \pm r_2)a + (s_1 \pm s_2)b$, $r_1, r_2, s_1, s_2 \varepsilon D$

$$\text{and } t(ra + sb) = (tr)a + (ts)b,\ t\varepsilon D.$$

It follows that H is an ideal of D. Since every ideal in a Euclidean domain is principal, it follows that $H = (d)$, for some $d\varepsilon D$. Then since every element of H is a multiple of d, we have $d \mid (ra + sb)$ for all r and $s\varepsilon D$.

So if we take $s = 0$, $r = 1$, we find that $d \mid a$ and taking $r = 0$, $s = 1$ we have d / b.

Let c be any other divisor of a and b i.e., c / a and c / b. Then clearly c / ra and c / sb for all $r, s\varepsilon D$. Hence $c / (ra + sb)$ for all $r, s\varepsilon D$. Thus c divides all elements of H and as $d\varepsilon H$, it follows that c/d. Hence d is a gcd of a and b. And from the nature of construction of H it follows that there exist λ, $\mu\varepsilon D$ such that

$$d = \lambda a + \mu b.$$

Theorem 5.6 : (Euclidean Algorithm). Let D be a Euclidean domain with Euclidean valuation v and let a, b be non zero elements of D. Let $r_1\varepsilon D$ be such that

$$a = bq_1 + r_1,\ q_1\varepsilon D$$

where either $\quad r_1 = o \quad$ or, $v(r_1) < v(b)$.

Let r_2 be such that $b = r_1q_2 + r_2$; $q_2, r_2\varepsilon D$, where either $r_2 = 0$, or, $v(r_2) < v(r_1)$.

In general, let r_{i+1} be such that

$$r_{i-1} = q_{i+1}\, r_i + r_{i+1};\ q_{i+1},\ r_{i+1}\varepsilon D$$

where either $\quad r_{i+1} = 0$, or, $v(r_{i+1}) < v(r_i)$

Then the sequence r_1, r_2 .... must terminate with some $r_s = 0$ and then r_{s-1} is a gcd of a and b.

Proof. : Since $v(r_i) < v(r_{i-1})$ and $v(r_i)$ is a non-negative integer, it is clear that after some finite steps we must arrive at $r_s = 0$.

Now, if $r_1 = 0$, then $a = bq_i$, so that b is clearly a gcd of a and b. Suppose that $r_1 \neq 0$. Then if d / a and d / b, we have $d \mid (a - bq)$, $q\varepsilon D$, i.e., $d \mid r_1$.

However, if $c \mid r_1$ and $c \mid b$, then $c \mid (bq + r_1)$, i.e., c $\mid a$. Hence the set of common divisors of a and b is the same set as the set of common divisors of b and r_1.

By a similar argument, if $r_2 \neq 0$, the set of common divisors of b and r_1 is the same set as the set of common divisors of r_1 and r_2. Continuing this process, we see finally that the set of common divisors of a and b is the same set as the set of common divisors of r_{s-2} and r_{s-1}.

Thus when $r_s = 0$, a gcd of r_{s-2} and r_{s-1} is also a gcd of a and b.

But when $r_s = 0$, we have $r_{s-2} = q_s r_{s-1}$ and this shows that r_{s-1} is a gcd of a and b.

Definition 5.6 : In a Euclidean domain D a non unit a is said to be **Prime element** of D if whenever $a = xy$, where $x, y \varepsilon D$, then one of x and y is a unit in D.

Thus, according to this definition, a prune element is an element in D which cannot be factored in a nontrivial way.

Theorem 5.7 : Let D be a Euclidean domain. Every non-zero element in D is either a unit in D or can be written as a product of finite number of prime elements of D.

Proof. : The proof is by induction on $v(a)$. If $v(a) = v(1)$ then a is a unit in D and so in this case the statement of the theorem is correct.

We assume that the theorem is true for all non zero elements $x \varepsilon D$ such that $v(x) < v(a)$. On the basis of it we shall prove the theorem for a.

If a is a prime element of D, there is nothing to prove. So suppose that $a = bc$, where neither b nor c is a unit in D.

Now, $\quad v(b) < v(bc) = v(a)$ and $v(c) < v(bc) = v(a)$

Thus, by our induction hypothesis b and c can be written as the product of finite number of prime elements of D;

$$b = p_1 p_2 \ldots p_n,\ c = p'_1 p'_2 \ldots p'_m,$$

where p'_s and p''_s are prime elements of D. Consequently, $a = bc = P_1 P_2 \ldots P_n p'_1 p'_2 \ldots p'_m$ and in this way a has been factored as a product of finite number of prime elements. This completes the proof.

Definition 5.7 : In any Euclidean domain D, a and b in D are said to be relatively prime if their gcd is a unit of D.

Definition 5.8 : If R is a commutative ring with unity, then two elements a and b in R are said to be associates if $b = an$ for some unit u in R.

Since any associate of a gcd is a gcd and since 1 is an associate of any unit, it follows that if a and b are relatively prime, we may assume $(a, b) - 1$, i.e., ged of a and b is 1.

Theorem 5.8 : Let D be a Euclidean domain. Suppose that for $a, b, c \varepsilon D$, $a \mid bc$ but $(a, b) = 1$. Then $a \mid c$

Proof. : We have already established that gcd of a and b can be obtained in the form $\lambda a + \mu b$, λ, $\mu \varepsilon D$. So $(a, b) = 1$ means that $\lambda a + \mu b = 1$.

Multiplying this equation by c we get

$$\lambda ac + \mu bc = c$$

Now, $a \mid \lambda ac$ always and $a \mid \mu bc$, since it is given that $a \mid bc$. Hence $a \mid (\lambda ac + \mu bc)$, or, $a \mid c$

Theorem 5.9 : If p is a prime element in the Euclidean domain D and $p \mid ab$, where $a, b \varepsilon D$, then p divides at least one of a and b.

Proof. : Suppose that p does not divide a. Then $(p, a) = 1$. Hence by Th. 5.8, $p \mid b$.

An immediate extension of Th. 5.9 is that if p is a prime element in the Euclidean domain and $p \mid a_1 a_2 \dots a_n$, then p divides at least one of $a_1, a_2, \dots a_n$.

Theorem 5.10 : (Unique Factorization Theorem). Let D be a Euclidean domain and $a (\neq 0)$ be a nonunit in D. Suppose that $a = p_1 p_2 \cdots p_n = p'_1 p'_2 \cdots p'_m$, where p_i and p'_j are prime elements of D. Then $n = m$ and each p_i $(1 \leq i \leq n)$ is an associate of some p'_j $(1 \leq j \leq m)$ and conversely.

Proof. : We have $a = p_1 p_2 \cdots p_n = p'_1 p'_2 \cdots p'_m$.

But $\quad p_1 \mid p_1 p_2 \cdots p_n$, hence $p_1 \mid p'_1 p'_2 \cdots p'_m$.

By Th. 5.9, p_1 must divide some p'_j; since both p_1 *and* p'_j are prime elements of D and $p_1 \mid p'_j$, they must be associates,

and $\quad p'_j = u_1 p_1$, where u_1 is a unit in D.

Thus, $\quad p_1 p_2 \cdots p_n = u_1 p_1 \; p'_1 p'_2 \cdots p'_{j-1} \, p'_{j+1} \cdots p'_m$

$$\Rightarrow p_2 \cdots p_n = u_1 p'_1 p'_2 \cdots p'_{j-1} \, p'_{j+1} \cdots p'_m.$$

Repeating this argument, after n steps we get 1 on the left hand side and a product of certain factors of p' on the right hand side. Since p', s are prime, we must have $m = n$. This proves the theorem.

5.2. Polynomial Rings :

Let R be a commutative ring. By the ring of polynomials in the indeterminate x, written as $R[x]$, we mean the set of ail symbols

$$a_0 + a_1x + \ldots + a_mx^m = \sum_{k=0}^{m} a_kx^k \qquad \text{......... (1)}$$

where m can be any non-negative integer and where the coefficients a_0, a_1, a_2 ... a_m are all in R.

Let us denote the polynomial in (1) by $a(x)$, so that $a(x) = a_0 + a_1x + \ldots + a_mx^m$.

The canonical form of the polynomial $a(x)$ is defined as follows :

(i) pick the largest k in (1) with $a_k \neq 0$, say $k = n$, and then

(ii) rewrite (i) as

$$a(x) = a_0 + a_1x + \ldots + a_nx^n;\ a_n \neq 0 \qquad \text{......... (2)}$$

In the exceptional case that when all $a_k = 0$, the canonical form is zero. The degree of $a(x)$ is the integer n in (2) and is symbolized by deg a; if $a(x)$ has the canonical form zero, it is said to have degree $- \infty$.

Definition 5.9 : If $a(x) = a_0 + a_1x + \ldots + a_mx^m$ and $b(x) = b_0 + b_1x + \ldots + b_nx^n$ are in $R[x]$, then $a(x) = b(x)$ if and only if for every integer $i \geq 0$, a/ $= a_i = b_i$.

Thus two polynomials will be said to be equal if and only if their corresponding coefficients are equal.

Definition 5.10 : If $a(x) = a_0 + a_1x + \ldots + a_mx^m$ and $b(x) = b_0 + b_1x + \ldots + b_nx^n$ are both in $R[x]$, then

$$a(x) + b(x) = c_0 + c_1x + \ldots + c_1x^1,$$

where for each i, $c_i = a_i + b_i$

Definition 5.11 : If $a(x) = a_0 + a_1x + \ldots + a_mx^m$ and $b(x) = b_0 + b_1x + \ldots + b_nx^n$, then

$$a(x)b(x) = c_0 + c_1x + \ldots + c_kx^k$$

where $\quad c_t = a_tb_0 + a_{t-1}b_1 + \ldots + a_0b_t$

With these definitions of addition and multiplication of elements in $R[x]$, we have the following theorem.

Theorem 5.11 : The set $R[x]$ of all polynomials in an indeterminate x with coefficients from a commutative ring R, is a commutative ring under polynomial addition and multiplication. Further, if R has unity 1, then 1 is also the unity *for* $R[x]$.

The proof of this theorem is simple and we leave it to the readers for checking that the conditions for being a ring are satisfied.

Example 5.3 : Let $R = \{0, 1\} = Z_2$,

and let $a(x) = 1 + ox + 1x^2$

$b(x) = 1 + 1x + 1x^2$

$a(x) + b(x) = (1 + ox + 1x^2) + (1 + 1x + 1x^2)$

$= 0 + 1x + 0x^2 = x$

and $a(x)\ b(x) = (1 + ox + 1x^2)\ (1 + 1x + 1x^2)$

$= 1 + 1x + 1x^3 + 1x^4$

Trivially, the degrees of the sum and product of two polynomials satisfy the inequalities

$$\deg\ (a + b) \leq \deg\ a + \deg\ b$$

$$\deg\ (ab) \leq \deg\ a + \deg\ b.$$

Lemma : If D is an integral domain, then in $D[x]$, we have deg (ab) = deg a + deg b.

Proof. : Let $a(x)$, $b(x)$ $\varepsilon D[x]$ and

$a(x) = a_0 + a_1x + ... + a_mx^m$

$b(x) = b_0 + b_1x + ... + b_nx^n$

with $a_m \neq 0$, $b_n \neq 0$. So that deg $a = m$, deg $b = n$.

By definition $a(x)\ b(x) = c_0 + c_1x + ... + c_kx^k$, where $c_t = a_tb_0 + a_{t-1}b_1 + ... + a_0b_t$

We have $c_{m+n} = a_mb_n$. Since $a_m \neq 0$, $b_n \neq 0$ and D is an integral domain, it follows that $c_{m+n} \neq 0$. We shall now show that $c_i = 0$ for $i > m + n$. Since c_i is the sum of the terms of the form a_jb_{i-j} and since $i = j + (i - j) > m + n$, then *either* $j > m$ or, $i - j > n$. But then one of a_j and b_{i-j} is zero and so that $a_jb_{i-j} = 0$. Since c_i is the sum of such terms, $c_i = 0$ for $i > m + n$.

Hence deg $(ab) = m + n$ = deg a + deg b.

Theorem 5.12 : (The Division Algorithm). Let F be a field and the polynomials $a(x)$ and $b(x)$ belong to $F[x]$, where $b[x] \neq 0$. Then there exist two polynomials $q(x)$ and $r(x)$ $\varepsilon F[x]$ such that

$$a(x) = b(x)\ q(x) + r(x)$$

where either $\quad r(x) = 0$, or $\deg r(x) < \deg b(x)$.

Proof. : Let $a(x) = \sum_{k=0}^{m} a_k x^k$ and $b(x) = \sum_{k=0}^{n} b_k x^k$ and let $b(x)$ be in canonical form with $b_n \neq 0$. Then there are two cases to consider.

Case I. If $m < n$, we can set $q(x) = 0$ and $r(x) = a(x)$, so that

$$a(x) = ob(x) + a(x)$$

where $\deg a(x) < \deg b(x)$

Case II. If $m \geq n$, we can define $a_1(x) = a(x) - b_n^{-1} a_m x^{m-n} b(x)$

$$= a(x) - q_1(x) b(x)$$

The polynomial $a_1(x)$ is of degree $m - 1$ at most, since $a(x)$ and $q_1(x)\ b(x)$ have the same leading coefficient $b_n^{-1} a_m b_n = a_m$.

We now use induction on m, repeating the process at most $m - n + 1$ times, until the remainder $a_k(x) = a(x) - [\Sigma q_i(x)] b(x)$

$$= a(x) - q(x)\ b(x)$$

has degree less than n. Then

$$a(x) = q(x)\ b(x) + a_k(x),$$

where either $a_k(x) = 0$, or $\deg a_k(x) < \deg b(x)$.

Since $F[x]$ is a ring of polynomials whose coefficients are taken from a field F, $F[x]$ is a commutative ring with unity element 1. Again $a(x)\ b(x) = 0$ will imply either $a(x) = 0$ or $b(x) = 0$. Hence $F[x]$ is a commutative ring with unity element and without a zero divisor. So we conclude that

$F[x]$ is an integral domain.

Theorem 5.13 : $F[x]$ is a Euclidean domain.

Proof. We have just observed that $F[x]$ is an integral domain. Now, for any $a(x)$ $\varepsilon F[x]$, if we assign valuation $v(a) = \deg a$, then from the result that

$$\deg(ab) = \deg a + \deg b,\ a(x), b(x)\ \varepsilon F[x]$$

we have, $\deg (ab) \geq \deg a$

i.e., $v(ab) \geq v(a)$

and also from the division algorithm we have for $a(x), b(x) \varepsilon F[x]$ with $b(x) \neq 0$, there exist $q(x)$ and $r(x)$ $\varepsilon F[x]$ such that

$$a(x) = q(x)\ b(x) + r(x)$$

where either $r(x) = 0$, or $\deg r(x) \leq \deg b(x)$

i.e., $v(r(x)) \leq v(b(x))$

The above arguments show that $F[x]$ is a Euclidean domain.

Since by Th 5.4, every Euclidean domain is a *PID*, is follows that ***F*[x] is a *PID*.**

Definition 5.12 : Let F be a field and let $F[x]$ be the ring of polynomials in x over F. A polynomial $f(x)$ $\varepsilon F[x]$ is called irreducible, if the degree of $f(x) \geq 1$ and whenever $f(x) = g(x)\ h(x)$, where $g(x), h(x)$ $\varepsilon F[x]$, then $g(x)$ εF or, $h(x)$ εF. If the polynomial is not irreducible, it is called reducible.

Remark : The irreducibility of a polynomial depends on the nature of the field.

For example, the polynomial $x^2 + 1$ is irreducible over R but reducible over C.

Theorem 5.14 : An ideal $(p(x)) \neq \{0\}$ of $F[x]$ is maximal if and only if $p(x)$ is irreducible over F.

Proof. : Suppose that $(p(x)) \neq \{0\}$ is a maximal ideal of $F[x]$. Then $(p(x)) \neq F[x]$. Again if $p(x) = a\varepsilon F$, then as $(p(x))$ is an ideal, $a^{-1}a\varepsilon(p(x))$, i.e., $1\ \varepsilon\ (p(x))$. Hence for any $f(x)$ $\varepsilon F[x]$, $1.f(x)$ $\varepsilon(p(x)$, which implies that $F[x] = (p(x))$. Since $(p(x))$ is a maximal ideal, $F[x] \neq (p(x))$, hence $p(x) \neq a\varepsilon F$. We want to show that $p(x)$ is irreducible over F.

Let $p(x) = f(x)\ g(x)$ be a factorization of $p(x)$ in $F[x]$.

Since $(p(x))$ is a maximal ideal and hence is also a prime ideal,

$$f(x)\ g(x)\ \varepsilon(p(x))$$

$$\Rightarrow f(x)\ \varepsilon(p(x)\}\ \text{or,}\ g(x)\ \varepsilon(p(x)).$$

That is either $f(x)$ or $g(x)$ has $p(x)$ as a factor. But then we connote have the degrees of both $f(x)$ and $g(x)$ less then the degree of $p(x)$. This shows that $p(x)$ is irreducible over F.

Conversely, suppose that $p(x)$ is irreducible over F and suppose that N is an ideal of $F[x]$ such that

$$(p(x)) \subseteq N \subseteq F[x]$$

Since $F[x]$ is Euclidean domain and every ideal in a Euclidean domain is principal, it follows that $N = (g(x))$ for some $g(x) \in F[x]$.

Then $(p)(x)) \subseteq N \Rightarrow p(x) \in N$

$$\Rightarrow p(x) = g(x)\ q(x) \text{ for some } q(x) \in F[x]$$

But by assumption, $p(x)$ is irreducible over F, so either $g(x)$ or $q(x)$ is of degree zero, i.e., either $g(x)$ or $q(x)$ is a non zero constant in F.

If $g(x)$ is a non zero constant in F, then $g(x)$ is a unit in $F[x]$, so $(g(x)) = N = F[x]$. Also, if $q(x)$ is a non zero constant $c \in F$, then $g(x) = c^{-1}p(x)$ is in $(p(x))$ and so

$$N = (p(x)).$$

Hence, by definition, $(p(x))$ is a maximal ideal of $F[x]$.

EXERCISES

1. Prove that $x^2 + 1$ is irreducible over the ring of integers mod 7.
2. Prove that ideal $(x^4 + 4)$ is not a prime ideal in the polynomial ring $Q[x]$ over the field of rational numbers.
3. Find the greatest common divisor of the following polynomials over Q, the field of rational numbers :

$$x^2 + 1 \text{ and } x^6 + x^3 + x + 1$$

4. Show that the ideal $(x^3 - x - 1)$ in the polynomial ring Z / (3) [x] over the field Z / (3) is a maximal ideal and therefore, *a* prime ideal.

Unit : 6 ❑ Extensions of Fields, Splitting Fields

If a field F is isomorplic to a subfield of a field G, then G is called an extension of the field F. Thus, extension of given field F corresponds to the monomorphisms from F into larger fields. The complex field C is an extension of the real field R and R is the extension of the rational field Q.

Let G and F be fields. A 1 - 1 homomorphism of the field F into a field G is called an embedding of F into G.

Henceforth, whenever there is an embedding of a field F into a field G, we say that G is an extension of F.

The complex field C consists of all complex numbers of the form $z = x + iy$, ($i = \sqrt{-1}$, $x, y \in R$).

Addition and multiplication in C are defined by the rules

$$(x_1 + iy_1) + (x_2 + iy_2) = (x_1 + x_2) + i(y_1 + y_2)$$
$$(x_1 + iy_1)\ (x_2 + iy_2) = (x_1x_2 - y_1y_2) + i(x_1y_2 + x_2y_1)$$

The assignment $x \rightarrow x + iO$ is clearly a monomorphism from R into C, which embeds R in C as a subfield. We can also consider C as a two-dimensional vector space spanned by the basis vectors $1 = 1 + iO$ and $i = O + i\ 1$.

Thus the field C is generated by i and R. This fact is expressed by writing $C = R\ [i]$.

Over any field F, the polynomials of degree n or less form an $(n + 1)$ dimensional vector space over F with basis $1, x, x^2, ..., x^n$. The commutative ring $F\ [x]$ of all polynomials over F is an infinite dimensional vector space over F. If a field G has finite dimension n, considered as a vector space over the subfield F, then G is said to be an extension of F of degree n, and we write $[G : F] = n$.

Hence, considering complex field C as an extension of degree 2 of real field R, we have $[C : R] = 2$.

In general, extensions of fields, which are generated in this way are called simple field extensions. That is, G is a simple extension of its subfield F, when, for some $c \in G$, $G = \{F, c\} = F[c]$ is generated by F and c.

Definition 6.1 : Let G be an extension of F. An element $\alpha \in$ G is said to be algebraic over F if there exist elements a_0, a_1, ... a_n ($n \geq 1$) of F, not all equal to zero, such that

$$a_0 + a_1\alpha + ... + a_n\alpha^n = 0$$

In otherwords, an elements $\alpha \in G$ is algebraic over F if there exists a nonconstant polynomical $p(x) \in F[x]$ such that $p(\alpha) = 0$.

Theorem 6.1 : If the field G is an extension of the field F of finite degree n, then every $c \in G$ is a root of a polynomial of degree at most n with coefficients in F.

Proof : Let G be any extension of finite degree $n = [G : F]$ and consider the $(n + 1)$ elements 1, c, c^2 ... c^n in G. Since $[G : F] = n$, these elements c^k must be linearly dependent over F; that is, there must exist a_0, a_2 ... $a_n \in F$, not all zero, such that

$$a_0 + a_1c + ... a_nc^n = 0$$

This shows that c is root of a polynomial of degree at most n with coefficients in F.

Theorem 6.2 : If G is a simple extension $G = F[c]$ of F of finite degree n, then $m(c) = c^n + a_1c^{n-1} + ... + a_n = 0$, all $a_i \in F$ for some monic polynomial (a monic polynomial is a polynomial with leading coefficient 1) $m(x)$ of degree $n = [G : F]$

Proof : Let G be a simple extension of the field F and let $[G : F] = n$. Then $G = F[c]$ is generated by F and some $c \in G - F$. Then by Th. 6.1, there exist a_0, a_1, ... $a_n \in F$ such that

$$a_nc^n + a_{n-1}c^{n-1} + ... + a_0 = 0 \qquad ...... (1)$$

Let a_s be the nonzero coefficient of the highest degree term in c in equation (1). Then multiplying equaion (1) by a_s^{-1} we get

$$m(c) \equiv c^s + b_{s-1}c^{s-1} + ... + b_0 = 0, \text{ all } b_i \in F \qquad ...... (2)$$

Equation (2) shows that three exists a monic polynomical $m(x)$ of degree $s(\leq n)$ such that $m(c) = 0$.

The theorem will be proved if we can show that $s = n$.

Let us choos $m(x)$ so as to minimize s. Since $G = F[c]$ is simple extension, every element $p \in G$ can be written in the form

$$p = \sum_{K=0}^{t} a_K c^K$$

of a polynomical in c with coefficients $a_K \in F$. Now, since $m(c) = 0$, we have

$$
\begin{aligned}
c^s &= -\ b_{s-1}c^{s-1} - \dots - b_0 \\
\therefore\ c^{s+1} &= c.c^s = -\ b_{s-1}c^s - b_{s-2}c^{s-1} - \dots - b_0c \\
&= -\ b_{s-1}(-b_{s-1}c^{s-1} - \dots - b_0) - b_{s-2}c^{s-1} - \dots - b_0c \\
&= d_{s-1}c^{s-1} + d_{s-2}c^{s-2} + \dots + d_0,\ d_i \in F
\end{aligned}
$$

This shows that every monomial c^t for $t \geq s$ can be expressed in terms of powers of c which are less than s. Thus every element $p \in G$ can be expressed as a linear combination of $1, c, \dots, c^{s-1}$. Hence, considered as a vector space, G will be spanned by $1, c, \dots, c^{s-1}$ and so $[G : F] \leq s$, i.e. $n \leq s$.

But we have noted above that $s \leq n$. Combining these results we get $s = n$.

Hence from equation (2) we see that

$$m(c) \equiv c^n + b_{n-1}c^{n-1} + \dots + b_0 = 0,\ b_i \in F.$$

Theorem 6.3 : In Th. 6.2, the monic polynomical $m(x)$ is irreducible and $G = F[x]/(m(x))$, i.e. G is isomerphic to the quotient ring of polynomial ring $F[x]$ over the principal ideal $(m(x))$ of all multiples of $m(x)$.

Proof : First of all we shall show that $m(x)$ is irreducible. Let $m(x) = p(x)q(x)$, $p(x)$, $q(x) \in F[x]$. Then since $m(c) = 0$, we have $p(c)q(c) = 0$. Since G is a field and hence, is an integral domain, either $p(c) = 0$ or $q(c) = 0$. But $m(x)$ has the minimum degree among all polynomials $p(x)$ in $F[x]$ such, that $p(c) = 0$.

Hence either $deg\ p(x) \geq deg\ m(x)$,

or, $\quad deg\ q(x) \geq deg\ m(x)$.

Hence $m(x)$ cannot be factored into polynomials of lower degree than $m(x)$ and so, by definition, $m(x)$ is irreducible. Next, we shall establish the second part of the theorem.

Let us define the mapping $\psi : p(x) \to p(c)$ from $F[x]$ to the set of all sums $\sum\limits_{K=0}^{v} a_K c^K$, $a_K \in F$. The mapping ψ is clearly a homomorphism. For, if $p(x)$, $q(x) \in F[x]$, then $\psi(p(x)) = p(c)$, $\psi(q(x)) = q(c)$.

Clearly $f(x) = p(x) + q(x) \in F[x]$, and

$$g(x) = p(x)\ q(x) \in F[x]$$

So, that $\psi(f(x)) = f(c)$, $\psi(g(x)) = g(c)$

Hence $\psi\{p(x) + q(x)\} = p(c) + q(c) = \psi(p(x)) + \psi(q(x))$

and $\psi(p(x) + q(x)) = g(c) = p(c)\,q(c) = \psi(p(x))\ \psi(q(x))$.

Since $G = F[c]$, ψ is an epimorphism of F $[x]$ on G. The kernel K_ψ of this epimorphism is the ideal of all polynomicals $p(x)$ with $p(c) = 0$. Since $F[x]$ is a principal idea! ring, the kernel K_ψ consists of multiples of any polynomial $p(x)$ of least degree with $p(c) = 0$.

This means that K_ψ consists of multiples of $m(x)$, i.e. $K_\psi = (m(x))$.

Since $m(x)$ is irreducible, so $(m(x))$ is a maximal ideal of $F[x]$. Hence $F[x]/(m(x))$ is a field.

Again we know that every homomorphic image of a ring R is isomorphic to some quotient ring of R, it follows that G, as the homomorphic image of $F[x]$, is isomorphic to the quotient ring $F[x]/(m(x))$ of $F[x]$.

Finite fields : By definition a finite field is a field having finite number of elements. The purpose of this section is to determine the structure of all finite fields, We shall show that for every prime p and positive integer n, there is exactly one finite field (upto isomorphism) of order p^n. This field $GF(p^n)$ is usually referred to as the Galois field of order p^n.

Theorem 6.4 : Let G be a finite extension of degree n over a finite field F. If F has m elements, then G has m^n elements.

Proof : Let $\{x_1, x_2 \dots x_n\}$ be a basis for G as a vector space over F. Then every $c \in G$ can be uniquely determined in the form

$$c = a_1x_1 + a_2x_2 + \dots + a_nx_n,\ a_i \in F$$

Since each a_i may be any of the m elements of F the total number of such distinct linear combinations of the x_i's is m^n.

Definition 6.2 : Prime field : A field is called prime if it has no proper subfield.

Clearly every field F contains a prime field—namely, the intersection of the family of its subfields, called the prime field of F.

Theorem 6.5 : If R is a ring with unity 1, then the mapping $\phi : Z \to R$ given by

$\varphi(n) = n.1$ for $n \in Z$ is a homomorphism of Z into R.

Proof : It is obvions that for $m, n \in Z$

$$\varphi(m + n) = (m + n).1 = m.1 + n.1$$
$$= \varphi(m) + \phi(n)$$

and $$\varphi(m\ n) = (mn).1 = (m.1)\ (n.1)$$
$$= \varphi(m) + \phi(n)$$

Hence φ is a homomorphism.

Corollary : If R is a ring with unity and characteristic $n > 1$, then R contains a subring isomorphic to Z_n.

If R has characteristic 0, then R contains a subring isomorphic to Z.

Proof : By Th, 6.5 the mapping $f : Z \rightarrow R$ given by $\varphi(m) = m.1$ for $m \in Z$ is a homomorphism. The kernel of ϕ must be an ideal of Z. All ideals of Z are of the form tZ for some $t \in Z$.

Now, if R has characteristic $n > 0$, then kernel of ϕ is $K_\varphi = nZ$.

Then from the fundamental theorem of homomorphism for ring, the image $\varphi(Z)$ $(\subset R)$ is isomorphic to $Z/nZ \simeq Z_n$.

If the characteristic of R is zero, then $m.1 \neq 0$ for all $m \neq 0$, so K_ϕ is $\{0\}$. Thus, the image $\phi(Z)$ $(\subseteq R)$ is isomorphic to Z.

Theorem 6.6 : A field F is either of prime characteristic p and contains a subfield isomorphic to Z_p or of characteristic zero and contains a subfield isomorphic to Q.

Proof : If the characteristic of F is not zero, then F contains a subfield isomorphic to Z_n. Then n must be a prime p or F would have zero divisors. If F is of characteristic zero then F must contain a subring isomorphic to Z. In this case F must contain a field of quotients of this subring and this field of quotients must be isomorphic to Q.

Thus every field contains a subfield isomorphic to Z_p for some prime p or a subfield isomorphic to Q.

The fields Z_p and Q are **prime fields.**

Theorem 6.7 : Every element x of $G = GF(p^n)$ satisfies the polynomial equation $x^q = x$, where $q = p^n$.

Proof : The set G^* of nonzero elements of G forms a multiplicative group of order $p^n - 1$ under the field multiplication.

For $x \in G^*$ the order of x in this group divides the order $p^n - 1$ of G^*, by Lagranges Theorem. Thus, for every $x \in G^*$ we have

$$x^{p^n-1} = 1$$

Multiplying through by x, we have

$$x^{p^n} = x, \quad \text{or, } x^q = x$$

also for $x = 0$, $x^q = x$ is trivially true.

Hence the result.

Corollary : For any $x_i \in G$, $(x - x_i) \mid (x^q - x)$ in the polynomial ring $G[x]$.

Proof : We first establish the following proposition. Given $a \in R$, a commutative ring, $(x - a) \mid p(x)$ in $R[x]$ if and only if $\tilde{p}(a) = 0$ in R, where $\tilde{p}$ is the polynomial function

$\tilde{p} : R \to R$ such that for $y \in R$

$\tilde{p}(y) \in R$

If $(x - a) \mid p(x) \in R[x]$, then $p(x) = (x - a)\, q(x)$ for some $q(x) \in R[x]$ and thus $\tilde{p}(a) = (a - a)\, \tilde{q}(u) = 0 \in R$.

Conversely for any $p(x)$

$$p(x) - \tilde{p}(a) = \sum_{K=0}^{n} p_K x^K - \sum_{K=0}^{n} p_K a^K = \sum_{K=0}^{n} p_K (x^K - a^K)$$

$$= \sum_{K=0}^{n} p_K \left[(x-a)(x^{K-1} + x^{K-2}a + \ldots + a^{K-1})\right]$$

$$= (x - a)\, q(x),\ q(x) \in R[x]$$

This shows that $(x-a) \mid \{p(x) - \tilde{p}(a)\}$ for any $a \in R$.

Consequently, if $\tilde{p}(a) = 0$, then $(x - a) \mid p(x)$

Let us now prove the statement in the corollary. Here $p(x) = x^q - x$ and since $x_i \in G$ satisfies $x^q = x$, we have, $\tilde{p}(x_i) = x_i^q - x_i = 0$ in G and hence $(x - x_i) \mid (x^q - x)$

Theorem 6.8 : In $G = GF(q)$, $q = p^n$, $x^q - x = \prod_{x_i \in G} (x - x_i)$

Proof : If $x_i \neq x_j$, then $(x - x_j, x - x_j) = 1$, hence $x - x_i$ and $x - x_j$ are relatively prime.

Since $(x - x_i) \mid (x^q - x)$ and $(x - x_i)$'s are relatively prime, it follows that

$$\left[\prod_{x_i \in G}(x - x_i)\right] / (x^q - x)$$

Since the divisor $\prod(x - x_i)$ has the same degree q as that of $x^q - x$ and since both the polynomials are monic, the two must be equal and so

$$x^q - x = \prod_{x_i \in G}(x - x_i)$$

Theorem 6.9 : In any finite field $GF(p^n)$, the multiplicative group G^* of all nonzero elements is cyclic.

Proof : Clearly, the multiplicative group G^* is abelian of order $q - 1$. Suppose that the group G^* is not cyclic. Then it follows that there would be some proper divisor r of $q - 1$ such that $x_i^r = 1$ for all $x_i \in G^*$. The argument used to prove that $\prod_{x_i \in G}(x - x_i) \mid (x^q - x)$ could then be applied to prove that

$$\prod_{x_i \in G^*}(x - x_i) \mid (x^r - x),\ r < q - 1.$$

But this is impossible, since a polynomial of degree $(q - 1)$ cannot divide a polynomial of degree $r < q - 1$. hence G^* is cyclic.

Roots of a polynomial

Definition 6.3 : Let F be any field and let $p(x)$ be any polynomial in $F[x]$. We say tht an element a lying in some extension field of F is a root of $p(x)$ if $p(a) = 0$.

Theorem 6.10 : Let $p(x)$ be a polynomial in $F[x]$ of degree $n \geq 1$ and $p(x)$ be irreducible over F. Then there exists an extension G of F such that $[G : F] = n$, in which $p(x)$ has a root.

Proof : Consider the ideal $V = (p(x))$ of $F[x]$. Since $p(x)$ is irreducible over F, V is a maximal ideal and hence $G = F[x]/V$ is a field. We shall show that G satisfies the conclusion of the theorem.

Consider the mapping ψ from $F[x]$ to $F[x]/V = G$ defined by $\psi(f(x)) = f(x) + V$.

It can be checked that y is a homomorphism. Let F^* be the image of F in G i.e. $F^* = \{\alpha + V \mid \alpha \in F\}$. Then the restriction of ψ to F induces isomorphism of F onto F^*. Using this isomorphism we can consider G to be an extension of F. Also, the elements $1 + V, x + V, x^2 + V, \dots x^{n-1} + V$ form a basis of G and so G is an extension of F of degree $n = deg\ p(a)$

Now, $f(x) \in F[x]$ and $f(x) = a_0 + a_1x + \dots + a_nx^n$

then since ψ is a homomorphism,

$$\psi(f(x)) = \psi(a_0) + \psi(a_1)\psi(x) + \dots + \psi(a_n)\psi(x^n)$$

$$= (V + a_0) + (V + a_1)\,(V + x) + \dots + (V + a_n)\,(V + x^n)$$

Since $\quad (V + x)^i = V + x^i$, we have

$$\psi(f(x)) = (V + a_0) + (V + a_1)\,(V + x) + \dots + (V + a_n)\,(V + x)^n$$

Since the element $a_i \in F$ corresponds to the element $a_i + V \in G$, it follows that

$$\psi(f(x)) = f(V + x) \qquad \dots\dots (1)$$

Now, $\quad p(x) \in V$, so that $\psi(p(x)) = 0 \qquad \dots\dots (2)$

but since $p(x) \in F[x]$, we have, by (1)

$$\psi(p(x)) = p(V + x)$$

From (1) and (2) it follows that

$$p(V + x) = 0.$$

This shows that the element $V + x \in G$ is a root of $p(x)$. Hence the theorem is proved.

Corollary : If $f(x) \in F[x]$ then there is a finite extension G of F in which $f(x)$ has a root. Moreover, $[G : F] \leq deg\, f(x)$

Proof : Let $p(x)$ be an irreducible factor of $f(x)$; any root of $p(x)$ is a root of $f(x)$. So the $deg\ p(x) \leq deg\ f(x)$.

By theorem 6.10, there is an extension G of F with $[G : F] = deg\, p(x)$ such that $p(x)$ has a root in G. It follows that there is an extension G of F with degree $[G : F] \leq deg\ f(x)$ such that $f(x)$ has a root in G.

Theorem 6.11 : If G_1 is a finite extension field of a field F and G_2 is a finite extension of G_1, then G_2 is a finite extension of F and

$$[G_2 : F] = [G_2 : G_1]\,[G_1 : F]$$

Proof : Let $\{a_i \mid i = 1, 2, \ldots n]$ be a basis for G_1 as a vector space over F and let $\{\beta_j \mid j = 1, 2, \ldots m\}$ be a basis for G_2 as a vector space over G_1.

Let γ be any element of G_2. Since the β_j form a basis for G_2 over G_1, we have

$$\gamma = \sum_{j=1}^{m} b_j \beta_j,\ b_j \in G_1,$$

Since α_i form a basis for G_1 over F, we have

$$b_j = \sum_{i=1}^{n} a_{ij}\alpha_i,\ a_{ij} \in F$$

Then $$\gamma = \sum_{j=1}^{m}\left(\sum_{i=1}^{n} a_{ij}\alpha_i\right)\beta_j = \sum_{i,j} a_{ij}(\alpha_i\beta_j)$$

So the mn vectors $\alpha_i\beta_j$ span G_2 over F. We shall show that the mn vectors $\alpha_i\beta_j$ are linearly independent over F.

Let us assume that $\sum C_{ij}(\alpha_i\beta_j) = 0,\ C_{ij} \in F$.

Then $$\sum_{j=1}^{m}\left(\sum_{i=1}^{n} C_{ij}\alpha_i\right)\beta_j = 0$$

and $$\left(\sum_{i=1}^{n} C_{ij}\alpha_i\right) \in G_1$$

Since the vectors β_j are linearly independent, we have

$$\sum_{i=1}^{n} C_{ij}\alpha_i = 0 \text{ for } j = 1, 2 \ldots m.$$

But now, α_i are linearly independent over F, so $\sum_{i=1}^{n} C_{ij}\alpha_i = 0 \Rightarrow C_{ij} = 0$ for all i and j.

Thus, the mn vectors $\alpha_i\beta_j$ are linearly independent and they span G_2 over F. Hence they form a basis for G_2 over F.

Hence it follows that $[G_2 : F] = mn = [G_2 : G_1]\,[G_1 : \mathrm{F}]$.

Theorem 6.12 : Let $f(x) \in F[x]$ be of degree $n \geq 1$. Then there is an extension G of F of degree at most $n\,!$ in which $f(x)$ has n roots.

Proof : We note here that in the statement of the theorem, a root of multiplicity m is counted as m roots.

By the corollary to Th. 6.10, there is an extension G_0 of F with $[G_0 : F] \leq n$ in which $f(x)$ has a root α. Thus, in $G_0[x]$, $f(x) = (x - a)\, q(x)$, where $q(x)$ is of degree $(n - 1)$.

Using induction, there is an extension G of G_0 of degree at most $(n - 1)!$ in which $q(x)$ has $(n - 1)$ roots. Since any root of $f(x)$ is either a or a root of $q(x)$ we obtain, in G all n roots of $f(x)$.

Now, $[G : F] = [G : G_0]\, [G_0 : F] \leq (n - 1)!\; n = n\,!$

Definition 6.4 : If $f(x) \in F[x]$, a finite extension G of F is said to be a splitting field over F for $F[x]$ if over G (i.e. in $G[x]$), but not over any proper subfield of G, $f(x)$ can be factored as a product of linear factors.

Th. 6.11 guarantees the existence of splitting fields.

Alternatively, we may say that G is a splitting field of $f(x)$ over F if G is a minimal extension of F in which $F(x)$ has n roots, where $n = deg\, f(x)$.

Example 6.1 : If Q denotes the field of rational numbers, determine the degree of the splitting field of the polynomial $x^3 - 2$ over Q.

Let $f(x) = x^3 - 2 \in Q[x]$

The roots of $f(x)$ are $2^{1/3}$, $\omega 2^{1/3}$, $\omega^2 2^{1/3}$, $\omega = (-1 + \sqrt{3}\, i)2$. Here $2^{1/3}$ is the only real root of $f(x)$. Evidently $f(x)$ is irreducible over the field q of reational numbers.

Also, $deg\, f(x) = 3$.

It follows that the extension field $Q(2^{1/3})$ of Q is such that

$$[Q(2^{1/3}) : Q] = 3$$

Suppose that G is a splitting field *of* $f(x)$ over Q, Now the field $Q(2^{1/3})$ cannot split $f(x)$, since, as a subfield of real field it cannot contain the complex number $\omega 2^{1/3}$. Hence $Q(2^{1/3})$ must be a proper subfield of G. Therefore,

$$[G : Q] > [Q(2^{1/3}) : Q] = 3$$

Also, $[G : Q] \leq 3! = 6$

Again, $\{G : Q\} = [G : Q(2^{1/3})]\, [Q(2^{1/3}) : Q]$

This shows that $[Q(2^{1/3}) : Q]$ is a divisor of $[G : Q]$. All these can hold if $[G : Q] = 6$.

BIBLIOGRAPHY

1. Herstein, I.N : Topics in algebra, Vikas Publishing House Pvi. Ltd.
2. Bhatiacharya, P.B., Jain, S.K. and Nagpaul S.R. : Basic abstract algebra, Cambridge University Press.
3. Fraleigh, John B : A first course in abstract algebra, Narosa publishing house.
4. BirkhotT, G and Bailee, T.C. : Modern applied algebra, McGravv Hill Book Company.
5. Ayres, F : Theory and problems of modern algebra, Schaum's outline series
6. Singh, Surjeet and Zameeruddin, Q : Modem algebra, Vikas Publishing house Pvi. Ltd.
7. Lang, Serge : Algebra, Addison - Wesley Publishing Company Inc
8. Artin, Michael : Algebra, Prentice-Hal! of India Pvt. Ltd.

PG (MT) 01 : Group B

Unit : 1 ❑ Vector Spaces

Introduction : In the study of mathematics we encounter many examples of mathematical objects that can be added to each other and multiplied by real numbers. First of all, the real numbers themselves are such objects. Other examples are real valued functions, the complex numbers, infinite series, vectors in n-dimensional spaces, and vector valued functions. Here we introduce a general mathematical concept, called a vector space, which includes all these examples and many others as special cases.

We have already studied algebraic structures like groups and rings ; the former has its motivation in the set of one-to-one mappings of a set onto itself, the latter, in the set of integers. The present algebraic structure which we are about to consider—vector space—can, in large part trace its origins to topics in grometry and physics.

A vector space differs from the structures of groups and rings due to the fact that one of the product defined on it uses elements outside of the set itself.

The importance of vector spaces lies on the fact that so many models arising in the solutions of specific problems turn out to be vector spaces. For this reason, the basic concepts introduced in them have a certain universality. The fundamental motions like linear dependence, basis and dimension which will be discussed here are potent and effetive tools in all branches of mathematics.

Briefly, a vector space is a set of elements of any kind on which certain operations can be performed. In defining a vector space, we do not specify the nature of the elements, nor do we tell how the operations are to be performed on them. Instead, we require that the operations have certain properties that we take as axioms for a vector space.

The precise definifion of a vector space is given below :

Definition 1.1 : A nonempty set V is said to be a vector space over a field F if V is an abelian group under an operation which we denote by +, and if for every $a \in F$, $\alpha \in V$ there is an element, written $a\,\alpha$, in V subject to

1. $a\,(\alpha + \beta) = a\,\alpha + a\beta$
2. $(a + b)\,\alpha = a\,\alpha + b\,\alpha$
3. $a(b\alpha) = (ab)\alpha$

4. $1\alpha = \alpha$

for all $a, b, \in F$, and $\alpha, \beta \in V$. Here 1 repersents the unit element of F under multiplication.

We shall call elements $\alpha. \beta, \ldots$ of the vector space V as vectors and the elements of the field F as scalars.

Note that in Axions 1 above, the operation '+' is that of V, whereas on the left hand side of Axiom 2, the '+' is that of field F and on the right hand side, it is of V.

EXAMPLES

Ex. 1. Let F be a field and V be the totality of ordered n-tuples $(a_1, a_2, \ldots a_n)$ where tha $a_i \in F$, for $i = 1, -, n$, Two elements $(a_1, a_2, \ldots a_n)$ and $(b_1, b_2, \ldots b_n)$ of V are declared to be equal if and only if $a_i = b_i$ for each $i = 1, 2, \ldots n$, We define operation in V such that

(i) $(a_1, a_2, \ldots a_n) + (b_1, b_2, \ldots b_n) = (a_1 + b_1, a_2, + b_2, \ldots a_n, + b_n)$

(ii) $c\,(a_1, a_2, \ldots a_n) = (ca_1, ca_2, \ldots ca_n)\ c \in F$

It is easy to verity that with these operations, V is a vector space over F

Ex. 2. Let F be any field and let $V = F[x]$, be the set of polynomials in x over F. In V, two polynomials can be added and a polynomial can always be multiplied by an element of F. With these natural operations, $V = F\,[x]$ is a Vector space over F.

Ex. 3. Let V be the set of all real-valued functions defined on the interval $[a, b] = \{x \mid a \leq x \leq b\}$.

For two functions f and $g \in V$, we define $f = g$ if and only if $f(x) = g(x)$, $\forall\ x \in [a,b]$

Addition in V is defined by

$$(f + g)\,(x) = f(x) = g(x)$$

and scalar multiplication, by

$$(cf)\,(x) = c(f(x)),\ c \in R, \text{ the set of real numbers}$$

The zero vector in V is the function 0 whose value is 0 for every x and the additive inverse of f is the function $-f$ which satisfies

$$(-f)\,(x) = -(f(x))$$

It is easy to verify any of the remaining axioms of the vecter space for the set V. For example, to show that addition in V is associative, we calculate

$$\begin{aligned} ((f + g) + h)(x) &= (f + g)(x) + h(x) \\ &= [f(x) + g(x)] + h(x) \\ \text{and } (f + (g + h)(x) &= f(x) + (g + h)(x) \\ &= f(x) + [g(x) + h(x)] \end{aligned}$$

and then observe that the last two expressions are equal by associativity in R.

Thus $f + (g + h) = (f + g) + h$.

Note : If the field F is the field of real numbers, the vector space V will be called a real vector space. Similarly, a complex vector space is defined analogously by using the field of complex numbers for F.

Ex. 4. Let V be the set of all real valued functions defined on $[a, b]$ and having derivatives of all orders on $[a, b]$. We may verify that V is a vector space over R under usual operations. It is the space of infinitely differentiable functions on $[a, b]$.

Note : We observe that for a vertor space V defined over a field F there is the zero element of the field F as well as the zero element of the abelian group V. To aviod confusion, we shall denote by 0, the 0 the zero element of the field F and by θ, the zero element of the vector space V.

Theorem 1.1: In a vector space V over a field F the following results hold.

1. $0\alpha = \theta$ *for all* $\alpha \in V$
2. $c\theta = \theta$ *for all* $c \in F$
3. $-1\ \alpha = -\alpha,\ \alpha \in V$
4. $c\alpha = \theta \Rightarrow c = 0$ *or* $\alpha = \theta$.

Proof : 1. In F, we have $0 + 0 = 0$

$\therefore$ For any $\alpha \in$ V, we have

$(0 + 0)\ \alpha = 0\alpha$

Or, $0\alpha + 0\alpha = 0\alpha$ by axiom 2 of vector space

Or, $0\alpha + 0\alpha = 0\alpha + \theta$

$\Rightarrow 0\alpha = \theta$, by left cancellation rule in the abelian group.

2. We have, in V, $\theta + \theta = \theta$

$\therefore$ For any $c \in F$,

$c(\theta + \theta) = c\theta$

By axiom 1 of vector space, $c\,\theta + c\,\theta = c\,\theta$

Or, $c\theta + c\theta = c\theta + \theta$

So by left concellation rule in V, $c\,\theta = \theta$.

3. We have in F, $0 = 1 + (-1)$

$\therefore$ For any $\alpha \in V$, $0\,\alpha = (1 + (-1))\,\alpha$

$\Rightarrow$ $0\,\alpha = 1\,\alpha + (-1)\,\alpha$

Now $0\alpha = \theta$ *and* $1\alpha = \alpha$, *so*

$\theta = \alpha + (-1)\,\alpha$

Or, $-\alpha + \theta = \alpha + \alpha + (-1)\,\alpha$

$\Rightarrow$ $-\alpha = (-1)\,\alpha$

4. Let $c\,\alpha = \theta$, and suppose that $c \neq 0$. Then $\overset{-1}{c}$ exists in F,

Hence $\overset{-1}{c}\,(c\,\alpha) = \overset{-1}{c}\,\theta$

By result 1 of Theorem 1.1

$\Rightarrow$ $(\overset{-1}{C}\,c)\,\alpha = \theta$, by axiom 3 of vector space

$\Rightarrow$ $1\,\alpha = \theta$

$\Rightarrow$ $\alpha = \theta$, by axiom 4 of vector space

Again, if $c = 0$, then $c\,\alpha = \theta$ for any $\alpha \in V$

Therefore, $c\,\alpha = \theta$ implies either $c = 0$ or $\alpha = \theta$.

Definition1.2 : If B is a vector space over a field F and $W \subseteq V$, then W is a subspace of V if under the operations of V, W itself a forms vector space over F.

It is clear that $\{\theta\}$ and V are both subspaces of V. These are trivial subspaces. Any subspace of a vector space V other than $\{\theta\}$ and V itself is called a proper subspace of V.

Theorem 1.2 : A non- empty subset W of a vector space V over a field F is a subspace of V if and only if

(i) $\alpha \in W, \beta \in W \Rightarrow \alpha + \beta \in W$

(ii) $\alpha \in W, c \in F \Rightarrow c\alpha \in W$,

Proof : Let us suppose that the conditions (i) and (ii) hold in W. Let $\alpha, \beta \in W$. Since F is a field, $1 \in F$ and so $-1 \in F$.

By condition (ii), we have $(-1)\beta \in W$

$$\text{or, } -\beta \in W$$

Hence by condition (i), $\alpha + (-\beta) \in W$

$$\text{or, } \alpha - \beta \in W \quad \text{whenever } \alpha, \beta \in W$$

This shows that $(W, +)$ is a subgroup of the additive group $(V. +)$. But since V is abelian, W is abelian. Hence using condition (ii) and the heredity property, we can conclude that W is a vector space over the field F. Hence W is a subspace of V over F. Thus the sufficiency of the conditions is established.

The necessity of the conditions (i) and (ii) follows from the definition of a vector space.

Note : The above theorem may be stated in the alternative form giving only one condition as follows :

A nonempty set W of a vector space V over a field F is a subspace of V if and only if $a\,\alpha + b\beta \in W$ for all $\alpha, \beta \in W$ and all $a, b \in F$.

EXAMPLES

Ex. 1. Let S be the subset fo R^3 defined by

$S = \{(x, y, z) \in R^3 \mid y = z = 0\}$. Then S is a nonempty subset of R^3, since $(0, 0, 0) \in S$.

Let $\alpha = (a, 0, 0)$ and $\beta = (b, 0, 0) \in S$, where $a, b \in R$

Then $\alpha + \beta = (a, 0, 0,) + (b, 0, 0,) = (a + b, 0, 0) \in S$

Thus, $\alpha \in S, \beta \in S \Rightarrow \alpha + \beta \in S$

Let $c \in R$, then $c\,\alpha = c\,(a, 0, 0) = (ca, 0, 0) \in S$

Hence, by definition, S is a subspace of R^3.

Ex. 2. Let $V = R^3$ and $W \subseteq V$ such that $W = \{(a, b, c) \in Q : a, b, c\}$ i.e. W consists of those vectors whose components are rational numbers.

Clearly, $\alpha = (4, 2, 3) \in W$, and $\sqrt{2} \in \mathrm{R}$.

But $\sqrt{2}\,\alpha = \sqrt{2}(4, 2, 3) = (4\sqrt{2}, 2\sqrt{2}, 3\sqrt{2}) \notin W$

Hence the condition (ii) for being a subspace is violated. So W as defined above is not a subspace of V.

Ex. 3. Let V be the vector space of all functions from the real field R into R, and let

$W = \{f : f(-x) = -f(x), x \in R\}$, i.e. W consists of the odd real valued functions defined over R.

The function f which assigns θ to every $x \in R$, is the zero function, denoted by O. Thus $O(x) = 0 \ \forall x \in R$

Then clearly, $O(-x) = 0 = -0 = -O(x)$

This implies that the zero function $O \in W$.

Suppose $g(x)$ and $h(x) \in W$, then $g(-x) = -g(x)$, $h(-x) = -h(x)$. Then for any $a, b \in R$, we have

$$\begin{aligned}(ab + bh)(-x) &= a\,g(-x) + bh\,(-x) = -ag(x) - bh(x)\\ &= -(ag\,(x) + bh(x))\\ &= -(ag + bh)\,(x)\end{aligned}$$

Hence $g, h, \in W \Rightarrow ag + bh \in W$; $a, b \in R$

So W is a subspace of V over the field of real numbers R.

Theorem 1.3 : Intersection of two subspaces of a vector space V over a field F is a subspace of V over F.

Proof : Let W_1 and W_2 be two subspaces of a vector space V over a field F. Let $W = W_1 \cap W_2$

Suppose that $\alpha, \beta \in W$, Then $\alpha, \beta \in W_1$ and W_2. But W_1 and W_2 are subspaces of V. Hence,

$$\alpha, \beta \in W_1 \text{ and } W_2$$

and for any $c \in F$, $c\alpha \in W_1$, and W_2

So, for $\alpha, \beta \in W$, we have $\alpha + \beta \in W_1 \cap W_2 = W$ and $c\,\alpha \in W_1 \cap W_2 = W$, $c \in F$

Hence W is a subspace of V.

Note : The union of two subspaces of a vector space is not, in general, a subspace of V.

Let W_1 and W_2 be the two subspaces of the vector space R^2 such that

$$W_1 = \{\, (x, y), \in R^2 \,|\, y = 0 \,\},$$

and $$W_2 = \{\, (x, y) \in R^2 \,|\, x = 0 \,\}$$

Let $\alpha = (1, 0)$ and $\beta\ (0, 1)$.

Clearly, $\alpha \in W_1$ and $\beta \in W_2$

Now, $\alpha + \beta = (1, 0) + (0, 1) = (1.1)$

But $\alpha + \beta = (1.1) \notin W_1$ and $\alpha + \beta \notin W_2$

Hence $\alpha + \beta \notin W_1 \cup W_2$, although $\alpha \in W_1 \cup W_2$ and $\beta \in W_1 \cup W_2$.

This shows that $W_1 \cup W_2$ is not a subspace of V.

Definition 1.3 : Let V be a vector space over a field F. Let $\alpha_1, \alpha_2 \ldots \alpha_n \in V$ and $c_1, c_2 .. c_n \in F$, then

$$c_1\alpha_1 + c_2\alpha_2 + \ldots + c_n\alpha_n$$

is called a linear conbination of $\alpha_1, \alpha_2 \ldots \alpha_n$

Definition 1.4 : Let $S \subseteq V$. The linear span of a subset S of the vector space V, denoted by $L(S)$, is the set of all linear combinations of the vectors in S.

In otherwords, if S is a subset of V, then the linear span of S is the set

$\{c_1\alpha_1 + c_2\alpha_2 + \ldots + c_n\alpha_n \mid c_1, c_2 \ldots c_n$ are scalars $\in F$, n is any positive integer and $\alpha_1, \alpha_2 \ldots \alpha_n \in S\}$

Theorem 1.4 : Let S be a nonempty subset of a vector space V. Then the linear span $L(S)$ of S is a subspace of V and it is the smallest subspace containing the set S.

Proof : Let $\alpha = c_1\alpha_1 + c_2\alpha_2 + \ldots + c_n\alpha_n \qquad \alpha_i \in S, c_i \in F$

$i = 1, 2 \ldots n$

and $\beta = d_1\beta_1 + d_2\beta_2 + .. + d_m\beta_m \qquad \beta_j \in S, d_j \in F$

$j = 1. 2 \ldots m.$

be any two elements of $L(S)$, To prove the theorem we have to prove that $\alpha + \beta \in L(S)$ and for any $c \in F$, $c\,\alpha \in L(S)$ [see Theorem 1.2]

We have, $\alpha + \beta = c_1\alpha_1 + .. + c_n\alpha_n + d_1\beta_1 + \ldots + d_m\beta_m$

Clearly, $\alpha + \beta$ is a finite linear combination of the vectors $\alpha_1, \alpha_2 \ldots \alpha_n, \beta_1, \beta_2 \ldots \beta_m \in S$, and so $\alpha + \beta \in L(S)$

Again, $\quad c\alpha = cc_1\alpha_1 + cc_2\alpha_2 + .. + cc_n\alpha_n$

$\qquad = c_1'\alpha_1 + c_2'\alpha_2 + .. + c_n'\alpha_n,\ c_i' = cc_i$

Since $c_i' = cc_i \in F$ for $i = 1, 2 \ldots n$, it follows that $c\,\alpha \in L(S)$ Hence $L(S)$ is a subspace of V.

To show that $L(S)$ is the smallest subspace containing S, we consider any other subspace W of V containing S and we shall show that $L(S) \subseteq W$

Let $\xi \in L(S)$. Then $\xi = x_1\alpha_1 + x_2\alpha_2 + .. + x_n\alpha_n$ for some x_i (i =1. 2 .. n) $\in F$.

Since W is a subspace of V containing α_i and $x_i \in F$, $x_i\alpha_i = W$ for i =1. 2 .. n

Since W is a subspace and $x_i\alpha_i \in W$, follows that

$$x_1\alpha_1 + x_2\alpha_2 + .. + x_n\alpha_n \in W$$

or, $\xi \in W$

Thus $\xi \in L(S) \Rightarrow \xi \in W$

Hence $L(S) \subseteq W$

In other words, $L(S)$ is the smallest subspace of V containing S.

The subspace $L(S)$ is called the subspace spanned or generated by S.

EXAMPLES

Ex. 1. Let $S = \{(1, 2, 1), (1, 1, -1), (4, 5, -2)\}$. Examine if the vector $(2, -1, -8)$ is in $L(S)$.

The vector $(2, -1, -8) \in L(S)$, if there exist scalars c_1, c_2, c_3 such that $(2, -1, -8) = c_1(1, 2, 1) + c_2(1, 1, -1) + c_3(4, 5, -2)$

i.e. if $(2, -1, -8) = (c_1 + c_2 + 4c_3, 2c_1 + c_2 + 5c_3, c_1 - c_2 - 2c_3)$

This will be true if

$$c_1 + c_2 + 4c_3 = 2$$
$$2c_1 + c_2 + 5c_3 = -1$$
$$c_1 - c_2 - 2c_3 = -8$$

Solving we get $c_1 = -\frac{14}{3}$, $c_2 = 0$, $c_3 = \frac{5}{3}$

This shows that the vector $(2, -1, -8)$ is a linear combination of the vectors (1, 2, 1) and) (4, 5, -2) such that

$$(2, -1, -8) = -\frac{14}{3}(1, 2, 1) + \frac{5}{3}(4, 5, -2)$$

Hence the vector $(2, -1, -8) \in L(S)$.

Ex. 2. Show that the yz plane $W = \{(0, y, z)\}$ in R^3 is generated by $(0, 1, 1)$ and $(0, 2, -1)$

To show this we have to find real numbers c_1 and c_2 such that, any vector $\alpha = (0, y, z) \in W$ can be expressed as a linear combination of $(0, 1, 1)$ and $(0, 2, -1)$. That is, to find c_1, c_2 such that

$$(0, y, z) = c_1 (0, 1, 1) + c_2 (0, 2, -1)$$

This is true if

$$y = c_1 + 2c_2$$

and $$z = c_1 - c_2$$

$$\Rightarrow \quad c_1 = \frac{y+2z}{3} \qquad c_2 = \frac{y-z}{3}$$

Thus for any given vector $\alpha = (0, y, z)$ α can be expressed ans a linear combination of vectors $(0, 1, 1)$ and $(0, 2, -1)$. Hence W is generated by $(0, 1, 1)$ and $(0, 2, -1)$.

Definition 1.5: Let U and W be subspaces of a vectorspace V. The sum of U and W. written as $U + W$ is defined as follows

$$U + W = \{ u + w \mid u \in U, w \in W\}$$

Theorem 1.5 : The sum $U + W$ of the subspaces U and W of V is a subspace of V.

Proof : Since $0 \in U$ and $0 \in W$, $\quad 0 = 0 + 0 \in U + W$.

Further, suppose that $u + w \in U + W$ and $u' + w' \in U + W$ with $u, u' \in U$ and $w, w' \in W$.

Then $(u + w) + (u' + w') = (u + u') + (w + w') \in U + W$

and for any scalar c, $c(u + w) = cu + cw$

Since $cu \in U$ and $cw \in W$, we have $c(u + w') \in U + W$.

Hence $U + W$ is a subspace of V.

Definition 1.6 : The vector space V is said to be the direct sum of its subspaces U and W, denoted by

$$V = U \oplus W$$

if every $v \in V$ can be written in one and only one way sa $v = u + w$ where $u \in U$ and $w \in W$.

EXAMPLES

Ex. 1. If $V^3 (R)$ is a vector space over R and $W_1 = \{ (a, b, 0) : a, b \in R\}$ and $W_2 = \{(0, 0, c) : c \in R\}$ are two subspaces of $V^3(R)$, then

$$V^3(R) = W_1 \oplus W_2.$$

In the context of direct sum we state the following theorem.

Theorem 1.6 The vector space V is the direct sum of its subspaces U and W if and only if

$$(i)\ V = U + W$$

and $$(ii)\ U \cap W = \{\theta\}$$

Proof :

Let us first assume that

$$V = U \oplus W.$$

Then by definition, every $\alpha \in V$ can be expressed uniquely as

$$\alpha = \alpha_1 + \alpha_2 \text{ with } \alpha_1 \in U, \alpha_2 \in W.$$

Hence we must have

$$\alpha \in U + W$$

Let $\beta \in U \cap W$. As $\beta \in U$, $\beta = \beta + \theta$, say, where $\beta \in U$, $\theta \in W$.

Also $\beta \in W$, so $\beta = \theta + \beta$, say, where $\theta \in U$, $\beta \in W$.

Hence by the uniqueness theorem,

$$\beta \equiv \theta,$$

$$\therefore \quad U \cap W = \{\theta\}$$

Conversely, let

(i) $V = U + W$ and (ii) $U \cap W = \{\theta\}$

If possible, let any $\alpha \in V$ be expressed in two ways, nemely, $\alpha = \alpha_1 + \alpha_2$ and $\alpha = \beta_1 + \beta_2$ with $\alpha_1, \beta_2 \in U$, $\alpha_2, \beta_2 \in W$.

Then $\alpha_1 - \beta_1 = -(\alpha_2 - \beta_2)$

belongs to both U and W. Consequntly Then $\alpha_1 - \beta_1 = -(\alpha_2 - \beta_2) = \theta$, by hypothesis.

$\therefore \alpha_1 = \beta_1, \ \alpha_2 = \beta_2$

Thus every $\alpha \in V$ can be expressed uniquely as $\alpha = \alpha_1 + \alpha_2$ with $\alpha_1 \in$ U, $\alpha_2 \in$ W.

Consequently, V = U ⊕ W.

This completes the proof.

Linear dependence and linear independence of vectors

Definition 1.7 : For a vector space V defined over a field F, the n vectors α_1, α_2 .. α_n of V are said to be linearly dependent if there exists a set of scalars $c_1, c_2 \ldots c_n$ of F, not all zero (where zero is the additive identity of F), such that

$$c_1\alpha_1 + c_2\alpha_2 + \ldots + c_n\alpha_n = \theta$$

Definition 1.8 : For a vector space V defined over a field F, the n vectors α_1, α_2 .. α_n of V are said to be linearly independent if and only if

$$c_1\alpha_1 + c_2\alpha_2 + .. + c_n\alpha_n = \theta, \qquad c_i \in F \qquad (i = 1.\ 2 \ldots n)$$

implies that $c_1 = c_2 = \ldots = c_n = 0$.

EXAMPLES

Ex. 1. The condinate vectors $\alpha_1 = (1, 1, 0)$, $\alpha_2 = (3, 2, 1)$ and $\alpha_3 = (2, 1, 1)$ are linearly dependent if there exists a set of scalars c_1, c_2, c_3, not all zero, such that

$c_1 (1, 1, 0) + c_2 (3, 2, 1) + c_3 (2, 1, 1) = (0, 0, 0)$;

This requires that

$$c_1 + 3c_2 + 2c_3 = 0$$
$$c_1 + 2c_2 + c_3 = 0$$
$$c_2 + c_3 = 0$$

The system of homogeneous linear equations has non zero solution as the rank of the coefficient matrix is 2(<3). We may also directly solve to check thar $c_1 = 1$, $c_2 = -1$, $c_3 = 1$ is a solution to the system. Hence

$$(1)\ \alpha_1 + (-1)\ \alpha_2 + (1)\ \alpha_3 = 0$$

Thus vectors $\alpha_1, \alpha_2, \alpha_3$ are linearly dependent and any one of the vectors can be written as a linear combination of the other two. For example, $\alpha_1 = \alpha_2 - \alpha_3$.

Ex. 2. The coordinate vectors (3, 2, 1), (0, 1, 2) and (1, 0, 2) are linearly indepenedent.

Suppose that c_1, c_2, c_3 are scalrs such that

$$c_1(3, 2, 1) + c_2(0, 1, 2) + c_3(1, 0, 2) = \theta = (0, 0, 0)$$

This requires that

$$3c_1 + c_3 = 0$$
$$2c_1 + c_2 = 0$$
$$c_1 + 2c_2 + 2c_3 = 0$$

It may be checked that the rank of the coefficient matrix is 3 = number of variables. Hence the only solution for the system of equations is $c_1 = c_2 = c_3 = 0$. (We can also directly obtain this solution)

Hence, by definition, it follows that the given vectors are linealy independent.

Basis of a vectors space

We have learnt that a given set of vectors spans or generates a certain vector space when every vector in the space can be expressed as a linear combination of the given set. Obviously there can be more than one set that can span a certain vector space; moreover the number of vectors in each generating set can vary.

Ex. 3. The set {(1, 0), (0, 1)} spans the vector space of all two dimensional real coordinate vectors ; also, the set ((1,2); (2,1), (3, 3)} spans the same vector space. Every vector in the space can be expresed as a linera combination of each of the given sets, and therefore each set is a spanning set even though the sets are different in both content and number of elements.

There is, however, a distinction between two kinds of spanning sets; that is, for a given vector space, some spanning sets are linearly independent and some are linearly dependent. Those spanning sets that are linearly independent are very important in the study of linear algebra ; such a set is called a basis of the vector space.

Definifition 1.9 : A basis of a vector space is a set of vectors of the space that (i) are linearly independent and (ii) span the vector space.

EXAMPLES

Ex. 1. The set {(1, 0, 0), (0, 1, 0), (0, 0, 1)} is a basis for the vector space of all coordinate vectors with three real components over the field of real numbers because the set is linearly independent and the set spens the specce.

Ex. 2. The set $\{(0, 1), (1, 0), (1, 1)\}$ is not a basis of the vector space of two dimensional real coordinate vectors over the field of real numbers, because the set is not linearly independent. The set does, however, span the space.

Ex. 3. The set $\left\{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}\right\}$ is not a basis for the vector space of all 2 × 2 real matrices, because the set does not span the space. The set is, however, linearly independent.

The set $\left\{\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right\}$ is a basis of the above vector space.

Theorem 1.7 If $\{\beta_1, \beta_2, .. \beta_n\}$ is a basis of a vector space V over a field F. then each vector in V can be expressed uniquely as a linear combination of $\beta_1, \beta_2, .. \beta_n$.

Proof : Since $\{\beta_1, \beta_2, .. \beta_n.\}$ spans the vector space V, any vector $\beta \in V$ can be expressed as a linear combination of $\beta_1, \beta_2, .. \beta_n$.

Let us suppose that

$$\beta = a_1\beta_1 + \ldots + a_n\beta_n \text{ and } \beta = b_1\beta_1 + \ldots + b_n\beta_n.$$

where $a_i' s$ and $b_i' s \in F$, are two such expressions of β. Theorem will be proved if we can show that $a_i = b_i$, $\quad i = 1, 2, \ldots n$

We have $\beta = a_1\beta_1 + a_2\beta_2 + .. + a_n\beta_n = b_1\beta_1 + b_2\beta_2 + .. + b_n\beta_n$

$$\Rightarrow (a_1 - b_1)\,\beta_1 + (a_2 - b_2)\,\beta_2 + .. + (a_n - b_n)\,\beta_n = \theta$$

Since $\beta_1, \beta_2, .. \beta_n$ are linearly independent, it follows that $a_1 - b_1 = 0, a_2 - b_2 = 0, .. a_n - b_n = 0,$

$$\therefore a_i = b_i, \; i = 1, 2, \ldots n.$$

Hence the theorem is proved.

Theorem 1.8 (Replacement Theorem). If $\{\alpha_1, \alpha_2 .. \alpha_n\}$ be a basis of a vector space V over a field F and a nonzero vetor β of V is expressed as $\beta = c_1\alpha_1 + c_2\alpha_2 + .. + c_n\alpha_n$, $c_i \in F$, then if $c_j \neq 0$, $\{\alpha_1, \alpha_2 \ldots \alpha_{j-1,}\, \beta, \alpha_{j+1} .. \alpha_n\}$ is a new basis of V. [The is, β can replace α_j in the basis]

Proof : $\beta = c_1\alpha_1 + c_2\alpha_2 + .. + c_{j-1}\alpha_{j-1} + c_j\alpha_j + c_{j+1}\alpha_{j+1} + .. + c_n\alpha_n$

$$\therefore c_j\alpha_j = \beta - c_1\alpha_1 - c_2\alpha_2 - - c_{j-1}\alpha_{j-1} - c_{j+1}\alpha_{j+1} - - c_n\alpha_n \; .... (1)$$

Since $c_j \neq 0$, c_j^{-1} exists in F

$\therefore$ Multiplying both sides of (1) by c_j^{-1} we get

$$\alpha_j = c_j^{-1}\beta - c_j^{-1}c_1\alpha_1 - c_j^{-1}c_2\alpha_2 - \dots - c_j^{-1}c_{j-1}\alpha_{j-1} - c_j^{-1}c_{j+1}\alpha_{j+1} - \dots - c_j^{-1}c_n\alpha_n$$
$$= p_1\alpha_1 + p_2\alpha_2 + p_{j-1}\alpha_{j-1} + p_{j+1}\alpha_{j+1} + \dots + p_n\alpha_n$$

where $p_i = -c_j^{-1}c_i$, *if* $i \neq j$

$$= c_j^{-1}, \textit{ if } i = j$$

We first prove that $\{\alpha_1, \alpha_2\ \alpha_{j-1}, \beta, \alpha_{j+1}, .. \alpha_n\}$ is linearly independent.

Let us suppose that

$$d_1\alpha_1 + d_2\alpha_2 + .. + d_{j-1}\alpha_{j-1} + d_j\beta + d_{j+1}\alpha_{j+1} +.. + d_n\alpha_n = \theta$$

$\therefore d_1\alpha_1 + d_2\alpha_2 + . + d_{j-1}\alpha_{j-1} + d_j(c_1\alpha_1 + c_2\alpha_2 + .. + c_{j-1}\alpha_{j-1} + c_j\alpha_j + c_{j+1}\alpha_{j+1} +.. + c_n\alpha_n) + d_{j+1}\alpha_{j+1} +.. + d_n\alpha_n = \theta$

or, $(d_1 + d_jc_1)\,\alpha_1 + (d_2 + d_jc_2)\,\alpha_2 + \dots + (d_{j-1} + d_jc_{j-1})\,\alpha_{j-1} + d_j\,c_j\,\alpha_j$
$+ (d_{j+1} + d_jc_{j+1})\,\alpha_{j+1} + (d_n + d_jc_n)\,\alpha_n = \theta$

Since $\{\alpha_1, \alpha_2 \dots \alpha_n\}$ is linearly independent, we have

$$d_1 + d_jc_1 = 0,\ d_2 + d_jc_2 = 0,\ d_{j-1} + d_jc_{j-1} = 0,\ d_jc_j = 0$$
$$d_{j+1} + d_jc_{j+1} = 0,\ ..\ d_n + d_jc_n = 0$$

Since $d_jc_j = 0$, but $c_j \neq 0$ we havt $d_j = 0$, hence it follows that $d_1 = d_2 = \dots = d_{j-1} = d_{j+1} = \dots = d_n = \theta$

This proves that, $\{\alpha_1, \alpha_2 \dots \alpha_{j-1}, \beta, \alpha_{j+1}, .. \alpha_n\}$ is a linearly independent set of vectors.

We now prove that $L\ \{\alpha_1, \alpha_2 \dots \alpha_{j-1}, \beta, \alpha_{j+1}, .. \alpha_n\} = V$.

Let δ be any arbitray vector in V. Since $\{\alpha_1, \alpha_2 .. \alpha_n\}$ is a basic of V. there exit $k_1, k_2 .. k_n \in F$ such that

$$\delta = k_1\alpha_1 + k_2\alpha_2 + .. + k_n\alpha_n$$

or, $\delta = k_1\alpha_1 + .. + k_{j-1}\alpha_{j-1} + k_j(p_1\alpha_1 + .. + p_{j-1}\alpha_{j-1} + p_j\beta + p_{j+1}\alpha_{j+1} +.. + p_n\alpha_n) + k_{j+1}\alpha_{j+1} +.. + k_n\alpha_n$

$= s_1\alpha_1 + s_2\alpha_2 + .. + s_{j-1}d_{j-1} + s_j\beta + s_{j+1}d_{j+1} +.. + s_n\alpha_n$

where $s_i = k_i + kp_i, \quad i \neq j,$

$= kp_i \qquad i = j$

$\therefore \quad \delta \in L\{\alpha_1, \alpha_2 .. \alpha_{j-1}, \beta, \alpha_{j+1}, .. \alpha_n\}$

$\therefore \quad V \subseteq L\{\alpha_1, \alpha_2, \alpha_{j-1}, \beta, \alpha_{j+1}, \dots \alpha_n\} \quad .... (2)$

But L $\{\alpha_1, \alpha_2, .. \alpha_{j-1}, \beta, \alpha_{j+1}, \alpha_n\}$being the smallest subspace containing the set $\{\alpha_1, \alpha_2, \alpha_{j-1}, \beta, \alpha_{j+1} \dots \alpha_n\}$.

$\therefore \quad L\{\alpha_1, \alpha_2\, \alpha_{j-1}, \beta, \alpha_{j+1}, .. \alpha_n\} \subseteq V \quad .... (3)$

Combining (2) and (3),

$V = L\{\alpha_1, \alpha_2\, \alpha_{j-1}, \beta, \alpha_{j+1}, .. \alpha_n\}$

Hence $\{\alpha_1, \alpha_2\, \alpha_{j-1}, \beta, \alpha_{j+1}, .. \alpha_n\}$is a basis of V.

Dimension of a vector space

If a vector space V over a field F has a basis consisting of a finite number of elements, then the space is said to be **finite dimensional ;** the number of elements in a basis is called the **dimension** of the space and is is written as dim V. If dim $V = n$, V is said to be n-dimensional. If V is not finite dimensional, it is called **infinite dimensional.** The null vector space has no basis. It is also said to be of dimension zero.

Theorem 1.9 : If $\{\alpha_1, \alpha_2 \dots \alpha_n\}$ be a basis of a finite dimensional vector space V over a field F, then any set of lineraly independent vectors of V contains at most n vectors.

Proof : Let $\{\beta_1, \beta_2 \dots \beta_m\}$ be a linearly independent set of vectors in V. None of β_i is a zero vector. Since $\{\alpha_1, \alpha_2 \dots \alpha_n\}$ is a basis of V and β_1 is a nonzero vector in V,

$$\beta_1 = c_1\alpha_1 + c_2\alpha_2 + .. + c_n\alpha_n$$

where $c_1, c_2, .. c_n$ are scalars in F, not all of which are zero. Let $c_i \neq 0$, then by the Replacement Theorem 1·8, $\{\alpha_1, \alpha_2 .. \alpha_{i-1}, \beta_1, \alpha_{i+1}, .. \alpha_n\}$is a basis of V.

Since β_2 is a non-zero vector of V and $\{\alpha_1, \alpha_2 .. \alpha_{i-1}, \beta_1, \alpha_{i+1}, .. \alpha_n\}$is a basis of V.

$$\beta_2 = d_1\alpha_1 + .. + d_{i-1}\alpha_{i-1} + d_i\beta_1 + d_{i+1}\alpha_{i+1} + .. + d_n\alpha_n$$

where d_i's are scalars, not all zero.

It is clear that at least one of $d_1, d_2, .. d_{i-1}, d_{i+1}, \dots d_n$ is nonzero. For, if is not true, then $\beta_2 = d_i\beta_1$ which shows that β_1 and β_2 are linearly dependent, which is a contradiction to the initial assumption that $\beta_1, \beta_2 .. \beta_n$ are linearly independent.

Let $d_j \neq 0, j \neq i$

Then, by the Replacement Theorem 1.8,

$\{\alpha_1, \alpha_2 .. \alpha_{i-1}, \beta_1, \alpha_{i+1}, .. \alpha_{j-1}, \beta_2, \alpha_{j+1}, .. \alpha_n\}$ is a new basis of V.

Proceeding in this way we observe that at each step one α is replaced by one β, and a new set of basis of V is obtained.

The following possiblities occur :

(a) $\beta_1, \beta_2 .. \beta_m$ all came to the new basis containing α's. In this case we have m < n.

(b) $\beta_1, \beta_2 .. \beta_m$ exhaust all α's and form a new basis. In this case $m = n$.

The case $m > n$ is not possible. For, it m > n. then $\beta_1, \beta_2 .. \beta_n$ will replace $\alpha_1, \alpha_2 .. \alpha_n$ and form a new basis.

Then the remaining β's viz $\beta_{n+1}, \beta_{n+2}, .. \beta_m$ we be linear combination of $\beta_1, \beta_2, .. \beta_n$, which means that $\{\beta_1, \beta_2, .. \beta_m\}$ is a linearly dependent set of vectors, a contradition.

Hence $m \leq n$.

Theorem 1.10 : Any two bases of a finite dimensional vector space V have the same number of vectors.

Proof : Let $\{\alpha_1, \alpha_2 \dots \alpha_n\}$ and $\{\beta_1, \beta_2 .. \beta_m\}$ be the two bases of a finite dimensional vector space V.

Since $\{\alpha_1, \alpha_2 \dots \alpha_n\}$ is a basis if V and $\{\beta_1, \beta_2 .. \beta_m\}$ is a linearly independent set of vector, we have, by Theorem 1.9, $m \leq n$ .....(1)

Again $\{\beta_1, \beta_2 .. \beta_m\}$ is a basis of V and $\{\alpha_1, \alpha_2 \dots \alpha_n\}$ is a linearly independent set of vector, we have, $n \leq m$ .... (2)

Combining (1) and (2), $n = m$, and the theorem is proved.

Theorem 1.11 : (Extension Theorem)

A linearly independent set of vectors in a finite dimensional vector space V over a field F is either a basis of V, or can be extended to a basis of V.

Proof : Let S = $\{\alpha_1, \alpha_2 \ldots \alpha_m\}$ be a linearly independent set of vectors in V. $L(S)$ being the smallest subspace containing S, hence

$$L(S) \subseteq V.$$

If $L(S) = V$, then S is a basis of V and there is nothing to prove.

So we suppoes that $L(S) \subset V$.

Let $\quad \beta \in V - L(S)$

We shall prove that the set of vector $\{\alpha_1, \alpha_2 \ldots \alpha_m, \beta\}$ is linearly independent.

Let us suppose that $c_1, c_2, \therefore c_m$, b $\in F$ such that

$$c_1\alpha_1 + c_2\alpha_2 + .. + c_m\alpha_m + b\beta = \theta \qquad \ldots (1)$$

It follows that b = 0. For if b $\neq$ 0, then b^{-1} exsits in F, and multiplying eqn (1) by b^{-1}, we obtain.

$$\beta = -b^{-1}c_1\alpha_1 - b^{-1}c_2\alpha_2 .. - b^{-1}c_m\alpha_m$$

This shows that β is a linear combination of $\alpha_1, \alpha_2 \ldots \alpha_m$ and hence $\beta \in L(S)$, which is not true, as $\beta \in V - L(S)$.

Hence b = 0

Then from (1), $c_1\alpha_1 + c_2\alpha_2 + .. + c_m\alpha_m = \theta$

But $\alpha_1, \alpha_2 .. \alpha_m$ are linearly independent vectors, hence, $c_1 = c_2 = .. = c_m = 0$.

This shows that the set $S_1 = \{\alpha_1, \alpha_2 .. \alpha_m, \beta\}$ is a linearly independent set.

Now, if $L(S_1) = V$, then S_1 is a basis of V. If $L(S_1) \neq V$, then we suppose that $\gamma \in V - L(S_1)$. Then proceeding exactly in the same way as above, we construct a set

$S_2 = \{\alpha_1, \alpha_2 \ldots \alpha_m, \beta, \gamma\}$ of linearly independent vectors.

We check if $L(S_2) = V$, then S_2 is a basis of V; if not, we proced as above. And after a finite number of steps we can find a basis of V containing S.

EXAMPLES

Ex. 1. Prove that the set S = { (1, 0, 1), (0, 1, 1), (1, 1, 0)} is a basis of R^3.

Let $\alpha_1 = (1, 0, 1), \alpha_2 = (0, 1, 1), \alpha_3 = (1, 1, 0)$

We have to show that $\alpha_1, \alpha_2, \alpha_3$ are linearly independent and that $L(S) = R^3$.

To show that $\alpha_1, \alpha_2, \alpha_3$ are linearly independent, we suppose that there exist c_1, $c_2, c_3 \in F$ such that

$$c_1\alpha_1 + c_2\alpha_2 + c_3\alpha_3 = \theta$$

$\Rightarrow \quad c_1(1, 0, 1) + c_2(0, 1, 1) + c_3(1, 1, 0) = (0, 0, 0)$

$\Rightarrow \quad c_1 + c_3 = 0, \qquad c_2 + c_3 = 0, \qquad c_1 + c_2 = 0,$

$\Rightarrow \quad c_1 = c_2 = c_3 = 0,$

Hence $\alpha_1, \alpha_2, \alpha_3$ are linearly independent.

To show that $L(S) = R^3$, we suppose that $\xi = (a, b, c)$ be any vector in R^3.

ξ will belong to $L(S)$ if we can find $r_1, r_2, r_3 \in F$ such that

$\xi = r_1\alpha_1 + r_2\alpha_2 + r_3\alpha_3.$

This requires, $(a, b, c) = r_1(1, 0, 1) + r_2(0, 1, 1) + r_3(1, 1, 0)$

$$\Rightarrow \quad \left.\begin{aligned} a &= r_1 + r_3 \\ b &= r_2 + r_3 \\ c &= r_1 + r_2 \end{aligned}\right\} \qquad (1)$$

The system of equations (1) will have unique solution if the coefficient determinant is non zero.

Now the

$$\text{Coeff. determinant} = \begin{vmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{vmatrix} = -2 \neq 0$$

nce we can find $\gamma_1, \gamma_2, \gamma_3$ from (1) such that

$\xi = \gamma_1\alpha_1 + \gamma_2\alpha_2 + \gamma_3\alpha_3$

Hence $\xi \in L(S)$.

Since $L(S)$ is the smallest subspace of R^3 containing S. We have $L(S) = R^3$ containing S, we have $L(S) = R^3$

Hence the given set forms a basis of R^3.

Ex. 2. Find a basis and the dimension of the subspace W of R^3 where

$W = \{(x, y, z) \in R^3 / x + y + z = 0\}$

Let $\xi = (a, b, c)$ be an arbitrary vector of W.

Then from the structure of W, we have $a + b + c = 0$

or, $\quad c = -a - b$

Hence $\xi = (a, b, -a, -b) = a(1, 0, -1) + b(0, 1, -1)$

Let $\alpha = (1, 0, -1)$ and $\beta = (0, 1, -1)$

Then ξ is a linear combination of vectors α and β

$\therefore \xi \in L\{\alpha, \beta\}$

We shall show that α and β are linearly independent. Let us suppose that scalars c_1, c_2, exists in F such that

$$c_1\alpha + c_2\beta = \theta$$

$$\Rightarrow \quad c_1(1, 0, -1) + c_2(0, 1, -1) = \theta = (0, 0, 0)$$

$$\Rightarrow \quad c_1 = 0, c_2 = 0, \quad -c_1 - c_2 = 0$$

$$\therefore c_1 = 0, c_2 = 0$$

Hence α, β are linearly independent

$\therefore \{\alpha, \beta\}$ is a basis of W and dim $W = 2$.

Theorem 1.12 : Let U and W be two subspaces of a finite dimensional vector space V over a field F. Then dim (U + W) = dim U + dim W – dim (U $\cap$ W)

Proof : Since U, W are subspaces of a finite dimensional vector space V, dim U. dim W, dim, (U + W), dim (U $\cap$ W) are each finite.

Let $S = \{\alpha_1, \alpha_2 \ldots \alpha_r\}$ be a basis for U $\cap$ W and let S be supplemented by additional vectors $\beta_1, \beta_2 .. \beta_s$ to make up a basis for U and similarly let S be supplemented by additional vectors $\gamma_1, \gamma_2, \gamma_t$ to make up a basis for W.

We want to prove that

$S_1 = \{\alpha_1, \alpha_2 .. \alpha_r, \beta_1, \beta_2 .. \beta_s, \gamma_1, \gamma_2, .. \gamma_t\}$ is a basis for U + W.

Clearly, $L(S_1) \subseteq U + W$

Let $\xi \in U + W$, then $\qquad \xi = u + w$, where $u \in U$ and $w \in W$.

Since $u \in U$, and $\{\alpha_1, \alpha_2 .. \alpha_r, \beta_1, \beta_2 . \beta_s\}$ is a basis of U. we have,

$$u \in L\{\alpha_1, \alpha_2 \alpha_r, \beta_1, \beta_2 \beta_s\}$$

and similarly, $w \in L\{\alpha_1, \alpha_2 \alpha_r, \gamma_1, \gamma_2, .. \gamma_t\}$

Hence $\qquad U + W \subseteq L(S_1)$

Thus, $\qquad L(S_1) = U + W$

To show that, $\{\alpha_1, \alpha_2 .. \alpha_r, \beta_1, \beta_2 .. \beta_s, \gamma_1, \gamma_2, \gamma_t\}$ is a basos of U + W we have to establish linear independence of the set of vectors S_1.

Suppose that for some scalars $a_1, a_2 .. a_r, b_1, b_2 .. b_s, c_1, c_2 .. c_t \in F$

$$a_1\alpha_1 + a_2\alpha_2 + .. + a_r\alpha_r, + b_1\beta_1 + .. + b_s\beta_s + c_1\gamma_1 + .. + c_t\gamma_t \in \theta \qquad \text{(i)}$$

Then $a_1\alpha_1 + a_2\alpha_2 + .. + a_r\alpha_r, + b_1\beta_1 + .. + b_s\beta_s$

$$= -(c_1\gamma_1 + .. + c_t\gamma_t) \qquad \text{(ii)}$$

The left hand vector in (ii) belongs to U, right hand vector belongs to W, but as these are equal, if follows that $-(c_1\gamma_1 + .. + c_t\gamma_t)$ belongs to $U \cap W$.

But $\{\alpha_1, \alpha_2 .. \alpha_r\}$ is a basis of $U \cap W$, hence there exist scalars $d_1, d_2 .. d_r \in F$ such that

$$-(c_1\gamma_1 + c_2\gamma_2 + .. + c_t\gamma_t) = d_1\alpha_1 + d_2\alpha_2 + .. + d_r\alpha_r$$

or, $\quad d_1\alpha_1 + d_2\alpha_2 + .. + d_r\alpha_r + c_1\gamma_1 + c_2\gamma_2 + .. + c_t\gamma_t) = \theta$

But $\quad \{\alpha_1, \alpha_2 .. \alpha_r, \gamma_1, \gamma_2 .. \gamma_t\}$is a basis for W,

hence $\quad d_1 = d_2 = .. = d_r = c_1 = c_2 = .. = c_t = \theta$

Hence from (ii) we have

$$a_1\alpha_1 + a_2\alpha_2 + .. + a_r\alpha_r + b_1\beta_1 + .. + b_s\beta_s) = \theta$$

But $\{\alpha_1, \alpha_2 .. \alpha_r, \beta_1, \beta_2 .. \beta_s\}$is a basis of U.

so, $\quad a_1 = a_2 = .. = a_r = b_1 = b_2 = .. = b_s = 0$

Since all the scalars a's, b's, c's are zero in (i), it follows that $\{\alpha_1, \alpha_2\ \alpha_r, \beta_1, \beta_2 .. \beta_s, \gamma_1, \gamma_2, \gamma_t\}$ is an independent set of vectors and hence, is a basis for U + W.

$$\therefore \quad \dim(U + W) = r + s + t = (r + s) + (r + t) - r$$
$$= \dim U + \dim W - \dim(U \cap W)$$

Ex. 3. Suppose U and W are the xy plane and yz plane respectively, in R^3.

$U = \{(a, b, o)\}$, $\qquad W = \{(0, b, c)\}$

Since $\quad R^3 = U + W. \qquad \dim(U + W) = 3$

Also dim U = 2 and dim W = 2

We have

$$\dim(U + W) = \dim u + \dim w - \dim(U \cap W)$$

$$\Rightarrow \quad 3 = 2 + 2 - \dim(U \cap W)$$

$$\therefore \quad \dim(U \cap W) = 1$$

Ex. 4. Let $\{(1, 1, 1, 1), (1, 2, 1, 2)\}$ be a linearly independent subset of a vector space V_4 of dimension 4 over a field F. Extend it to a basis for V_4.

Let $S = (1, 1, 1, 1), (1, 2, 1, 2)\}$

Then $L(S) = \{c_1 (1, 1, 1, 1) + c_2(1, 2, 1, 2) / c_1, c_2 \in F\}$

$= \{(c_1 + c_2, c_1 + 2c_2; c_1 + c_2, c_1 + 2c_2)/c_1, c_2 \in F\}$

We observe that is L(S), the first and the 3rd coodinates are equal. It follows that the vector (0, 3, 2, 3) is not in L(S)

Thus we have an enlarged independent set

$S_1 = \{(1, 1, 1, 1), (1, 2, 1, 2), (0, 3, 2, 3)\}$

$L(S_1) = \{d_1 (1, 1, 1, 1) + d_2(1, 2, 1, 2) + d_3(0, 3, 2, 3) \backslash d_1, d_2, d_3 \in F\}$

$=\{d_1 + d_2, d_1 + 2d_2 + 3d_3, d_1 + d_2 + 2d_3, d_1 + 2d_2 + 3d_3/d_1, d_2, d_3 \in F\}$

If follows from the structure of $L(S_1)$ that the vector (2, 6, 4, 5) is not in $L(S_1)$. Hence the vectors (1, 1, 1, 1), (1, 2, 1, 2), (0, 3, 2, 3), (2, 6, 4, 5) are linearly independent and since V_4 is of dimension 4, they form a basis of V_4.

EXERCISES

1. Check whether the following set of vectors is linearly dependent or linearly independent

   $\{(1, 0, 1), (1, 1, 0), (1, -1, 1), (1, 2, -3)\}$

2. Let $V = R^3$. Show that W is a subspace of V, where :

   (i) $W = \{(a, b, 0) ; a, b \in R\}$

   (ii) $W = \{(a, b, c) ; a + b + c = 0\}$

3. Let $V = R^3$. Show that W is not a subspace of V. where :

   (i) $W = \{(a, b, c) : a \geq 0\}$

   (ii) $W = \{(a, b, c) : a^2 + b^2 + c^2 \leq 1\}$

4. Write the matrix $E = \begin{pmatrix} 3 & 1 \\ 1 & -1 \end{pmatrix}$ as a linear combination of the matrices

   $A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}$ and $C = \begin{pmatrix} 0 & 2 \\ 0 & -1 \end{pmatrix}$.

5. If $\{\alpha_1, \alpha_2, \alpha_3\}$ be a basias of real vector space V and $\beta_1 = \alpha_1 + \alpha_3$, $\beta_2 = 2\alpha_1 + 3\alpha_2 + 4\alpha_3$, $\beta_3 = \alpha_1 + 2\alpha_2 + 3\alpha_3$, prove that $\{\beta_1, \beta_2 \beta_3\}$is also a basis of V.

6. Find a basis for the vector space R^3 that contains the vectors (1, 2, 1) and (3, 6, 2)

7. Find the dimension of the subspace S of R^3 defined by

   (i) $S = \{(x, y, z) \in R^3 : 2x + y - z = 0\}$

   (ii) $S = \left\{(x, y, z) \in R^3 : \begin{matrix} x+2y=z \\ 2x+3z=y \end{matrix}\right\}$

8. Prove the set $\{(1, 1, 1), (1, -1, 1), (0, 1, 1)\}$ is a basis for R^3.

9. Given S as a finite subsef of a vector space V, prove that if S is linearly independent and every proper superset of S in V is linearly dependent then S is a basis for V.

Unit : 2 □ Inner Product Spaces

A Vector space involves a set of vectors and an associated field. If we restrict this field to be either the field of complex numbers or the field of real numbers, and if we define a certain mapping known as the **inner product function** with respect to the given vector space then we create still another important algebraic system known as an inner product space. Such a special vector space is very useful because within this structue we can define abstractly such common concepts as 'length', 'distance' and 'angle'.

One example of an inner product is the scalar or dot product of vectors in R^3. The scalar product of the vectors

$$\alpha = (x_1, x_2, x_3) \text{ and } \beta = (y_1, y_2, y_3)$$

in R^3 is the real number.

$$(\alpha, \beta) = x_1y_1 + x_2y_2 + x_3y_3$$

Geometrically, this dot product is the product of the length of α, the length of β and the cosine of the angle between α and β. It is therefore possible to define the geometric concepts of length and angle in R^3 by means of the algebraically defined scalar product.

An inner product on a vector space is a function with properties similar to the dot product in R^3, in terms of such an inner product we can also define 'length' and 'angle'. Our comments about the general notion of angle will be restricted to the concept of perpendicularity (or orthogonality) of vectors.

We begin with a definition.

Definition : Let F be the field of real numbers or the field of complex numbers, and V be a vector space over F. An inner product on V is a function which assigns to each ordered pair of vectors α, β in V a scalar (α, β) in F in such a way that for all α, β, γ in V and all scalars $c \in F$.

(a) $(\alpha + \beta, \gamma) = (\alpha, \gamma) + (\beta, \gamma)$

(b) $(c\alpha, \beta) = c(\alpha, \beta)$

(c) $(\beta, \alpha) = (\alpha, \beta), \overline{(\alpha, \beta)}$ the overhead bar denotes complex conjugation

(d) $(\alpha, \alpha) \geq 0$, $(\alpha, \alpha) = 0$ if and only if $\alpha = 0$.

It should be observed that conditions (a), (b) and (c) imply that

(e) $(\alpha, c\beta + \gamma) = c\ \bar{c}(\alpha, \beta) + ((\alpha, \gamma)$

EXAMPLES

Ex. 1. On F^n, there is an inner product which we shall call the standard inner product. It is defined on $\alpha = (x_1, x_2 .. x_n)$ and $\beta = (y_1, y_2 .. y_n)$ by

$$(\alpha, \beta) = \sum_i x_i \overline{y_i}$$

When $F = R$, this may also be written as

$$(\alpha, \beta) = \sum_i x_i y_i$$

In the real case, the standard inner product is often called the dot or scalar product and is denoted by $\alpha.\ \beta$.

We note here from the definition that (α, α) is always real. We use the notation.

$$\|\alpha\| = \sqrt{(\alpha, \alpha)}$$

This non-negative real number $\|\alpha\|$ is called the norm or length of α.

A real vector space V for which a real inner product is defined in called a **Euclidean space.**

A complex vector space V for which a complex inner product is defined is called a **Unitary space.**

Enclidean spaces and unitary spaces are collectively called **inner product space.**

Ex. 2. Let V be the vector space of all real valued continuous functions of $[a, b]$ and let $F = R$.

For $x, y \in V$, let us define

$$(x, y) = \int_a^b x(t)y(t)dt \qquad ...(1)$$

Then conditions (a), (b) and (c) of the inner product are easy to verify. Also

$$(\mathrm{x}, \mathrm{x}) = \int_a^b (x(t))^2\, dt \quad \geq 0$$

If equality holds, then using the fact that $x(t)$ is continuous it can be proved that $x(t) = 0$ for all t. Equation (1) defines an niner product and the vector space V is an inner product space.

Ex. 3. Let C be the field of complex numbers. Let $\mathrm{x} = (x_1, x_2, .. x_n)$ and $y = (y_1, y_2, .. y_n)$ be two elements of the vectors space C^n.

We define

$$(x, y) = x_1\overline{y}_1 + x_2\overline{y}_2 + .. + x_n\overline{y}_n$$

This inner product on C^n defines a unitany space.

Ex. 4. Let V be the vector space of infinte sequences of real numbers $\{a_n\}$, satisfying.

$$\sum_{n=1}^{\infty} a_n^{\ 2} < \propto$$

i.e. the sum converges.

Addition and multiplication are defined componentwise :

$$(a_1, a_2 \ldots) + (b_1, b_2 \ldots) = (a_1 + b_1, a_2 + b_2, \ldots)$$

$$\text{and } k(a_1, a_2 \ldots) = (ka_1, ka_2 \ldots)$$

An inner product is defnied in V by

$$((a_1, a_2 \ldots), (b_1, b_2 \ldots) = (a_1b_1 + a_2b_2 + ..$$

The above sum converges absolutely for any pair of points in V, hence the inner product is well defined. This inner product space is called l_2 space (or Hilbert space).

Problem : Show that in any inner product space V,

$$(0, x) = 0, (x, 0) = 0, x \in V.$$

Solution : We have, by property (a) of inner product

$$(0 + 0, x) = (0, x) + (0, x)$$

$$\Rightarrow \quad (0, x) = (0, x) + (0, x)$$

$$\Rightarrow \quad (0, x) = 0$$

Also, $(x, 0) = \overline{(0, x)} = \overline{0} = 0$.

We have already defined the norm or length of *a* vector α in an inner product space by.

$$\| \alpha \| = \sqrt{(\alpha, \alpha)}$$

Theorem : The norm has the following properties :

(a) $\|c\alpha\| = |c|\,\|\alpha\|, \alpha \in V, c \in F$

(b) $\|\alpha\| > 0$ if $\alpha \neq 0$ and $\|\alpha\| = 0$ if $\alpha = 0$.

(c) $\|\alpha + \beta\| \leq \|\alpha\| + \|\beta\|$, $\alpha, \beta \in V$

Proof : $\|\alpha\| \geq 0$ follows from the definition of inner product. We have

$$(c\alpha, c\alpha) = c(\alpha, c\alpha) = \bar{c}\,c\,(\alpha, \alpha)$$

$$\therefore \quad \|c\alpha\|^2 = |c|^2 \|\alpha\|^2$$

$\sqcup$ $\|c\alpha\| > 0$, We have, $\|c\alpha\| = |c|\,\|\alpha\|$, which proves, part (*a*)

To prove part (*c*), we first prove the **Schwarz inequality**.

$$|(\alpha, \beta)| \leq \|\alpha\|\,\|\beta\| \qquad \text{.....} (1)$$

Clearly, the inequality i (1) is true when α = 0.

If α ≠ 0, let us write

$$\gamma = \beta - \frac{(\beta, \alpha)}{\|\alpha\|^2}\alpha$$

$$\text{Then} \quad (\gamma, \alpha) = \left(\beta - \frac{(\beta, \alpha)}{\|\alpha\|^2}\alpha, \alpha\right)$$

$$= (\beta, \alpha) - \frac{(\beta, \alpha)(\alpha, \alpha)}{\|\alpha\|^2}$$

$$= (\beta, \alpha) - \frac{(\beta, \alpha)\|\alpha\|^2}{\|\alpha\|^2} = 0$$

Again $\quad \|\gamma\|^2 \geq 0$

$$\Rightarrow \quad |(\gamma, \gamma)| \geq 0$$

$$\Rightarrow \quad \left(\beta - \frac{(\beta, \alpha)}{\|\alpha\|^2}\alpha, \beta - \frac{(\beta, \alpha)}{\|\alpha\|^2}\alpha\right) \geq 0$$

$$\Rightarrow \qquad \left(\beta-\frac{(\beta,\alpha)}{\|\alpha\|^2}\alpha,\beta\right)-\left(\beta-\frac{(\beta,\alpha)}{\|\alpha\|^2}\alpha,\frac{(\beta,\alpha)}{\|\alpha\|^2}\alpha\right)\geq 0$$

$$\Rightarrow \qquad (\beta,\beta)-\frac{(\beta,\alpha)(\alpha,\beta)}{\|\alpha\|^2}-\frac{\overline{(\beta,\alpha)}}{\|\alpha\|^2}(\beta,\alpha)+\frac{(\beta,\alpha)\overline{(\beta,\alpha)}}{\|\alpha\|^2\|\alpha\|^2}(\alpha,\alpha)\geq 0$$

Since $(\alpha,\beta)=\overline{(\beta,\alpha)}$, we have

$$\|\beta\|^2-\frac{|(\beta,\alpha)|^2}{\|\alpha\|^2}-\frac{|(\beta,\alpha)|^2}{\|\alpha\|^2}+\frac{|(\beta,\alpha)|^2}{\|\alpha\|^2}\geq 0$$

$$\Rightarrow \qquad \|\beta\|^2\geq\frac{|(\beta,\alpha)|^2}{\|\alpha\|^2}$$

Hence $\quad |(\alpha,\beta)|^2=|(\beta,\alpha)|^2\leq\|\alpha\|^2\|\beta\|^2$

Now by using (1) we have

$$\begin{aligned}\|\alpha,\beta\|^2=(\alpha+\beta,\alpha+\beta) &= (\alpha,\alpha)+(\alpha,\beta)+(\beta,\alpha)+(\beta,\beta)\\ &= \|\alpha\|^2+(\alpha,\beta)+\overline{(\alpha,\beta)}+\|\beta\|^2\\ &= \|\alpha\|^2+2\mathrm{Re}(\alpha,\beta)+\|\beta\|^2\end{aligned}$$

Re (α,β) is the real partof (α,β)

$$\begin{aligned}&\leq\|\alpha\|^2+2\|\alpha\|\|\beta\|+\|\beta\|^2\\ &=(\|\alpha\|+\|\beta\|)^2\end{aligned}$$

Hence $\quad \|\alpha+\beta\|\leq\|\alpha\|+\|\beta\|$

Definition : Let α and β be vectors in an inner product space *V*. Then α is orthogonal to β if $(\alpha,\beta)=0$; since this implies β is **orthogonal to** α we often simply say that α and β are orthogonal.

If, *S* is a set of vectors in *V*, *S* is called an **orthogonal set** provided all pairs of distinct vectors in *S* are orthogonal.

An **orthogonal set** is an orthogonal set *S* with the additional property that $\|\alpha\|=1$ for every α in *S*.

The zero vector is orthogonal to every vector in *V* and is the only vector with this property. It is appropriate to think of an orthonormai set as a set of mutually perpendicular vectors, each having length 1.

Ex. 5. The vector (x, y) in R^2 is orthogonal to $(-y, x)$ with respect to standard inner product, for

let $\alpha = (x, y)$, $\beta = (-y, x)$
so that $(\alpha, \beta) = x(-y) + yx = 0$

Problem : For the complex inner product space C^3 find $\|\alpha\|$ and $\|\alpha - \beta\|$ where $\alpha = (i, 2, 0)$ and $\beta = (0, -i, 6)$.
Are α and β orthogonal?

Solution : We have $\|\alpha\|^2 = (\alpha, \alpha)$

$$= i\bar{i} + 2.\bar{2} + 0.\bar{0}$$

$$= \text{i}(-\text{i}) + 2.2 + 0$$

$$= 5$$

$$\therefore \ \|\alpha\| = \sqrt{5}$$

$$\|\alpha - \beta\|^2 = (\alpha - \beta, \alpha - \beta) \ = (\alpha, \alpha) - (\alpha, \beta) + (\beta, \beta)$$

$$= \|\alpha\|^2 - 2\text{Re}(\beta, \alpha) + \|\beta\|^2$$

Now, $\|\alpha\|^2 = 5$, $\|\beta\|^2 = 0^2 - \text{i}^2 + 6^2 = 37$

$$(\beta, \alpha) = 0.\bar{\text{i}} + (-1)\bar{2} + 6.\bar{0} = -2\text{i} = 0 - 2\text{i}$$

$\therefore$ $\quad$ $\text{Re}\,(\beta, \alpha) = 0$

$\therefore$ $\quad$ $\|\alpha - \beta\|^2 = 5 - 2.0 + 37 = 42$

$$\therefore \ \|\alpha - \beta\| = \sqrt{42}$$

Since $(\beta, \alpha) \neq 0$, α, β are not orthogonal.

Theorem : (Pythagoras), If α, β be two orthogonal vectors in a Euclidean space V, then

$$\|\alpha + \beta\|^2 = \|\alpha\|^2 + \|\beta\|^2$$

Proof : We have, $\|\alpha + \beta\|^2 = (\alpha + \beta, \alpha + \beta) = \|\alpha\|^2 + (\alpha, \beta) + (\beta, \alpha) + \|\beta\|^2$

$$= \|\alpha\|^2 - 2\text{Re}(\alpha, \beta) + \|\beta\|^2$$

$$= \|\alpha\|^2 + 2(\alpha, \beta) + \|\beta\|^2,$$

(since in a Euclidean space $(\alpha, \beta$ is real)

Since the vectors α, β are orthogonal $(\alpha, \beta) = 0$

$$\therefore \|\alpha+\beta\|^2 = \|\alpha\|^2 + \|\beta\|^2 .$$

Theorem : (Parallelogram law). If α, β be any two vectors in Euclidean space V then ,

$$\|\alpha+\beta\|^2 + \|\alpha-\beta\|^2 = 2\|\alpha\|^2 + 2\|\beta\|^2$$

Proof : We have $\|\alpha+\beta\|^2 = \|\alpha\|^2 + 2(\alpha,\beta) + \|\beta\|^2$

and $\|\alpha-\beta\|^2 = \|\alpha\|^2 - 2(\alpha,\beta) + \|\beta\|^2$

Adding $\|\alpha+\beta\|^2 + \|\alpha-\beta\|^2 = 2\|\alpha\|^2 + 2\|\beta\|^2$

Theorem : An orthogonal set of non-zero vectors is linearly independent.

Proof : Let S be a finite or infinite orthogonal set of non-zero vectors in a given inner product space. Suppose $\alpha_1, \alpha_2 .. \alpha_m$ are distinct vectors in S and that

$$\beta = c_1\alpha_1 + c_2\alpha_2 + .. + c_m\alpha_m \qquad ..... (1)$$

where $c_1, c_2 .. c_m$ are scalars from the field on which the vector space is defined.

Then $(\beta, \alpha_k) = (c_1\alpha_1 + c_2\alpha_2 + .. + c_m\alpha_m, \alpha_k)$

$$= c_1(\alpha_1, \alpha_k) + c_2(\alpha_2, \alpha_k) + .. + c_k(\alpha_k, \alpha_k) + .. + c_m(\alpha_m, \alpha_k)$$

Since $\alpha_1, \alpha_2 .. \alpha_m$ are pairwise orthogonal,

$$(\alpha_i, \alpha_k) = 0 \qquad \text{for } i \neq k$$

$$\therefore \qquad (\beta, \alpha_k) = c_k(\alpha_k, \alpha_k)$$

Since $(\alpha_k, \alpha_k) \neq 0$. it follows that

$$c_k = \frac{(\beta,\alpha_k)}{(\alpha_k,\alpha_k)} = \frac{(\beta,\alpha_k)}{\|\alpha_k\|^2}, 1 \leq k \leq m \qquad .... (2)$$

Thus, when $\beta = 0$, $\quad c_k = \dfrac{(\beta,\alpha_k)}{\|\alpha_k\|^2} = 0 \qquad k = 1, 2, \ldots m.$

Hence from (1) we conclude that the vectors $\alpha_1, \alpha_2 .. \alpha_m$ are linearl independent.

Corollary : If a vector β is a linear combination of a orthogonal sequence of non-zero vectors $\alpha_1, \alpha_2 .. \alpha_m$, then β is the particular linear combination.

$$\beta = \sum_{k=1}^{m} \frac{(\beta, \alpha_k)}{\|\alpha_k\|^2} = \alpha_k$$

The corollary follows simply from equations (1) and (2) in the proof of the theorem.

Gram-Schmidt Orthogonalization :

The basis $\{(1, 0), (0, 1)\}$ of the vector space of all two-dimensional coordinate vectors, and the basis $\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ of the vector space of all three dimensional coordinate vectors, have the property that the standard inner product of every pair of vectors within each basis is zero ; because of this property these two bases are illustrations of orthogonal bases. In both of these examples, a geometric interpretation is that for each orthogonal pair of vectors the corresponding geometric vectors are perpendiculars to each other.

Theorem : Let V be an inner product space and let $\beta_1, \beta_2, \ldots \beta_n$ be independent vectors in V. Then we can construct orthogonal vectors $\alpha_1, \alpha_2 .. \alpha_n$ in V such that for each $k = 1, 2, .. n$, the set

$$\{\alpha_1, \alpha_2 .. \alpha_k\}$$

is a basis for the subspace spanned by $\beta_1, \beta_2, \ldots \beta_k$.

Proof : The vectors $\alpha_1, \alpha_2 .. \alpha_n$ will be obtained by means of a construction known as **Gram-Schmidt orthogonalization process.**

Proof : We shall use induction to prove the theorem. First let $\alpha_1 = \beta_1$.

As induction hypothesis, suppose $\alpha_1, \alpha_2 .. \alpha_m$ $(1 \leq m \leq n)$ have been chosen so that for every k.

$$\{\alpha_1, \alpha_2 .. \alpha_k\}, \ 1 \leq k \leq m$$

satisfy the statement of the theorem. i. e. $\{\alpha_1, \alpha_2 .. \alpha_k\}$, $1 \leq k \leq m$, is an orthogonal basis for the subspace of V, spanned by $\beta_1, \beta_2, .. \beta_k$.

On the basis of this hypothesis, we shall show that a vector α_{m+1} can be constructed such that

$$\{\alpha_1, \alpha_2 .. \alpha_{m+1}\}$$

is a set of orthogonal vectors satisfying the statement of the theorem.

Let us write $\alpha_{m+1} = \beta_{m+1} - \sum_{k=1}^{m} \frac{(\beta_{m+1}, \alpha_k)}{\|\alpha_k\|^2} = \alpha_k$

Then clearly $\alpha_{m+1} \neq 0$. For, it $\alpha_{m+1} = 0$, then

$$\beta_{m+1} = \sum_{k=1}^{m} \frac{(\beta_{m+1}, \alpha_k)}{\|\alpha_k\|^2} = \alpha_k$$

In otherwords, β_{m+1} is a linear combination of the vectors $\alpha_1, \alpha_2 .. \alpha_m$ and hence a linear combination of $\beta_1, \beta_2, .. \beta_m$, which is not true, as by assumption $\beta_1, \beta_2, .. \beta_n$ are linearly independent.

Hence $\alpha_{m+1} \neq 0$.

Again, for $\mathrm{i} \leq j \leq m$, we have

$$\begin{aligned}(\alpha_{m+1}, \alpha_j) &= (\beta_{m+1}, \alpha_j) - \sum_{k=1}^{m} \frac{(\beta_{m+1}, \alpha_j)(\alpha_k, \alpha_j)}{\|\alpha_k\|^2} \\ &= (\alpha_{m+1}, \alpha_j) - \frac{(\beta_{m+1}, \alpha_j)(\alpha_j, \alpha_j)}{\|\alpha_j\|^2} \quad (\text{since } (\alpha_k, \alpha_j) = 0 \text{ for } k \neq j) \\ &= (\beta_{m+1}, \alpha_j) - (\beta_{m+1}, \alpha_j) = 0\end{aligned}$$

This shows that $\{\alpha_1, \alpha_2 .. \alpha_{m+1}\}$ is as orthogonal set consisting of m + 1 non-zero vectors in the subspace spanned by $\beta_1, \beta_2 .. \beta_{m+1}$. Since an orthogonal set of nonzero vectors is linerly independent. $\{\alpha_1, \alpha_2 .. \alpha_{m+1}\}$ is a basis for the subspace spanned by $\beta_1, \beta_2 .. \beta_{m+1}$. Hence by induction it follows that the vectors $\alpha_1, \alpha_2 .. \alpha_n$ may be constructed one after the other in accordance with the formula (1).

In particular, when $n = 4$, we have.

$$\alpha_1 = \beta_1.$$

$$\alpha_2 = \beta_2 - \frac{(\beta_2, \alpha_1)}{\|\alpha_1\|^2}\alpha_1$$

$$\alpha_3 = \beta_3 - \frac{(\beta_3, \alpha_1)}{\|\alpha_1\|^2}\alpha_1 - \frac{(\beta_3, \alpha_2)}{\|\alpha_2\|^2}\alpha_2$$

$$\alpha_4 = \beta_4 - \frac{(\beta_4, \alpha_1)}{\|\alpha_1\|^2}\alpha_1 - \frac{(\beta_4, \alpha_2)}{\|\alpha_2\|^2}\alpha_2 - \frac{(\beta_4, \alpha_3)}{\|\alpha_3\|^2}\alpha_3$$

Corollary : Every finite dimensional inner product space has the orthonormal basis.

Proof : Let V be a finite dimensional inner product space and $\{\beta_1, \beta_2 .. \beta_n\}$ be a basis of V. Applying the Gram-Schmidt process of orthogonalization, we can construct an orthogonal basis $\{\alpha_1, \alpha_2 .. \alpha_n\}$ for V. Then to obtain an orthonormal basis, we simply replace each vector α_k by ${}^{\alpha_k}\!/\!{}_{\|\alpha_k\|}$

Theorem : Any orthonormal subset of a finite dimensional inner product space V can be extended to an orthonormal basis of V.

Proof : Let $\alpha_1, \alpha_2 .. \alpha_m$ form an orthonormal subset of V. We append the vectors in some basis of V to form $\alpha_1, \alpha_2 .. \alpha_n$ and call the new vectors $\alpha_{m+1}, \alpha_{m+2} .. \alpha_n$. By Gram-Schmidt orthogonalization process applied to $\alpha_1, \alpha_2 .. \alpha_n$ we get $\beta_1, \beta_2 .. \beta_n$.

Clearly the non-null vector among $\beta_1, \beta_2 .. \beta_n$ form an orthonormal basis of V. This basis contains $\alpha_1, \alpha_2 .. \alpha_n$.

Ex. 6 : Consider the vectors $\beta_1 = (3, 0, 4)$, $\beta_2 = (-1, 0, 7)$ and $\beta_3 = (2, 9, 11)$ in R^3 equipped with the standard inner product. Applying the Gram-Schmidt process to $\beta_1, \beta_2 \beta_3$ we obtain the following vectors.

$$\alpha_1 = \beta_1 = (3, 0, 4)$$

$$\alpha_2 = \beta_2 - \frac{(\beta_2, \alpha_1)}{\|\alpha_1\|^2}\alpha_1$$

$$= (-1, 0, 7) - \frac{((-1,0,7),(3,0,4))}{\|(3,0,4)\|^2}(3,0,4)$$

$$= (-1, 0, 7) - (3, 0, 4)$$

$$= (-4, 0, 3)$$

$$\alpha_3 = \beta_3 - \frac{(\beta_3, \alpha_1)}{\|\alpha_1\|^2}\alpha_1 - \frac{(\beta_3, \alpha_2)}{\|\alpha_2\|^2}\alpha_2$$

$$= (2, 9, 11) - \frac{((2,9,11),(3,0,4))}{\|(3,0,4)\|^2}(3,0,4)$$

$$- \frac{((2,9,11),(-4,0,3))}{\|(-4,0,3)\|^2}(-4,0,3)$$

$$= (2, 9, 11) -2(3, 0, 4) - (-4, 0, 3)$$

$$= (0, 9, 0)$$

These vectors are evidently noe-zero and mutually orthogonal. Hence $\{\alpha_1, \alpha_2 .. \alpha_3\}$ is an orthogonal basis for R^3.

Clearly an orthonormal basis is $\left\{\frac{\alpha_1}{\|\alpha_1\|}, \frac{\alpha_2}{\|\alpha_2\|}, \frac{\alpha_3}{\|\alpha_3\|}\right\}$

$$= \left\{\left(\frac{3}{5},0,\frac{4}{5}\right),\left(-\frac{4}{5},0,\frac{3}{5}\right),(0,1,0)\right\}$$

Now, an arbitrary vector $\gamma = (x_1, x_2, x_3)$ in R^3 can be expressed as a linear combination of α_1, α_2 α_3 as

$$\gamma \quad = \sum_{k=1}^{3} \frac{\gamma, \alpha_k}{\|\alpha_k\|^2} \alpha_k$$

So, (x_1, x_2, x_3) $$= \frac{((x_1,x_2,x_3),(3,0,4))}{25}\alpha_1 + \frac{((x_1,x_2,x_3),(-4,0,3))}{25}\alpha_2$$

$$+ \frac{((x_1,x_2,x_3),(0,9,0))}{81}\alpha_3$$

$$= \frac{3x_1+4x_3}{25}\alpha_1 + \frac{-4x_1+3x_3}{25}\alpha_2 + \frac{x_2}{9}\alpha_3$$

In particular if $(x_1, x_2, x_3) = (1, 2, 3)$, then

$$(1,2,3) = \frac{3}{5}(3,0,4) + \frac{1}{5}(-4,0,3) + \frac{2}{9}(0,9,0)$$

Ex. 7. Extend $\{(2, 3, -1), (1, -2, -4)\}$ to an orthogonal basis of the Euclidean space R^3 and then find an associated orthonormal basis.

Let $\alpha_1 = (2, 3, -1)$ and $\alpha_2 = (1, -2, -4)$

We have $(\alpha_1, \alpha_2) \quad = 2.1 + 3.(-2) + (-1).(-4)$

$= 2 - 6 + 4 = 0$

So, α_1, α_2 are orthogonal vectors

Let $\alpha_3 = (0, 0, 1)$

Then $\{\alpha_1, \alpha_2\ \alpha_3\}$ is linearly independent, because

$$\begin{vmatrix} 2 & 3 & -1 \\ 1 & -2 & -4 \\ 0 & 0 & 1 \end{vmatrix} = -7 \neq 0$$

So $\{\alpha_1, \alpha_2\ \alpha_3\}$ is a basis of R^3.

Let $\beta = \alpha_3 - c_1\alpha_1 - c_2\alpha_2$.

where $\quad c_1 = \dfrac{(\alpha_3, \alpha_1)}{\|\alpha_1\|^2} \qquad c_2 = \dfrac{(\alpha_3, \alpha_2)}{\|\alpha_2\|^2}$

Since $(\beta, \alpha_1) = 0$, $(\beta, \alpha_2) = 0$ β is orthogonal to α_1 and α_2 and $L\{\alpha_1, \alpha_2, \alpha_3\} = L\{\alpha_1, \alpha_2, \beta\}$

$\therefore \{\alpha_1, \alpha_2, \beta\}$ is an orthogonal basis of R^3.

Now, $\quad c_1 = \dfrac{(\alpha_3, \alpha_1)}{\|\alpha_1\|^2} = \dfrac{-1}{14}, \quad c_2 = \dfrac{(\alpha_3, \alpha_2)}{\|\alpha_2\|^2} = \dfrac{-1}{21}$

$$\therefore \beta = (0,0,1) + \frac{1}{14}(2,3,-1) + \frac{4}{21}(1,-2,-4) = \left(\frac{1}{3}, \frac{1}{6}, \frac{1}{6}\right)$$

There fore, the orthogonal basis is $\left\{(2,3,-1), (1,-2,-4), \left(\frac{1}{3}, \frac{1}{6}, \frac{1}{6}\right)\right\}$ and the associated orthonormal basis is $\left\{\frac{1}{\sqrt{14}}(2,3,-1), \frac{1}{\sqrt{21}}(1,-2,-4), \frac{1}{\sqrt{6}}(2,-1,1)\right\}$

Ex . 8. Use Gram-Schmidt process to obtain an orthonormal basis of the subspace of the Euclidean space R^4 with standard inner product generated by the linearly independent set $\{(1, 1, 0, 1), (1, 1, 0, 0), (0, 1, 0, 1)\}$.

Let $\alpha_1 = (1, 1, 0, 1)$, $\alpha_2 = (1, 1, 0, 0)$ $\alpha_3 = (0, 1, 0, 1)$

We set $\beta_1 = \alpha_1 = (1, 1, 0, 1)$

and $\beta_2 = \alpha_2 - \dfrac{(\alpha_2, \beta_1)}{\|\beta_1\|^2}\beta_1$

$$= (1,1,0,0) - \frac{((1,1,0,0),(1,1,0,1))}{3}(1,1,0,1)$$

$$= (1,1,0,0) - \frac{2}{3}(1,1,0,1)$$

$$= \left(\frac{1}{3}, \frac{1}{3}, 0, \frac{-2}{3}\right) = \frac{1}{3}(1,1,0,-2)$$

$$\beta_2 = \alpha_3 - \frac{(\alpha_3, \beta_1)}{\|\beta_1\|^2}\beta_1 - \frac{(\alpha_3, \beta_2)}{\|\beta_2\|^2}\beta_2$$

$$= (0,1,0,1) - \frac{2}{3}(1,1,0,1) + \frac{1}{2}\cdot\frac{1}{3}(1,1,0,-2)$$

$$= (0,1,0,1) - \frac{2}{3}(1,1,0,1) + \frac{1}{6}(1,1,0,-2)$$

$$= \frac{1}{2}(-1,1,0,0)$$

So the orthogonal basis of the subspace is

$$\left\{(1,1,0,1), \frac{1}{3}(1,1,0,-2), \frac{1}{2}(-1,1,0,0)\right\}$$

and the corresponding orthonormal basis is

$$\left\{\frac{1}{\sqrt{3}}(1,1,0,1), \frac{1}{\sqrt{6}}(1,1,0,-2), \frac{1}{\sqrt{2}}(-1,1,0,0)\right\}$$

EXERCISES

1. Prove that the set of vectors $\{(1, 2, 2), (2, -2, 1), (2, 1, -2)\}$ is an orthogonal basis of the Euclidean space R^3 with standard inner product.

   Express $(4, 3, 2)$ as a linear combination of these basis vectors.

2. Consider the inner product space C^2 with standard inner product

$$(\alpha, \beta) = a_1\bar{b}_1 + a_2\bar{b}_2 \,;\; \alpha = (a_1, b_1)\ \beta = (a_2, b_2)$$

Determine which of the following bases are orthogonal and then determine which are orthonormal.

(a) $\{(1, 0), (0, i)\}$

(b) $\{(i, 1), (2, i)\}$

(c) $\{(2, i), (-, 2i)\}$

3. Find the norm of each of the following vectors :

(a) $\alpha = \left(\frac{1}{2}, \frac{-1}{4}, \frac{1}{3}, \frac{1}{6}\right) \in R^4$

(b) $\beta = (1-2i, 3+i, 2-5i) \in C^3$

4. In R^3, let $\alpha = (a_1, a_2, a_3)$, $\beta = (b_1, b_2, b_3)$. Determine whether (α, β) is a real inner product for R^3 if (α, β) be defined by

(a) $(\alpha, \beta) = (a_1 + a_2 + a_3)(b_1 + b_2 + b_3)$

(b) $(\alpha, \beta) = a_1b_1 + (a_2 + a_3)(b_2 + b_3) + a_3b_3$

5. Prove that for all α, β in a Euclidean space V

(a) $(\alpha, \beta) = 0$ if and only if $\|\alpha+\beta\| = \|\alpha-\beta\|$

(b) $(\alpha, \beta) = 0$ if and only if $\|\alpha+\beta\|^2 = \|\alpha\|^2 + \|\beta\|^2$

(c) $(\alpha + \beta, \alpha - \beta) = 0$ if and only if $\|\alpha\| = \|\beta\|$

6. Apply the Gram-Schmidt process to the vectors $\beta_1 = (1, 0, 1)$, $\beta_2 = (1, 0, -1)$, $\beta_3 = (0, 3, 4)$ to obtain an orthonormal basis for R^3 with the standard inner product.

7. Consider C^3, with the standard inner product. Find an orthonormal basis for the subspace spanned by

$$\beta_1 = (1, 0, i), \text{ and } \beta_2 = (2, 1, 1 + i).$$

8. Let V be a finite dimensional inner product space, and let $\{\alpha_1, \alpha_2 .. \alpha_n\}$ be an orthonormal basis for V. Show that for any vertors α, β in V.

$$(\alpha, \beta) = \sum_{k=1}^{n} (\alpha, \alpha_k)\overline{(\beta, \alpha_k)}$$

9. Apply the Gram-Schmidt process to find an orthonormal basis for the Euclidean space R^3 with standard inner product that contains

(a) the vector $\left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 0\right)$

(b) the vectors $\left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{-1}{\sqrt{3}}\right)$, $\left(\frac{2}{\sqrt{6}}, \frac{-1}{\sqrt{6}}, \frac{1}{\sqrt{6}}\right)$.

Unit : 3 □ Linear Transformations

Definition : Let V and W be vector spaces over the field F. A transformation T from V to W is said to be a **linear transformation** if.

$$\left.\begin{array}{ll} & T(\alpha+\beta)=T(\alpha)+T(\beta) \\ \text{and} & T(k\alpha)=kT(\alpha) \end{array}\right\} \quad \ldots.(1)$$

where α and β are arbitrary elements of V and k is an element of F.

The two conditions in (1) can be replaced by a single condition.

$$T(k\alpha + \beta) = kT(\alpha) + T(\beta)$$

EXAMPLES

Ex. 1. If V is any vector space, the **identity transformation** I, defined by $I(\alpha) = \alpha$, $\alpha \in V$, is a linear transformation from V. into V. The **zero transformation** O, defined by $O\alpha = 0$ is a linear transformation from V into V.

Ex. 2. If F be a field and let V be the space of polynomial functions f from F into F, given by

$$f(x) = c_0 + c_1x + .. + c_kx^k$$

Let $(Df)(x) = c_1 + 2c_2x + .. + kc_kx^{k-1}$

The D is a linear transformation from V into V.

Ex. 3. For the vector space R^2, let the transformation $T. R^2 \rightarrow R^2$ be defined in such a way that

$$T((a, b)) = (a + 2, b)$$

This transformation is not linear, because

$$\begin{aligned} T((a_1, b_1) + (a_2, b_2)) &= T((a_1 + a_2, b_1 + b_2) \\ &= (a_1 + a_2 + 2, b_1 + b_2), \end{aligned}$$

$$\begin{aligned} \text{whereas } T((a_1, b_1) + T(a_2, b_2)) &= (a_1 + 2, b_1) + (a_2 + 2, b_2) \\ &= (a_1 + a_2 + 4, b_1 + b_2) \end{aligned}$$

$$\therefore T((a_1, b_1) + (a_2, b_2)) \neq T((a_1, b_1)) + T(a_2, b_2))$$

Also $T(k(a, b)) = T(ka, kb) = T(ka + 2, kb)$

and $kT((a, b)) = k(a, + 2, b) = (ka + 2k + kb)$

So $T(k(a, b)) \neq kT(a, b))$

Hence the transformation T is not linear.

Ex. 4. Let $T. R^3 \rightarrow R^3$ be defined by $T(a_1, a_2, a_3) = (a_1, a_2\ 0), (a_1, a_2, a_3) \in R^3$

Let $\alpha = (a_1, a_2, a_3)$ and $\beta = (b_1, b_2, b_3) \in R^3$

Then $\alpha + \beta = (a_1 + b_1, a_2 + b_2, a_3 + b_3)$

$$\begin{aligned} T(\alpha + \beta) .. &= (a_1 + b_1, a_2 + b_2, 0) \\ &= (a_1, b_1, 0) + (a_2, b_2, 0) \\ &= T(\alpha) + T(\beta) \end{aligned}$$

and for any $c \in R$,

$$c\alpha = c(a_1, a_2, a_3) = (ca_1, ca_2, ca_3)$$

$$\begin{aligned} \therefore \mathrm{T}(c\alpha) = (ca_1, ca_2, 0) &= c(a_1, a_2, 0) \\ &= cT(\alpha) \end{aligned}$$

Hence T is linear transformation.

Ex. 5. Let F be a field and let $k_1, k_2 .. k_n$ be fixed elements of F. Define the mapping T from F^n to F by

$$T((x_1, x_2 .. x_n)) = k_1x_1, + k_2x_2 + .. + k_nx_n.$$

Then it is easy to check that T is linear transformation from F^n to F. A linear transformation from F^n to F is called a **linear functional of F^n** or a **linear form in x_1, x_2 .. x_n.**

Theorem : Let V and W be vector spaces over a field F and $T : V \rightarrow W$ be a linear transformation.

Then (i) $T(0) = 0'$, where 0 and $0'$ are null elements in V and W respectively.

(ii) $T(-\alpha) = -T(\alpha)$ for all $\alpha \in V$.

Proof : In V we have $0 + 0 = 0$

$$T(0 + 0) = T(0)$$

But T is linear, so $T(0 + 0) = T(0) + T(0)$

$\therefore\ T(0) + T(0) = T(0) = T(0) + 0'$

$\Rightarrow \quad T(0) = 0'$

Now, $\quad \alpha + (-\alpha) = 0 \quad$ in V

$\therefore \quad T(\alpha + (-\alpha) = T(0) = 0'$

$\Rightarrow \quad T(\alpha) + T(-\alpha) = 0'$

$\Rightarrow \quad T(-\alpha) = -T(\alpha)$

Kernel and image of a linear transformation :

Definition : Let V and W be vector spaces over the field F and $T : V \to W$ be a linear transformation. The image of T. written Im T, is the set of image points in W :

Im $T = (\alpha' \in W \mid T(\alpha) = \alpha'$ for some $\alpha \in V)$

The kernel of T, written ker T, is the set of elements in V which map into $\alpha' \in W$:

ker $T = \left\{\alpha \,\varepsilon\, V \middle| T(\alpha) = 0'\right\}$ (0′ is the zero or null element in W)

Theorem : Let $T : V \to W$ be a linear transformation. Then ker T is subspace of V.

Proof : We have ker $T = \left\{\alpha \,\varepsilon\, V \middle| T(\alpha) = 0'\right\}$

Since $T(0) = 0'$, $0 \in$ ker T and so ker T is non empty. If ker T contains *0* alone then ker T is a subspace of V.

Let us suppose that ker T contains a nonzero element α of V, Then $T(\alpha) = 0'$.

Then for any $c \in F$ (the field on which V is defined) we have $T(c\alpha) = cT(\alpha) = c0' = 0'$

This implies that $c\,\alpha \in$ ker T.

Let $\alpha, \beta \in$ ker T, then $T(\alpha) = 0'$, $T(\beta) = 0'$

Now since T is a linear transformation,

$$T(\alpha + \beta) = T(\alpha) + T(\beta) = 0' + 0' = 0'$$

So, $\alpha, \beta \in$ ker $T \quad \Rightarrow \alpha + \beta \in$ ker T

and $\alpha \in$ ker $T \quad \Rightarrow c\alpha + \in$ ker T, $c \in F$

This shows that ker T is a subspace of V.

Theorem : Let $T : V \rightarrow W$ be a linear transformation. Then ker T is subspace of W.

Proof : Since $T(0) = 0' \in \text{Im } T$, it follows that Im T is nonempty.

Let Im T contain a nonzero element say α'. Then there exists an element $\alpha \in V$ such that $T(\alpha) = \alpha'$.

This implies that $T(c\alpha) = cT(\alpha) = c\alpha'$ for all $c \in F$.

$\therefore \; c\alpha' \in \text{Im } T$.

Agani, let $\alpha', \beta' \in \text{Im } T$, then there are elements $\alpha', \beta \in V$ such that

$$T(\alpha) = \alpha', \; T(\beta) = \beta'.$$

Now, $\quad T(\alpha + \beta) = T(\alpha) + T(\beta) = \alpha' + \beta'$

$\Rightarrow \quad \alpha' + \beta' \in \text{Im } T$

Hence $\quad \alpha', \beta' \in \text{Im } T \Rightarrow \alpha' + \beta' \in \text{Im } T$

and $\quad c\alpha' \in \text{Im } T, \; c \in F$

Hence $\quad$ Im T is a subspace of W.

Theorem : Let $T : V \rightarrow W$ be a linear transformation such that ker $T = \{0\}$. Then the images of a linearly independent set of vectors in V are linearly independent in W.

Proof : Let be $\{\alpha_1, \alpha_2 .. \alpha_n\}$ be a linearly independent set in V, we have to prove that $\{T(\alpha_1), T(\alpha_2) .. T(\alpha_n)\}$ is a linearly independent set of vectors in W.

Let $\quad c_1, c_2 .. c_n \in F$.

$$c_1T(\alpha_1) + c_2T(\alpha_2) + .. + c_nT(\alpha_n) = \underline{0}'$$

Since T is a linear transformation, we have

$$T(c_1\alpha_1 + c_2\alpha_2 + .. + c_n\alpha_n) = c_1T(\alpha_1) + .. + c_nT(\alpha_n)$$

$$\therefore \quad T(c_1\alpha_1 + c_2\alpha_2 + .. + c_n\alpha_n) = 0'$$

This implies that $c_1\alpha_1 + c_2\alpha_2 + .. + c_n\alpha_n \in \ker T$.

But ker $T = \{0\}$ is given.

Hence $\quad c_1\alpha_1 + c_2\alpha_2 + .. + c_n\alpha_n = 0$.

But since $\{\alpha_1, \alpha_2 .. \alpha_n\}$ in a linearly independent set of vectors in V, we must have

$$c_1 = c_2 = ... = c_n = 0$$

Thus it follows that

$$c_1T(\alpha_1) + c_2T(\alpha_2) + .. + c_nT(\alpha_n) = 0$$

iniplies that $c_1 = c_2 = ... = c_n = 0$

Hence $\{T(\alpha_1), T(\alpha_2) .. T(\alpha_n)\}$ is an independent set of vectors in W.

Theorem : Let $T : V \rightarrow W$ be a linear transformation and $\{\alpha_1, \alpha_2 .. \alpha_n\}$ be a basis of V. Then $\{T(\alpha_1), T(\alpha_2) .. T(\alpha_n)\}$ generates Im T.

Proof : Let $\alpha' \in$ Im T. Then exists at least one element, say α, in V such that $T(\alpha) = \alpha'$.

Since $\{\alpha_1, \alpha_2 .. \alpha_n\}$ is a basis of V, there exist $c_1, c_2 .. c_n \in F$ such that

$$\begin{aligned} \alpha &= c_1\alpha_1 + c_2\alpha_2 + .. + c_n\alpha_n \\ \therefore\ T(\alpha) &= T(c_1\alpha_1 + c_2\alpha_2 + .. + c_n\alpha_n) \\ &= c_1T(\alpha_1) + c_2T(\alpha_2) + .. + c_nT(\alpha_n), \text{ since } T \text{ is linear.} \end{aligned}$$

$$\therefore \quad \alpha' = c_1T(\alpha_1) + c_2T(\alpha_2) + .. + c_nT(\alpha_n)$$

Since each $T(\alpha_i) \in$ Im T, it follows that Im T is generated by $\{T(\alpha_1), T(\alpha_2) .. T(\alpha_n)\}$.

Nullity and Rank of a linear transformation :

Definition : Let V and W be vector spaces over the field F and let T be a linear transformation from V into W. The **nullity of** T is the dimensin of the subspace ker T of V and the **rank of T** is the dimension of the subspace Im T of W.

If V be a finite dimensional vector space then both ker T and Im T are finite dimensional.

Theorem : Let V and W be vector spaces over the field F and V is finite dimensional.

If $T : V \rightarrow W$ isa linear transformation then the nullity of T + rank of T = dim V.

Proof : Let dim $V = n$.

Let $\{\alpha_1, \alpha_2 .. \alpha_k\}$ be a basis of ker T. There are vectors $\alpha_{k+1}, \alpha_{k+2} .. \alpha_n$ in V such that $\{\alpha_1, \alpha_2 .. \alpha_n\}$ is a basis for V. We shall now prove that $\{T(\alpha_{k+1}), T(\alpha_{k+2}) ... T(\alpha_n)\}$ is a basis for Im T. We have noted that the vectors $T(\alpha_1), T(\alpha_2), .. T(\alpha_n)$ certainly span Im T, and since $\alpha_1, \alpha_2 .. \alpha_k \in$ ker T, we have $T(\alpha_1) = T(\alpha_2) = T(\alpha_k) = 0'$. Hence it follows that the voctors $T(\alpha_{k+1}), T(\alpha_{k+2}) \ldots T(\alpha_n)$ span Im T. We shall show that $T(\alpha_{k+1})$, $T(\alpha_{k+2}) \ldots T(\alpha_n)$ are linearly independent.

Suppose that there are $c_{k+1}, c_{k+2} \ldots c_n \in F$ such that

$$c_{k+1}\, T(\alpha_{k+1}) + c_{k+2}\, T(\alpha_{k+2}) + \ldots + c_n\, T(\alpha_n) = 0'$$

Since T is a linear transformation, it follows that

$$T(c_{k+1}\alpha_{k+1} + c_{k+2}\alpha_{k+2} + \ldots + c_n\alpha_n) = 0'$$

Hence $\alpha = c_{k+1}\alpha_{k+1} + c_{k+2}\alpha_{k+2} + \ldots + c_n\alpha_n \in \ker T$.

Again, by assumption, $\{\alpha_1, \alpha_2 \ldots \alpha_k\}$ is a basis of ker T so $\alpha = b_1\alpha_1 + b_2\alpha_2 + \ldots + b_k\alpha_k$, $b_1, b_2, \ldots b_k \in F$.

Thus $\alpha = b_1\alpha_1 + b_2\alpha_2 + \ldots + b_k\alpha_k = c_{k+1}\alpha_{k+1} + c_{k+2}\alpha_{k+2} \cdots c_n\alpha_n$

or, $b_1\alpha_1 + b_2\alpha_2 + \ldots + b_k\alpha_k - c_{k+1}\alpha_{k+1} - c_{k+2}\alpha_{k+2} - \ldots - c_n\alpha_n = 0$

But $\{\alpha_1, \alpha_2 \ldots \alpha_n\}$ is a basis of V, so we must have,

$$b_1 = b_2 = \ldots = b_k = c_{k+1} = \ldots c_n = 0$$

Hence $T(\alpha_{k+1}), T(\alpha_{k+2}) \ldots T(\alpha_n)$ are linearly independent and as they span Im T, $\{T(\alpha_{k+1}), T(\alpha_{k+2}) \ldots T(\alpha_n)\}$ is a basis of Im T.

Clearly dim (In T) = $n - k$ = rank of T.

Also, dim (ker T) = k = nullity of T.

Hence rank of T + nullity of $T = n = \dim V$.

Theorem : Let V and W are both finite dimensional vector spaces of the same dimension and T is a linear transformation from V to W, then T is one-one if and only if T is onto.

Proof : Let T be one-one then clearly $\ker T = \{0\}$. and so dim ker $T = 0$.

Now, ker T is a subspace of V and Im T is a subspace of W and also dim ker T + dim Im T = dim V.

Since dim ker $T = 0$, we have dim Im T = dim V = dim W (given)

$\therefore$ dim Im T = dim $W \Rightarrow$ Im $T = W$

Hence the transformation T is onto.

Connversely, suppose that T is onto.

Then $\operatorname{Im} T = W$

$\therefore$ $\dim \operatorname{Im} T = \dim W = \dim V.$

Since $\dim \ker T + \dim \operatorname{Im} T = \dim V$, we have

$\dim \ker T = 0.$ Hence $\ker T = \{0\}$ and T is one-one.

Theorem : If V and W be are finite dimensional vector spaces over a field F and if there exists a linear transformation $T : V \to W$ which is both one-one and on to, then $\dim V = \dim W$.

Proof : Since by assumption T is one-one, we have $\ker T = \{0\}$

$\therefore \dim \ker T = 0$

Again, since T is onto, we have $\operatorname{Im} T = W$

$\therefore \dim \operatorname{Im} T = \dim W$

Now, $\dim \ker T + \dim \text{In } T = \dim V$,

$\Rightarrow$ $0 + \dim W = \dim V.$

i. e. $\dim V = \dim W.$

Theorem : Let V be a finite dimensional vector space over the field F and let $\{\alpha_1, \alpha_2 .. \alpha_n\}$ be a basis for V. Let W be a vector space over the same field F and $\beta_1, \beta_2 .. \beta_n$ be any vecotors in W. Then there is precisely one linear transformation T from V into W such that

$$T(\alpha_i) = \beta_i, \qquad i = 1, 2, .. n,$$

Proof : To prove that there is some linear transformation T with $T(\alpha_i) = \beta_i$, we proceed as follows. Let α be an arbitrary element in V. Since $\{\alpha_1, \alpha_2 .. \alpha_n\}$ is a basis for V, there exist unique scalars $c_1, c_2 .. c_n$ in F such that

$$\alpha = c_1\alpha_1 + c_2\alpha_2 + .. + c_n\alpha_n$$

For this vector α we define

$$T(\alpha) = c_1\beta_1 + c_2\beta_2 + .. + c_n\beta_n$$

Then T is a well-defined rule for associating each vector $\alpha \in V$ to a vector $T(\alpha)$ in W.

From this definition it is clear that $T(\alpha_i) = \beta_i$ for each i.

Let us now examine whether T is linear.

Let $\gamma, \delta \in V$ and

$$\gamma = p_1\alpha_1 + p_2\alpha_2 + .. + p_n\alpha_n$$
$$\delta = q_1\alpha_1 + q_2\alpha_2 + .. + q_n\alpha_n$$

where p_i, q_i are unique scalars in F determined by the basis $\{\alpha_1, \alpha_2 .. \alpha_n\}$.

Then $\gamma + \delta = (p_1 + q_1)\alpha_1 + .. + (p_n + q_n)\alpha_n$

and $b\gamma = (bp_1)\alpha_1 + (bp_2)\alpha_2 + .. + (bp_n)\alpha_n, b \in F$

so, by definition

$$T(\gamma + \delta) = (p_1 + q_1)\beta_1 + .. (p_n + q_n)\beta_n,$$
$$T(b\gamma) = (bp_1)\beta_1 + .. + (bp_n)\beta_n,$$

and $T(\gamma) + T(\delta) = (p_1\beta_1 + .. + p_n\beta_n) + (q_1\beta_1 + .. + q_n\beta_n)$

$$= (p_1 + q_1)\beta_1 + .. + (p_n + q_n)\beta_n = T(\gamma + \delta),$$

$$\begin{aligned} bT(\gamma) &= b(p_1\beta_1 + p_2\beta_2 + .. + p_n\beta_n) \\ &= (bp_1)\beta_1 + .. + (bp_n)\beta_n \\ &= T(b\gamma) \end{aligned}$$

This shows that T is linear.

We have to show that T is unique linear transformation such that $T(\alpha_i) = \beta_i$, $i = 1, 2, .. n$,

Let us assume that there exists another linear transformation $T' : V \rightarrow W$ such that $T'(\alpha_i) = \beta_i$, $i = 1, 2, \; n$,

Now, $T'(\alpha) = T'(c_1\alpha_1 + c_2\alpha_2 + .. + c_n\alpha_n)$

$= c_1T'(\alpha_1) + c_2T'(\alpha_2) + .. + c_nT'(\alpha_n)$, (Q T' is linear)

$= c_1\beta_1 + c_2\beta_2 + .. + c_n\beta_n$

$= T(\alpha)$

Thus $T'(\alpha) = T(\alpha)$ for every $\alpha \in V$.

Hence $T' = T$

Thus T is unique linear transformation such that $T(\alpha_i) = \beta_i$, $i = 1, 2, .. n$

EXAMPLES

Ex. 1. Let V be the vector space of n-square matrices over field F. Let M be an arbitray matrix in V. Let $T: V \to W$ be defined by $T(A) = AM + MA$ where $A \in V$. Show that T is linear.

For any $A, B \in V$ and any $c \in F$, we have

$$\begin{aligned} T(A+B) &= (A+B)M + M(A+B) \\ &= AM + BM + MA + MB \\ &= (AM + MA) + (BM + MB) = T(A) + T(B) \end{aligned}$$

$$\begin{aligned} \text{and } T(cA) &= (cA)M + M(cA) \\ &= c(AM) + c(MA) = c(AM + MA) \\ &= cT(A) \end{aligned}$$

Accordingly, T is linear.

Ex. 2. Let V be the vector space of 2 × 2 matrices over K and let $M = \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}$.

Let $T: V \to W$ be the linear transformation defined by $T(A) = AM - MA$. Find the basis and the dimension of ker T.

We seek the set of $\begin{pmatrix} x & y \\ s & t \end{pmatrix}$ such that

$$T\left(\begin{pmatrix} x & y \\ s & t \end{pmatrix}\right) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Now,
$$\begin{aligned} T\left(\begin{pmatrix} x & y \\ s & t \end{pmatrix}\right) &= \begin{pmatrix} x & y \\ s & t \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix} - \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}\begin{pmatrix} x & y \\ s & t \end{pmatrix} \\ &= \begin{pmatrix} x & 2x+3y \\ s & 2s+3t \end{pmatrix} - \begin{pmatrix} x+2s & y+2t \\ 3s & 3t \end{pmatrix} \\ &= \begin{pmatrix} -2s & 2x+2y-2t \\ -2s & 2s \end{pmatrix} - \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \end{aligned}$$

$$\Rightarrow \quad s = 0 \quad \text{and} \quad x + y = t$$

Among the four variables x, y, s, t we observe that $s = 0$ and the variables x, y and t are connected by $x + y = t$. This means that there are only two free variables x and y. Hence dim ker $T = 2$.

To obtain a basis of ker T we set $x = 1, y = -1$

so that the solution is $x = 1, y = -1, t = 0, s = 0$ and the matrix is $\begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}$

Again, setting $x = 1, y = 0$ we get the matrix $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

Thus $\left\{\begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right\}$ is a basis of ker T.

The Algebra of Linear Transformations :

In the study of linear transform a vector space V into a vector space W, it is of fundamental importance that the set of these transformations inherits a natural vector space structure. The theorem below explores this concept.

Theorem : Let V and W be vector spaces over the field F. Let T and U be linear transformations from V into W.

The function $(T + U)$ defined by

$$(T + U)(\alpha) = T(\alpha) + U(\alpha), \alpha \in V.$$

is a linear transformation from V into V. If c is any element of F, the function (cT) defined by

$$(cT)(\alpha) = c(T(\alpha))$$

is a linear transformation from V into V. The set of all linear transformations from V into W, together with the addition and scalar multiplication defined above is a vector space over the field F.

Proof : Suppose T and U are linear transformations from V into W and define $(T + U)$ as above. Then

$$\begin{aligned}(T + U)(\alpha + \beta) &= T(\alpha + \beta) + U(\alpha + \beta), \alpha, \beta \in V, \\ &= T(\alpha) + U(\alpha) + T(\beta) + U(\beta) \\ &= (T + U)(\alpha) + (T + U)(\beta)\end{aligned}$$

$$\begin{aligned}\text{Also, } (T + U)(c\alpha) &= T(c\alpha) + U(c\alpha), \qquad c \in F \\ &= c(T(\alpha) + c(U(\alpha)) \\ &= c(T(\alpha) + U(\alpha)) \\ &= c(T + U)(\alpha))\end{aligned}$$

Thus $(T + U)$ is a linear transformation.

Now we show that the function (cT) is a linear transformation

We have, $(cT)(\alpha + \beta) = c(T(\alpha + \beta))$, by definition

$= c(T(\alpha) + T(\beta))$, since T is linear

$= c(T(\alpha)) + c(T(\beta))$

$= cT(\alpha) + cT(\beta)$

And for any $d \in F$,

$$\begin{aligned} (cT)(d\alpha) &= c(T(d\alpha) \\ &= c(dT(\alpha)) \\ &= cd(T(\alpha) = dc(T(\alpha)) \\ &= d[c(T(\alpha))] = d[cT(\alpha)] \end{aligned}$$

which shows that (cT) is a linear transformation. To verify that the set of linear transformations of V into W (together with these operations) is a vector space, we have to check each of the conditions on the vector addition and scalar multiplication. The task is lengthy but not difficult. The zero vector in this space will be the zero transformation which sends every vector of V into the zeto vector in W; each of the properties of the two operations follows from the corresponding property of the operations in the space W.

Theorem : Let V, W and Z be vector spaces over the field F. Let T be a linear transformation from V into W and U, a linear transformation from W into Z. Then the composed function UT defined by $(UT)(\alpha) = U(T(\alpha))$ is a linear transformation from V into Z.

Proof : We have. $(UT)(\alpha + \beta) = U(T(\alpha + \beta))$, $\quad \alpha, \beta \in V$

$= U(T(\alpha) + T(\beta))$, since T is linear on V.

$= U(T(\alpha)) + U(T(\beta))$,

since U in linear on W

$= UT(\alpha) + UT(\beta)$

and $UT(c\alpha) = U(T(c\alpha))$, $\quad c \in F$,

$= U(cT(\alpha))$

$$= c[U(T(\alpha))]$$
$$= c(UT(\alpha))$$

This shows that UT is a linear transformation from V to Z.

Definition : The function T from V into W is called **invertible** if there exists a function U from W into V such that the UT is the identity function on V and TU is the identity function on W. If T is invertible, the function U is unique and is denoted by T^{-1},

Funthemore, T is invertible if and only if

(a) T is one-one i.e. $T(\alpha) = T(\beta)$ implies $\alpha = \beta$

(b) T is onto, i.e. range of T is W.

Theorem : If V and W are vector spaces over the field F and let T be a linear transformation from V into W. If T is invertible, then the inverse function T^{-1} is a linear transformation from W onto V.

Proof : Let β_1, β_2 be vectors in W and c be a scalar. We wish to show that

$$T^{-1}(\beta_1 + \beta_2) = T^{-1}(\beta_1) + T^{-1}(\beta_2)$$

and $$T^{-1}(c\beta_1) = cT^{-1}(\beta_1)$$

Let $T^{-1}\beta_i = \alpha_i$, $i = 1, 2$

Then $T(\alpha_i) = \beta_i$, $i = 1, 2$

Since T is linear, $T(\alpha_1 + \alpha_2) = T(\alpha_1) + T(\alpha_2)$

$$= \beta_1 + \beta_2$$

Thus $\alpha_1 + \alpha_2$ is unique vector in V which is sent by T into $\beta_1 + \beta_2$ and so

$$T^{-1}(\beta_1 + \beta_2) = \alpha_1 + \alpha_2 = T^{-1}(\beta_1) + T^{-1}(\beta_2)$$

Also, $T(c\alpha_1) = cT(\alpha_1)$, since T is linear

$$= c\beta_1,$$

That is the unique vector $c\alpha_1$ is sent by T into $c\beta_1$ and so,

$$T^{-1}(c\beta_1) = c\alpha_1 = cT^{-1}(\beta_1)$$

$\therefore$ T^{-1} is linear.

Isomorphism : Let V and W be finite dimensional vector spaces over a field F. A linear transformation $T : V \to W$ is said to be an isomorphism if T is both one-one and onto.

Theorem : Two finite dimensional vector spaces V and W over a field F are isomorphic if and only if dim V = dim W.

Proof : Let V and W be isomorphic. Then there exists a linear transformation $T : V \to W$ such that T is both one-one and onto.

Since T is one-one, ker $T = \{0\}$

and since T is onto Im $T = W$.

Hence from the relation

$$\dim \ker T + \dim \operatorname{Im} T = \dim V, \text{ we get } \dim W = \dim V.$$

Conversely, let dim V = dim $W = n$.

Let $\{\alpha_1, \alpha_2 .. \alpha_n\}$ be a basis of V and $\{\beta_1, \beta_2 .. \beta_n\}$ be a basis of W.

Then, there exists a linear transformation $T : V \to W$ such that $T(\alpha_i) = \beta_i$, $i = 1, 2 .... n$.

Since $\{T(\alpha_1), T(\alpha_2) ... T(\alpha_n)\}$ generates Im (T), so $\{\beta_1, \beta_2 .. \beta_n\}$ generates Im (T).

i. e. $L\{\beta_1, \beta_2 .. \beta_n\} = \text{In}(T)$

But $L\{\beta_1, \beta_2 .. \beta_n\} = \text{W}$, since $\{\beta_1, \beta_2 .. \beta_n\}$is a basis of W.

Hence In (T) = W.

This means that T is onto.

Agains from the relation

$$\dim \ker T + \dim \operatorname{Im} T = \dim V.$$

we have, dim ker T + dim W = dim V.

But dim V = dim W, so, dim ker $T = 0$

$\Rightarrow$ ker $T = \{0\}$

$\Rightarrow$ T is one-one.

Thus, T being both one-one and onto, the vector spaces V and W are isomorphic.

Theorem : Every n-dimensional vector space over the field F is isomorphic to the space F^n.

Proof : Let $\{\alpha_1, \alpha_2 .. \alpha_n\}$ be on ordered basis of V. Then any vector ξ of V can be uniquely expresed as

$$\xi = c_1\alpha_1 + c_2\alpha_2 + .. + c_n\alpha_n, \text{ where } c_1, c_2 .. c_n \in F$$

Let us define a transformation $T : V \to F$ by

$$T(\xi) = \begin{pmatrix} c_1 \\ c_2 \\ \mathrm{M} \\ c_n \end{pmatrix}$$

Let $\quad \eta = b_1\alpha_1 + b_2\alpha_2 + .. + b_n\alpha_n \in V$

and $\quad \delta = d_1\alpha_1 + d_2\alpha_2 + .. + d_n\alpha_n \in V$

then $\quad \eta + \delta = (b_1 + d_1)\alpha_1 + (b_2 + d_2)\alpha_2 + .. + (b_n + d_n)\alpha_n \in V$

so that $\quad T(\eta) = \begin{pmatrix} b_1 \\ b_2 \\ \mathrm{M} \\ b_n \end{pmatrix} \qquad T(\delta) = \begin{pmatrix} d_1 \\ d_2 \\ \mathrm{M} \\ d_n \end{pmatrix}$

and $\quad T(\eta+\delta) = \begin{pmatrix} b_1 + d_1 \\ b_2 + d_2 \\ .. \\ b_n + d_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ .. \\ b_n \end{pmatrix} + \begin{pmatrix} d_1 \\ d_2 \\ .. \\ d_n \end{pmatrix} = T(\eta) + T(\delta) \qquad ....(1)$

Let $\quad a \in F$, then $a\eta \in V$ and

$a\eta = (ab_1)\alpha_1 + (ab_2)\alpha_2 + .. + (ab_n)\alpha_n$

$\therefore \quad \mathrm{T}(a\eta) = \begin{pmatrix} ab_1 \\ ab_2 \\ \mathrm{M} \\ ab_n \end{pmatrix} = a\begin{pmatrix} b_1 \\ b_2 \\ .. \\ b_n \end{pmatrix} = aT \qquad ....(2)$

Results (1) and (2) show that the transformation T is linear.

To prove that T is isomorphism, we have to show that T is one-one and onto.

Now, for any two vector η and $\delta \in V$, we have

$$T(\eta)=T(\delta)\Rightarrow\begin{pmatrix} b_1 \\ b_2 \\ \mathrm{M} \\ b_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \mathrm{M} \\ b_n \end{pmatrix}\Rightarrow b_1=d_1, b_2=d_2, \mathrm{L}\ \ b_n=d_n$$

$\Rightarrow \quad \eta=\delta$

Hence T is one-one

Now to prove that T is onto let $\begin{pmatrix} p_1 \\ p_2 \\ \mathrm{M} \\ p_n \end{pmatrix}$ be an arbitray element in F^n, then

$p_1\alpha_1 + p_2\alpha_2 + .. + p_n\alpha_n \in V$ and

$$T(p_1\alpha_1 + p_2\alpha_2 + .. + p_n\alpha_n) = \begin{pmatrix} p_1 \\ p_2 \\ p_n \end{pmatrix}$$

This implies that T is onto.

Hence T is an isomorphism.

EXAMPLES

Ex. 1. Let $T: R^3 \to R^3$ difined by $T(x, y, z) = (2x, 4x - y, 2x + 3y - z)$. Show that T is invertible and determine T^{-1}.

The kernel of the transformation ker T is the set of all (x, y, z) such that $T(x, y, z) = (0, 0, 0)$, i.e.

$$T(x, y, z) = (2x, 4x - y, 2x + 3y - z) = (0, 0, 0)$$

Thus ker T is the solution space of the homogeneous system.

$$2x = 0,\ 4x - y = 0,\ 2x + 3y - z = 0$$

which has only the trivial solution $(0, 0, 0)$.

Thus ker $T = \{0\}$. Hence T is nonsingular and so T is invertible.

Let (r, s, t) be the image of (x, y, z) under T ; then (x, y, z) is the image of (r, s, t) under T^{-1}:

$T(x, y, z) = (r, s, t)$ and $T^{-1}(r, s, t) = (x, y, z)$

We will find the values of x, y, z in terms of r, s and t, and then substitute in the above formula for T^{-1}.

From $T(x, y, z) = (2x, 4x - y, 2x + 3y - z) = (r, s, t)$

we have, $\quad 2x = r,\ 4x - y = s, \quad 2x + 3y - z = t$

$$\therefore \quad x = \frac{r}{2}, \quad y = 2r - s, \quad z = 7r - 3s - t$$

and so, T^{-1} is given by

$$T^{-1}(r,s,t) = \left(\frac{1}{2}r, 2r - s, 7r - 3s - t\right)$$

Matrix representation of a linear transformation :

Let V be an n-dimensional vector space over the field F and let W be an m-dimensional vector space over F. Let $\{\alpha_1, \alpha_2 .. \alpha_n\}$ be a basis for V and $\{\beta_1, \beta_2 .. \beta_n\}$ be a basis for W. If T is any linear transformation from V into W, then T is determined by its action on the vectors α_j, $j = 1, 2, .. n$. Each of the n vectors $T(\alpha_1)$, $T(\alpha_2)$ .. $T(\alpha_n)$ is uniquely expressible as a linear combination.

$$T(\alpha_j) = \sum_{i=1}^{m} A_{ij}\beta_i$$

of the β_i, the scalars $A_{1j}, A_{2j} .. A_{nj}$ being the coordinates of $T(\alpha_j)$ in the basis $\{\beta_1, \beta_2 .. \beta_m\}$. Accordingly the transformation T is determined by mn scalars A_{ij} via the formulas (1). The m × n matrices $A = (A_{ij})$ is called the matrix of T relative to the pair of ordered bases $\{\alpha_1, \alpha_2 .. \alpha_n\}$ and $\{\beta_1, \beta_2 .. \beta_m\}$

The jth column of A is the coordinate vector of $T(a_j)$ relative to the ordered basis $\{\beta_1, \beta_2 .. \beta_m\}$

Let $\xi = x_1\alpha_1 + x_2\alpha_2 + .. + x_n\alpha_n$ be an arbitrary vector of V and let

$T(\xi) \quad = y_1\beta_1 + y_2\beta_2 + .. + y_m\beta_m,\ x_1 .. x_n,\ y_1 .. y_m \in F$

Now $T(\xi) \quad = T(x_1\alpha_1 + x_2\alpha_2 + .. + x_n\alpha_n)$

$\quad = x_1T(\alpha_1) + x_2T(\alpha_2) + .. + x_nT(\alpha_n)$ $\qquad$ since T is a linear transfomation.

$\quad = x_1(A_{11}\beta_1 + A_{21}\beta_2 + .. + A_{m1}\beta_m) + x_2(A_{12}\beta_1 + A_{22}\beta_2 + .. + A_{m2}\beta_m) + .. + x_n(A_{1n}\beta_1 + A_{2n}\beta_2 + .. + A_{mn}\beta_m)$

$$= (A_{11}x_1 + A_{12}x_2 + .. + A_{1n}x_n)\beta_1 + (A_{21}x_1 + A_{22}x_2 + .. + A_{2n}x_{n.})\beta_2 + .. + (A_{m1}x_1 + A_{m2}x_2 + .. + A_{mn}x_m)\beta_m$$

$$= y_1\beta_1 + y_2\beta_2 + .. + y_m\beta_m$$

Since $\beta_1, \beta_2 .. \beta_m$ are linearly independent, we have

$$y_1 = A_{11}x_1 + A_{12}x_2 + .. + A_{1n}x_n$$

$$y_2 = A_{21}x_1 + A_{22}x_2 + .. + A_{2n}x_n$$

$$y_m = A_{m1}x_1 + A_{m2}x_2 + .. + A_{mn}x_n$$

In matrix notation we have

$$\begin{pmatrix} y_1 \\ y_2 \\ \mathrm{M} \\ y_m \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} & \mathrm{L} & A_{1n} \\ A_{21} & A_{22} & & A_{2n} \\ & & & \\ A_{m1} & A_{m2} & & A_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \mathrm{M} \\ x_n \end{pmatrix}$$

or, $Y = AX$, ... (2)

where X is the coordinate vector of an arbitray element ξ in V relative to the ordered basis $\{\alpha_1, \alpha_2 .. \alpha_n\}$ and Y is the coordinate vector of $T(\xi)$ in W relative to the ordered basis $\{\beta_1, \beta_2 .. \beta_m\}$.

(2) is the matrix representation of the linear transformation relative to the chosen ordered bases of V and W.

The coordinates of $T(\alpha_j)$ are given by the ith column of Λ for $j = 1, 2, .. n$.

The matrix of linear transformation T depends not only on the choice of basis of two vectorspaces V ans W but also on the order of the bases.

If we know the ordered bases for the spaces V and W, then the associated matrix for the linear transformation can be constructed.

Conversely, if a matrix $A = (A_{ij})$ of order $m \times n$ be selected first then there exists a unique linear transformation $T : V \rightarrow W$ whose matrix is A, because if we prescribe the jth column of $A = (A_{ij})$ as the coordinates of $T(\alpha_j)$ relative to basis $\{\beta_1, \beta_2 .. \beta_m\}$, then T is determined uniquely with $A = (A_{ij})$ as the associated matrix.

Thus there is one-one correspondence between the set of all linear transformations of V to W and the set of all $m \times n$ matrices with elements from the scalar field F.

Hence we have

Theorem : Let V be an n-dimensional vector space over the field F and W, an m-dimensional vector space over F. Let B be an ordered basis for V and B', an ordered basis for W. For each linear transformation T from V into W, there is an mxn matrix A with entries in F such that

$$[T(\alpha)]_{B'} = A[\alpha]_B$$

for every vector α in V. Furthermore, $T \rightarrow A$ is a one-one correspondence between the set of all linear transformations from V into W and the set of all mxn matrices over the field F.

Theorem : Let V and W be finite dimensional vector spaces over a field F and let $T : V \rightarrow W$ be a linear transformation. Then the rank of T = the rank of the matrix of T.

Proof : Let $\{\alpha_1, \alpha_2 .. \alpha_n\}$ and $\{\beta_1, \beta_2 .. \beta_m\}$ be a pair of ordered bases of V and W respectively. Let m (T) be the matrix of T relative to those bases such that

$$m(T) = \begin{pmatrix} C_{11} & C_{12} & \mathrm{L} & C_{1n} \\ C_{21} & C_{22} & & C_{2n} \\ & & & \\ C_{m1} & C_{m2} & & C_{mn} \end{pmatrix}, \quad c_{ij} \varepsilon F$$

Then $\quad T(\alpha_1) = c_{11}\beta_1 + c_{21}\beta_2 + .. + c_{m1}\beta_m$

$T(\alpha_1) = c_{12}\beta_1 + c_{22}\beta_2 + .. + c_{m2}\beta_m$

..

$T(\alpha_n) = c_{1n}\beta_1 + c_{2n}\beta_2 .. c_{mn}\beta_n$

Since $\{\alpha_1, \alpha_2 .. \alpha_n\}$ is a basis of V, $\{T(\alpha_1), T(\alpha_2) .. T(\alpha_n)\}$ generates Im T. Let rank of $T = r$, then by definition dim Im $T = r$.

So we may suppose that $\{T(\alpha_1), T(\alpha_2) .. T(\alpha_r)\}$ is a basis of Im T. Hence the remaining vertors $T(\alpha_{r+1}), .. T(\alpha_n)$ belong to $L\{T(\alpha_1), T(\alpha_2) .. T(\alpha_r)\}$

Let us consider the isomorphism $\phi : W \rightarrow F^m$ defined by

$$\phi(c_1\beta_1 + c_2\beta_2 + \mathrm{L} + c_m\beta_m) = \begin{pmatrix} c_1 \\ c_2 \\ \mathrm{M} \\ c_m \end{pmatrix}, c_i \varepsilon F$$

Then $\phi(T(\alpha_1)) = \begin{pmatrix} c_{11} \\ c_{21} \\ \text{M} \\ c_{m1} \end{pmatrix}$, $\phi(T(\alpha_2)) = \begin{pmatrix} c_{12} \\ c_{22} \\ \text{M} \\ c_{m2} \end{pmatrix}$, ..$\phi(T(\alpha_n)) = \begin{pmatrix} c_{1n} \\ c_{2n} \\ \text{M} \\ c_{mn} \end{pmatrix}$

Since $\{T(\alpha_1), T(\alpha_2) .. T(\alpha_r)\}$is a lineraly independent set and ϕ is an isomorphism $\{\phi(T(\alpha_1)), \phi(T(\alpha_2)) .. \phi(T(\alpha_r))\}$ is a linearly independent set of m truples of F^m.

Since each of $T(\alpha_{r+1})$, $T(\alpha_{r+2})$ .. $T(\alpha_n)$, belongs to $L\{T(\alpha_1), T(\alpha_2) .. T(\alpha_r)\}$ and ϕ is an isomorphism, each of $\phi(T(\alpha_{r+1}))$, $\phi(T(\alpha_{r+2}))$, .. $\phi(T(\alpha_n))$ belongs to $L\{\phi(T(\alpha_1)), \phi(T(\alpha_2)) .. \phi(T(\alpha_r))\}$.

This means that the first r columns of the matrix m(T) are linearly independent and the remaining $(n-r)$ columns are linear combinations of the first r columns. So the rank of $m(T) = r$.

Hence rank of T = rank of matrix of T.

Ex. 2. Let $T : R^3 \to R^3$ be a linear transformation defined by $T(x, y, z) = (3x + 2y - 4z, x - 5y + 3z)$.

Find the matrix of T relative to the bases $\alpha_1 = (1, 1, 1)$, $\alpha_2 = (1, 1, 0)$,

$\alpha_3 = (1, 0, 0)$ of R^3 and $\beta_1 = (1, 3)$, $\beta_2 = (2, 5)$ of R^2.

We have $T(\alpha_1) = (1, -1) = -7\beta_1 + 4\beta_2$

$T(\alpha_2) = (5, -4) = -33\beta_1 + 19\beta_2$

$T(\alpha_3) = (3, 1) = -13\beta_1 + 8\beta_2$

$\therefore$ The matrix of $T = \begin{pmatrix} -7 & -33 & -13 \\ 4 & 19 & 8 \end{pmatrix}$

Linear Operators : Let V be a vector space over a field F. We consider the special case of linear transformation $T : V \to W$. These transformations are called linear operators or linear transformations on V.

Theorem . Let T be a linear operator on an n-deimensional vector space over a field F. Then the following statements are equivalent.

(i) T is non singular.

(ii) T is one-one

(iii) T is onto

(iv) T maps a linearly independent set of V to another linearly independent set.

(v) T maps a basis of V to another basis.

Proof : Let (i) be true, Then T is one-one and onto.

Hence $(i) \Rightarrow (ii)$

Next let us suppose that (ii) holds, i.e. T is one-one. Then ker $T = \{0\}$ and so dim ker $T = 0$.

But $\quad$ dim ker T + dim Im T = dim V

So, $\quad 0$ + dim Im T = dim V

$\therefore$ dim Im $T = n$

Since Im T is a subset of V and dim Im = dim V + n it follows that Im $T = V$.

Hence T is onto.

$\therefore \quad (ii) \Rightarrow (iii)$

Let us now assume that (iii) holds, Let $\{\alpha_1, \alpha_2 .. \alpha_r\}$ be a linearly independent set in V. We shall prove that $\{T(\alpha_1), T(\alpha_2) .. T(\alpha_r)\}$ is also linearly independent set in V.

Since T is onto, dim Im T = dim V

so dim ker $T = 0 \Rightarrow$ ker $T = \{0\}$

To prove that $\{T(\alpha_1), T(\alpha_2) .. T(\alpha_r)\}$ is a linearly independent set in V, let us suppose that there exist $c_1, c_2 .. c_r \in F$ such that

$$c_1T(\alpha_1) + c_2T(\alpha_2) + .. + c_rT(\alpha_r)\} = 0$$

Since T is linear, we have

$$\mathrm{T}(c_1\alpha_1 + c_2\alpha_2 + .. + c_r\alpha_r) = 0$$

$$\Rightarrow \quad c_1\alpha_1 + c_2\alpha_2 + .. + c_r\alpha_r \;\varepsilon \text{ ker } T$$

But ker $T = \{0\}$, hence $c_1\alpha_1 + c_2\alpha_2 + .. + c_r\alpha_r = 0$

Again since $\alpha_1, \alpha_2 .. \alpha_r$ are linearly independent we must have $c_1 = c_2 = .. = c_r = 0$

Thus, $c_1T(\alpha_1) + c_2T(\alpha_2) + .. + c_rT(\alpha_r) = 0$

implies $c_1 = c_2 = .. = c_r = 0$

So $T(\alpha_1), T(\alpha_2) .. T(\alpha_r)$ are linearly independent

or, $(iii) \Rightarrow (iv)$

Let (iv) hold. Let us suppose that $\{\alpha_1, \alpha_2 .. \alpha_n\}$ is a basis of V. Then by (iv) it follows that $\{T(\alpha_1), T(\alpha_2) .. T(\alpha_n)\}$ is a linearly independent set with n elements. Since dim V = n, it follows that $\{T(\alpha_1), T(\alpha_2) .. T(\alpha_n)\}$ is a basis of V.

$\therefore (iv) \Rightarrow (v)$

Finally, we shall show that $(v) \Rightarrow (i)$.

Let us assume that (v) holds. Let $\{\alpha_1, \alpha_2 .. \alpha_n\}$ be a basis of V, then $\{T(\alpha_1), T(\alpha_2) .. T(\alpha_n)\}$ is another basis of V.

Let $\xi \in$ ket T and let $\xi = a_1\alpha_1 + a_2\alpha_2 + .. + a_n\alpha_n$, $a_i \in F$.

Since $\xi \in$ ket T, $T(\xi) = 0$

or, $T(a_1\alpha_1 + a_2\alpha_2 + .. + a_n\alpha_n) = 0$

or, $a_1T(\alpha_1) + a_2T(\alpha_2) + .. + a_nT(\alpha_n) = 0$

This implies that $a_1 = a_2 = .. = a_n = 0$

$\therefore \xi \in$ ket $T \Rightarrow \xi = 0$. Hence ker $T = \{0\}$

$\therefore T$ is one-one

Again dim ker T + dim Im T = dim V, gives

dim Im T = dim V.

But Im $T \subseteq V$ and dim Im T = dim V

So Im $T = V$

Hence T is onto.

$\therefore T$ is one-one and onto. Hence T is invertible and so nonsingular.

$\therefore (v) \Rightarrow (i)$.

Thus, all the statements are equivalent.

Matrox of the composite transformation :

Let $T : V \rightarrow W$ and $S : W \rightarrow U$ be linear transformations where V, W, U are finite dimensional vector spaces over a field F. Then relative to the choice of ordered bases matrix of the composite transformation ST

= matrix of the transforation S × matrix of the transformation T.

i. e. $m(ST) = m(S)\, m(T)$.

Proof : Let dim $V = m$, dim $W = n$ and dim $U = p$. let $\{\alpha_1, \alpha_2 .. \alpha_m\}$, $\{\beta_1, \beta_2 .. \beta_n\}$, $\{\gamma_1, \gamma_2 .. \gamma_p\}$ be respectively the ordered bases of V, W and U.

Relative to these bases let the transformation matrices be

$m(T) = (a_{ij})_{mxn}$, $m(S) = (b_{jk})_{nxp}$, $m(ST) = (c_{ik})_{mxp}$

Then $$T(\alpha_i) = \sum_{j=1}^{n} a_{ji} \beta_j, \qquad i = 1, 2 .. m$$

$$S(\beta_j) = \sum_{k=1}^{p} b_{kj} \gamma_k, \qquad j = 1, 2 .. n$$

$$ST(\alpha_i) = \sum_{k=1}^{p} c_{ki} \gamma_k, \qquad i = 1, 2 .. m$$

But $ST(\alpha_i) \;=\; S\left[T(\alpha_i)\right]$, since S and T are linear.

$$= S\left[\sum_{j=1}^{n} \alpha_{ji} \beta_j\right]$$

$$= \sum_{j=1}^{n} \alpha_{ji}\, S(\beta_j)$$

$$= \sum_{j=1}^{n} \alpha_{ji} \left(\sum_{k=1}^{p} b_{kj}\, \gamma_k\right)$$

$$= \sum_{k=1}^{p} \left(\sum_{j=1}^{n} a_{ji}\, b_{kj}\right) \gamma_k$$

$$\therefore\ c_{kj} = \sum_{j=1}^{n} b_{kj}\, a_{ji}$$

This shows that $m(ST) = m(S).\ m(T)$.

Theorem : Let V and W be finite dimensional vector spaces of dimension n over a field F. Let $T : U \rightarrow W$ be invertible. Then relative to a choice of ordered basis

$$m\left(T^{-1}\right) = \left[m(T)\right]^{-1}$$

Proof : Let $\{\alpha_1, \alpha_2 .. \alpha_n\}$ and $\{\beta_1, \beta_2 .. \beta_n\}$, be ordered bases of V and W respectively.

Let matrices of the transformations be

$m(T) = (a_{ij})_{nxn}$ and $m(T^{-1}) = (b_{jk})_{nxn}$.

Then $T(\alpha_i)\sum_{j=1}^{n} a_{ji}\, \beta_j$, $i = 1, 2 .. n$

Then $T^{-1}(\beta_j) = \sum_{i=1}^{n} b_{ij}\, \alpha_i$, $j = 1, 2 .. n$

Since T^{-1} is the inverse of the transformation T, we have $T^{-1}\, T = I_v$ and $TT^{-1} = I_w$; I_v and I_w being the identity transformations of V and W, respectively.

$$\begin{aligned}
\text{Now,}\quad T^{-1}T(\alpha_i) &= T^{-1}\left[T(\alpha_i)\right] \\
&= T^{-1}\left(\sum_{j=1}^{n} a_{ji}\, \beta_j\right) \\
&= \sum_{j=1}^{n} a_{ji}\, T^{-1}\left(\beta_j\right) \qquad \text{Q}\, T^{-1} \text{ is linear} \\
&= \sum_{j=1}^{n} a_{ji}\left(\sum_{i=1}^{n} b_{ij}\, \alpha_i\right) \\
&= \sum_{i=1}^{n}\left(\sum_{j=1}^{n} b_{ij}\, \alpha_{ji}\right)\alpha_i = \sum_{i=1}^{n} c_i\, \alpha_i
\end{aligned}$$

where $c_1 = \sum_{j=1}^{n} b_{ij}\, \alpha_{ji}$

Now, $T^{-1}T\,(\alpha_{ij}) = \alpha_i$, hence

$$\alpha_1 = \sum_{i=1}^{n} c_i \alpha_i$$

or, $$\sum \alpha_j \delta_{ij} = \sum_{j=1}^{n} c_j \alpha_j$$

or, $$\sum \left(c_j - \delta_{ij}\right) \alpha_j = 0$$

Since α_i's are linearly independent we have

$c_j = \delta_{ij}$, $\quad j = 1, 2, .. n.$

or, $c_i = \delta_{ij}$, $\quad i = 1, 2, .. n.$

or, $$\sum_{j=1}^{n} b_{ij}\, \alpha_{ji} = \delta_{ij}, \quad i = 1, 2, .. n.$$

This shows that $m(T^{-1})m(T) = I_n$

Similarly, $\quad TT^{-1} = I_w$ will give

$$m(T)m(T^{-1}) = I_n$$

$\therefore\ m(T)m(T^{-1}) = m(T^{-1})\, m(T) = I_n$

$\Rightarrow \quad m(T^{-1}) = [m(T)]^{-1}$

EXERCISES

1. Let $T : R^3 \rightarrow R^3$ difined by $T(x, y, z) = (x - y, x + xy, y + 3z)$, $(x, y, z) \in R^3$. Show that T is invertible and determine T^{-1}.

2. Let S and T be linear transformations of R^3 to R^3 defined by $S(x, y, z) = (z, y, x)$, $(x, y, z) \in R^3$ and $T(x, y, z) = (x + y + z, y + z, x)$, $(x, y, z) \in R^3$.

   (i) Determine TS and ST

   (ii) Prove that both S and T are invertible,

   Verify that $(ST)^{-1} = T^{-1}S^{-1}$

3. A linear mapping $T : R^3 \to R^3$ difined by

   $T(x, y, z) = (2x + y - z, y + 4z, x - y + 3z), (x, y, z) \in R^3$.

   Find the matrix of T relative to the order bases

   $(0, 1, 1), (1, 0, 1), (1, 1, 0)$ of R^3.

4. If k be a nonzero scalar, show that the linear transformation T is singular if and only if kT is sirtgular. Hence T is singular if and only if $-T$ is singular.

5. Let T be a linear operator on R^2 defined by $T(3, 1) = (2, -4)$ and $T(1, 1)= (0, 2)$. Find $T(a, b)$ and in particular find $T(7, 4)$.

6. Let F be a field and T be the operator on F^2 defined by $T(x_1, x_2) = (x_1, 0)$. Find the matrix of T with respect to the standard ordered basis $\{(1, 0), (0, 1)\}$ of F^2.

7. Let T be the linear transformation from R^3 into R^2 defined by

   $T(x_1, x_2, x_3) = (x_1 + x_2, 2x_3 - x_1)$

(a) If B is the standard ordered basis for R^3 and B' is the standard ordered basis for R^3, what is the matrix of T relative to the pair B, B'?

(b) If $B = \{\alpha_1, \alpha_2, \alpha_3\}$and $B' = \{\beta_1, \beta_2\}$where

   $\alpha_1 = (1, 0, -1), \alpha_2 = (1, 1, 1), \alpha_3 = (1, 0, 0), \beta_1 = (0, 1), \beta_2 = (1, 0)$

   what is the matrix of T relative to the pair B, B'?

8. Let T be a linear operator on R^3, the matrix of which in the standard ordered basis is

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

   Find a basis for the range of T and a basis for ker T.

9. Let T be a linear operator on R^3 defined by

   $T(x_1, x_2, x_3) = (3x_1 + x_3, -2x_1 + x_2, -x_1, + 2x_2 + 4x_3)$

(a) What is the matrix of T in the standard ordered basis for R^3?

(b) What is tha matrix of T in the ordered basis $\{\alpha_1, \alpha_2, \alpha_3\}$ where $\alpha_1 = (1, 0, 1), \alpha_2 = (-1, 2, 1)$ and $\alpha_3 = (2, 1, 1)$?

(c) Prove that T is invertible and give a rule for T^{-1} like the one which defines T.

Unit : 4 ❑ Reduction of Matrices to Diagonal/ Normal Form

Definition : Similar matrices—Let A and B be $n \times n$ matrices over the field F. We say that B is similar to A over F if there is an invertible $n \times n$ matrix P over F such that $B = P^{-1}AP$.

Definition : Congruent matrices—Let A and B be $n \times n$ matrices over the field F. A matrix A is said to be congruent to matrix B if there exists an invertible matrix P over F such that $B = P^TAP$.

EXAMPLES

Ex. 1. The matrix $\begin{bmatrix} 1 & 4 \\ 2 & 5 \end{bmatrix}$ is congruent to the matrix $\begin{bmatrix} 1 & 2 \\ 0 & -3 \end{bmatrix}$ because there exists and invertible matrix

$$P = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \text{ such that}$$

$$\begin{bmatrix} 1 & 2 \\ 0 & -3 \end{bmatrix} = P^T \begin{bmatrix} 1 & 4 \\ 2 & 5 \end{bmatrix} P$$

Ex. 2. The matrix $\begin{bmatrix} 1 & 2 \\ -1 & 4 \end{bmatrix}$ is similar to the matrix $\begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$ because there exists an invertible matrix $P = \begin{bmatrix} 2 & -3 \\ 1 & -3 \end{bmatrix}$ such that

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

Theorem : Similarity of matrices is an equivalence relation over the set of $n \times n$ matrices over F.

Proof : Let A, B, C, be $n \times n$ matrices over F. Let R be the relation of similarity between matrices. That is, ARB if there exists an invertible matrix P such that

$$B = P^{-1}AP$$

Since, $IA = AI$, I being the $n \times n$ identity matrix, it follows that

$$A = I^{-1}AI, \text{ for every } n \times n \text{ matrices } A$$

Hence ARA, $\forall$A in the set of $n \times n$ matrices

So the relation R is reflexive.

Now, suppose that ARB holds. Then there exists an invertible matrix P such that

$B = P^{-1}AP$

or, $PB = AP$

or, $A = PBP^{-1} = (P^{-1})^{-1}\, BP^{-1}$

This shows that BRA holds

Hence R is symmetric

Finally, suppose that ARB and BRC hold, then there exist invertible matrices P and Q such that

$$A = P^{-1}BP \text{ and } B = Q^{-1}CQ$$

Then $A = P^{-1}(Q^{-1}CQ)P = (QP)^{-1}\, C(QP)$

Since QP is invertible, it follows that

ARC holds

Hence R is transitive

So R is an equivalence relation.

Theorem : Let V be an n dimensional vector space over the field F and let B and B' be two ordered bases of V. Then there is a unique, necessarily invertible $n \times n$ matrix P with entries in F such that

(i) $[\alpha]_B = P\,[\alpha]_{B'}$

(ii) $[\alpha]_{B'} = P^{-1}\,[\alpha]_B$

for every vector α in V; $[\alpha]_B$ being the coordinate matrix of the vector α relative to the ordered basis B.

Proof : Let $B = \{\alpha_1, \ldots\ldots\ldots\ldots \alpha_n\}$ and $B' = \{\alpha_1', \ldots\ldots\ldots\ldots \alpha'_n\}$be tow ordered bases for V. There are unique scalars P_{ij}, such that

$$\alpha'_j = \sum_{i=1}^{n} P_{ij}\alpha_i, 1 \leq j \leq n \qquad \ldots\ldots\ldots (1)$$

Let $x'_1 x'_2, \ldots\ldots\ldots\ldots\ldots x'_n$, be the coordinates of a given ventor α in the ordered basis B'. Then.

$$\alpha = x'_1\alpha'_2, + \ldots\ldots\ldots\ldots + x'_n\alpha'_n,$$

$$= \sum_{j=1}^{n} x'_j \alpha = \sum_{j=1}^{n} x'_j \left(\sum_{i=1}^{n} P_{ij} \alpha_i \right)$$

or, $$\alpha = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} P_{ij} x'_j \right) \alpha_i$$

or, $$\sum_{i=1}^{n} x_i \alpha_i = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} P_{ij} x'_j \right) \alpha_i \qquad \text{......... (2)}$$

Since the coordinates x_1, x_2, ... x_n of α in the ordered basis B are uniquely determined, it follows from (2) that

$$x_i = \sum_{j=1}^{n} P_{ij} x'_j, 1 \leq i \leq n \qquad \text{......... (3)}$$

Let P ba the $n \times n$ matrix whose (i, j)th entry is P_{ij}, and let X and X' be the coordinate matrices of the vector α in the ordered basis B and B'. Then we may write form (3)

$$X = PX' \qquad \text{......... (4)}$$

Since B and B' are linearly independent sets, $X = 0$ if and only is $X' = 0$. Hence from (4), it follows that P is invertible, and so

$$X' = P^{-1}X \qquad \text{......... (5)}$$

Thus in terms of the coordinate martix of a vector relative to an ordered basis, (4) and (5) may be expressed as

$[\alpha]_B = P\,[\alpha]_{B'}$

$[\alpha]_{B'} = P^{-1}\,[\alpha]_B$

Theorem : Let V be a finite-dimensional vector space over the field F and let

$$B = \{\alpha_1, \alpha_2, \ldots\ldots \alpha_n\} \text{ and } B' = \{\alpha'_1, \alpha'_2 \ldots\ldots \alpha'_n\}$$

be ordered bases for V. Suppose T is a linear operator on V. If $P = [P_1, P_2, \ldots..P_n]$ is the $n \times n$ matrix with columns $P_j = [\alpha'_j]_B$, then

$$[T]_{B'} = P^{-1}\,[T]_B P$$

Proof : In the theorem just proved, we heve

$$[\alpha]_B = P^{-1}\,[\alpha]_{B'} \qquad \text{......(1)}$$

for every vector α in V. Here P is the unique invertible $n \times n$ matrix. If P_j is the jth column of the matrix P then

$$P = [P_1, P_2, ...P_n], \text{ where } P_j = [\alpha'_j]_B$$

By definition $[T[\alpha]_B = [T]_B [\alpha]_B$

Applying (1) to the vector $T(\alpha)$ we have

$$[T[\alpha]_B = P[T[\alpha]_{B'}$$

Combining (1), (2) and (3) we get

$$[T]_B P[\alpha]_{B'} = P[T(\alpha)]_{B'}$$

or, $\quad P^{-1}[T]_B \, P[\alpha]_{B'} = [T(\alpha)]_{B'}$

Hence $\quad [T]_{B'} = P^{-1} [T]_B P$

Ex. 3. Let T be a linear operator on R^2 defined by $T(x_1, x_2) = (x_1, 0)$. It can be shown that the matrix T in the standard ordered basis $B = \{(1, 0); (0, 1)\}$ is

$$[\mathrm{T}]_B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

Suppose B' is the ordered basis for R^2 consisting of the vectors (1, 1) and (2, 1), then

$$(1, 1) = (1, 0) + (0, 1)$$
$$(2, 1) = 2(1, 0) + (0, 1)$$

so that the matrix P is $\begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}$

$$\therefore \quad P^{-1} = \begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix}$$

Thus, $[T]_{B'} = P^{-1} [T]_B P$

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

Definition : An $n \times n$ matrix A is said to be orthogonally congruent (or orthogonally similar) to a matrix B if there exists an othogonal matrix P such that

$$B = P^T AP = P^{-1}AP$$

EXAMPLES

Ex. 1. The matrix $\begin{bmatrix}1 & 2\\ 2 & 1\end{bmatrix}$ is orthogonally congruent to $\begin{bmatrix}-1 & 0\\ 0 & 3\end{bmatrix}$ because there exists an orthogonal marix

$$P=\begin{bmatrix}\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\end{bmatrix} \text{ such that}$$

$$\begin{bmatrix}-1 & 0\\ 0 & 3\end{bmatrix}=P^T\begin{bmatrix}1 & 2\\ 2 & 1\end{bmatrix}P=P^{-1}\begin{bmatrix}1 & 2\\ 2 & 1\end{bmatrix}P$$

Characteristic equation : Let A be an $n \times n$ matrix over a field F. Then det $(A - xI_n)$ is said to be the characteristic polynomial of A and is denoted by $\psi_A(x)$, The equation $\psi_A(x) = 0$ is said to be the **characteristic equation if A.**

Let $A = [a_{ij}]_{n\times n}$, the

$$\psi_A(x) = \begin{bmatrix}a_{11}-x & a_{12} & a_{1n}\\ \cdots\cdots & \cdots\cdots & \cdots\cdots\\ a_{21} & a_{22}-x & a_{2n}\\ a_{n1} & a_{n2} & a_{mn-x}\end{bmatrix}$$

$$= C_0x^n + C_0x^{n-1} + \ldots.. + C_n$$

where $C_0 = (-1)^n$,

$C_r = (-1)^{n-r}$ [sum of the principal minors of A of order r]

$\therefore$ $C_1 = (-1)^{n-1}\{a_{11} + a_{22} + \ldots\ldots + a_{nn}\}$

C_n = det A.

Theorem : (Cayley Hamilton theorem) Every square matrix satisfies its own characteristic equation.

The is, if A is a square matrix of order n, and if $C_0x^n + C_1x^{n-1} + \ldots.. + C_n = 0$ be the characteristic equation for the matrix A, then the theorem states that

$$C_0A^n + C_1A^{n-1} + \ldots.. + C_nI_n = 0$$

Proof : We have

$$\psi_A(x) = |A\ xI_n| = C_0x^n + C_1x^{n-1} + \ldots.. + C_n =$$

Now, let $B(x)$ denote the adjoint of the martix $A - xI_n$. The elements of $B(x)$ are confactors of the matrix $A - xI_n$ and hence are polynomials in x of degree not exceeding $(n - 1)$. Hence

$$B(x) = B_0x^{n-1} + B_1x^{n-2} + ... + B_{n-1}$$

where each $B_i$is and $n \times n$ matrix

Now, $(A - xI_n)\, B(x) = \det\,(A - xI_n).\ I_n$ gives

$$(A - xI_n)\,(B_0x^{n-1} + B_1x^{n-2} + ... + B_{n-1})$$
$$= (C_0x^n + C_1x^{n-1} + ... + C_n)I_n$$

or, $A(B_0x^{n-1} + B_1x^{n-2} + ... + B_{n-1})$

$$-(B_0x^n + B_1x^{n-1} + ... + B_{n-1}x)$$
$$= C_0I_nx^n + C_1I_nx^{n-1} + ... + C_nI_n$$

Equating coeffients of like provers of x from both sides we get

$$-B_0 = C_0I_n$$
$$AB_0 - B_1 = C_1I_n$$
$$AB_1 - B_2 = C_2I_n$$
$$AB_{n-2} - B_{n-1} = C_{n-1}I_n$$
$$AB_{n-1} = C_nI_n$$

Multiplying the above relations by $A^n, A^{n-1}, ... A_n, I_n$ respectively and adding we get

$$C_0A^n + C_1A^{n-1} + ... + C_{n-1}A + C_nI_n = 0$$

i.e. $\psi_A(A) = 0$

Hence the theorem is proved.

Ex. 2. Verify Cayley Hamilton's theorem for the matrix A.

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

and hence compute A^{n-1}.

We have the characteristic equation $\psi_A(x) = |A - xI_3| = 0$

$$\text{or,} \quad \begin{pmatrix} 1-x & 0 & 0 \\ 1 & 2-x & 1 \\ 2 & 3 & 2-x \end{pmatrix} = 0$$

$$\Rightarrow \quad (1 - x)\,[4 - 4x + x^2 - 3] = 0$$

$$\Rightarrow \quad x^3 - 5x^2 + 5x - 1 = 0$$

We have $A^2 = \begin{pmatrix} 1 & 0 & 0 \\ 5 & 7 & 4 \\ 9 & 12 & 7 \end{pmatrix}$, $A^3 = \begin{pmatrix} 1 & 0 & 0 \\ 20 & 26 & 15 \\ 35 & 45 & 26 \end{pmatrix}$

So, $A^3 - 5A^2 + 5A - I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 20 & 26 & 15 \\ 35 & 45 & 26 \end{pmatrix} - \begin{pmatrix} 5 & 0 & 0 \\ 25 & 35 & 20 \\ 45 & 60 & 35 \end{pmatrix} + \begin{pmatrix} 5 & 0 & 0 \\ 5 & 10 & 5 \\ 10 & 15 & 10 \end{pmatrix} - \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

$$= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = 0$$

This shows that the matrix A satisfies its characteristic equation. Thus Cayley Hamilton's theorem is verified.

We have $A(A^2 - 5A + 5I_3) = I_3$

Hence $\quad A^{-1} = A^2 - 5A + 5I_3$

$$= \begin{pmatrix} 1 & 0 & 0 \\ 5 & 7 & 4 \\ 9 & 12 & 7 \end{pmatrix} - \begin{pmatrix} 5 & 0 & 0 \\ 5 & 10 & 5 \\ 10 & 15 & 10 \end{pmatrix} + \begin{pmatrix} 5 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 5 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & -1 \\ -1 & -3 & 2 \end{pmatrix}$$

Characteristic value or eigen value of a matrix.

A root of the characteristic equation $\psi_A(x) = 0$ of a square matrix is called a characteristic value or an eigen value of A.

Clearly for a singular matrix A (i.e for the square matrix A for which det $A = 0$), 0 is an eigen value of A. For, the characteristic equation is

$$\psi_A(x) = (C_0x^n + C_1x^{n-1} + \ldots.. + C_n = 0, \text{ where } C_n = \det A.$$

If det $A = 0$, then clearly $x = 0$ is a root of this equation.

Again, if the square maxtrix A is a diagonal matrix such that

$$A = \begin{pmatrix} a_1 & 0 & 0 \ldots\ldots 0 \\ 0 & a_2 & 0 \ldots\ldots 0 \\ \ldots\ldots & \ldots\ldots & \ldots\ldots \\ 0 & 0 & \ldots\ldots a_n \end{pmatrix}$$

Then characteristic equation is

$$\begin{vmatrix} a_1 - x & 0 \ldots\ldots\ldots 0 \\ 0 & a_2 - x \ldots\ldots 0 \\ \ldots\ldots\ldots & \ldots\ldots\ldots \\ 0 \ldots\ldots 0 & \ldots\ldots\ldots a_n - x \end{vmatrix}$$

Clearly roots of this equation are $\alpha_1, \alpha_2, \ldots\ldots \alpha_n$. Hence for a diagonal matrix, the eigen values are the diagonal elements.

Theorem : If A and B are similar matrices then A and B have the same characteristic polynomial and hence the same characteristic values.

Proof : Since A and B are similar, there exists an invertible matrix P such that

$$A = P^{-1}BP$$

$$\therefore \quad A - xI = P^{-1}BP - xI = P^{-1}BP - xP^{-1} = P^{-1}(B - xI)P$$

$$\therefore \quad \det(A - xI) = (\det P^{-1}) \det(B - xI) (\det P)$$

$$= \det(B - xI)$$

This shows that matrices A and B have the same characteristic polynomial and hence the same characteristic values.

Theorem : If λ be an eigen value of a non-singular matrix A, then λ^{-1} is an eigen value of A^{-1}.

Proof : Since A is non singular, $\lambda \neq 0$

Now, $\det\left(A^{-1} - \lambda^{-1}I\right) = \frac{1}{\lambda^n} \det\left(\lambda A^{-1} - I\right)$

$$= \frac{1}{\lambda^n} \cdot \frac{\det A \det\left(\lambda A^{-1} - I\right)}{\det A}, \det A \neq 0$$

$$= \frac{1}{\lambda^n} \frac{\det\left\{A\det\left(\lambda A^{-1} - I_n\right)\right\}}{\det A} = \frac{1}{\lambda^n} \cdot \frac{\det \lambda I - A}{\det A} = 0 \text{ , since } \lambda \text{ is an eigenvalue of } A.$$

This shows that λ^{-1} is an eigenvalue of A^{-1}.

Definition : If A is a square matrix of order n, then the values of λ for which the equation.

$$AX = \lambda X \text{ .... } (1)$$

has nontrivial solutions are called the eigen or characteristic values of A. If λ is an eigenvalue, then the nonzero vector X for which the equation (1) holds, is called the eigen vector or characteristic vector corresponding to the eigen value λ.

Theorem : To an eigen vector of A there corresponds a unique eigen value of A.

Proof : Let λ_1 and λ_2 be the eigen values corresponding to an eigenvector X of A.

Then $AX = \lambda_1 X$ and also $AX = \lambda_2 X$

$\Rightarrow \quad (\lambda_1 - \lambda_2)X = 0$

But X is non zero, hence $\lambda_1 = \lambda_2$

Theorem : Two eigenvectors of a square matrix A corresponding to two distinct eigen values of A are linearly independent.

Proof : Let X_1, X_2 be two eigen vectors of A corresponding to two distinct eigen values λ_1, λ_2 respectively.

Then $AX_1 = \lambda_1 X_1$ and $AX_2 = \lambda_2 X_2$

Let A be the square matrix over F. We have to show that

$$c_1X_1 + c_2X_2 = 0$$

will imply $c_1 = 0$ and $c_2 = 0$ $\quad c_1, c_2 \in F$

Now $\quad c_1X_1 + c_2X_2 = 0 \Rightarrow c_1AX_1 + c_2AX_2 = 0$

$\Rightarrow \quad c_1\lambda_1X_1 + c_2\lambda_2X_2 = 0$

$\Rightarrow \quad \lambda_1(-c_2X_2) + c_2\lambda_2X_2 = 0$

$\Rightarrow \quad c_2(\lambda_1 - \lambda_2)X_2 = 0$

Since $X_2 \neq 0$ and $\lambda_1 \neq \lambda_2$, we have $c_2 = 0$

$\therefore \; c_1 = 0$. Hence X_1, X_2 are linearly independent

Theorem : Let A be a square matutix of order n having K distinct eigenvalues λ_1, $\lambda_2 \ldots\ldots \lambda_k$. Let X_i be an eigenvector corresponding to the eigenvalue λ_i, $i = 1, 2 \ldots. K$. Then the set $\{X_1, X_2 \ldots\ldots X_k\}$ is linearly independent.

Proof : We shall prove this theorem by induction. We have seen in the theorem just proved above that if λ_1, λ_2 are distinct eigenvalues of the matrix A, then X_1, X_2 are linearly independent. Thus the statement of the present theorem is true for $K = 2$. Let us suppose that the statement be true for K = r, a positive integer.

So we assume that the eigenvectors $X_1, X_2 \ldots\ldots X_r$ corresponding to eigen value $\lambda_1, \lambda_2 \ldots\ldots \lambda_r$ are linearly independent. Let λ_{r+1} be distinct from $\lambda_1, \lambda_2 \ldots\ldots \lambda_r$. Let X_{r+1} be the eigen vector corresponding to λ_{n+1}. We shall prove that $X_1, X_2 \ldots\ldots X_{r+1}$ are linearly independent.

Let $C_1, C_2 \ldots\ldots C_{r+1}$ be scalars such that

$$C_1X_1 + C_2X_2 + \ldots\ldots + C_rX_r + C_{r+1}X_{r+1} = 0 \ldots\ldots (1)$$

or, $A(C_1X_1 + C_2X_2 + \ldots\ldots + C_{r+1}X_{r+1}) = 0$

or, $C_1\lambda_1X_1 + C_2\lambda_2X_2 + \ldots\ldots + C_{r+1}\lambda_{r+1}X_{r+1} = 0 \ldots\ldots\ldots\ldots (2)$

Multiplying (1) by λ_{r+1} and subracting the result from (2) we get

$$C_1(\lambda_1 - \lambda_{r+1})X_1 + C_2(\lambda_2 - \lambda_{r+1})X_2 + \ldots\ldots + C_r(\lambda_r - \lambda_{r+1})X_r = 0$$

Since λ_i's ($i = 1, 2 \ldots\ldots\ldots r + 1$) are distinct and $X_1, X_2 \ldots\ldots X_r$ are linearly independent, it follows that

$$C_1 = C_2 = .. = C_r = 0$$

Then from (1) we get $C_{r+1} = 0$, since $X_{r+1} \neq 0$

Hence $X_1, X_2 \ldots\ldots X_{r+1}$ are linearly independent

Thus, as the results actually holds for $r = 2$, by induction the result is true for $r =$ 3, 4 .... K.

Theorem : The eigen vectors of an $n \times n$ matrix A over a field F corresponding to an eigen value $\lambda \in F$, together with the null vector, form a vector space, a subspace of $V_n(F)$.

Proof : Let S be the set of all eigen vectors of A corresponding to a eigen value λ of A.

Let $X_1, X_2 \in S$

Then $AX_1 = \lambda X_1$ and $AX_2 = \lambda X_2$

$\therefore \quad A(X_1 + X_2) = \lambda(X_1 + X_2)$

This show that $X_1 + X_2$ is an eigenvector of A corresponding to the eigenvalue λ. Hence $X_1 + X_2 \in S$. Let $c\ (\neq 0) \in F$, then $AX_1 = \lambda X_1 \Rightarrow A(cX_1) = \lambda(cX_1)$

$\Rightarrow cX_1 \in S$ for $c \in F$

Thus the eigen vectors corresponding to λ together with the null vector form a vector space which is clearly a subspace of $V_n(F)$.

EXAMPLES

Ex. 1. Find the eigen values and eigen vectors of the matrix A, where

$$A = \begin{pmatrix} 1 & 0 & -2 \\ 0 & 0 & 0 \\ -2 & 0 & 4 \end{pmatrix}$$

The characteristic equation of A is

$$0 = \det(A - xI) = \begin{vmatrix} 1-x & 0 & -2 \\ 0 & -x & 0 \\ -2 & 0 & 4-x \end{vmatrix}$$

or, $\quad -x^3 + 5x^2 = 0$

Hence the eigenvalues are $\lambda_1 = 0, \lambda_2 = 0, \lambda_3 = 5$

Let $X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$, then $AX = \lambda X$

$$\Rightarrow \begin{pmatrix} 1 & 0 & -2 \\ 0 & 0 & 0 \\ -2 & 0 & 4 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \lambda \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

$$\Rightarrow x_1 - 2x_3 = \lambda x_1$$
$$0 = \lambda x_2$$
$$-2x_1 + 4x_3 = \lambda x_3$$

For $\lambda = 0$ these reduce to

$$x_1 = 2x_3$$

Hence any vector of the form $(2c, b, c)$ is an eigen vector. We have to choose two vectors of this form such that they are linearly independent. Clearly we may take

$$X_1 = (2, 0, 1)$$
$$X_2 = (0, 1, 0)$$

For $\lambda_3 = 5$ the conditons reduce to

$$-2x_1 = x_3$$
$$x_2 = 0$$

Hence any vector of the form $(a, 0, -2a)$ is an eigen vector, As a simple vector of this form choose

$$x_3 = (1, 0, -2)$$

as a characteristic vector assouiated with $\lambda_3 = 5$.

Theorem : The eigen values of a real symmetric matrix are all real.

Proof : Let A be an $n \times n$ real symmeetric matrix. The characteristic equation of A is an equation in x with real coefficients. Let λ be an eigen value of A.

Then $AX = \lambda X$, X is non-null

$\therefore$ $\left(\overline{AX}\right) = \overline{\lambda X}$, where the bar overhead denotes complex conjugate

Taking transpose,

$$\left(\overline{AX}\right)^T = \left(\overline{\lambda X}\right)^T$$

$$\Rightarrow \quad \overline{X}^T A^T = \overline{\lambda X}^T \quad \ldots\ldots (1)$$

But since A is real and symmetric, $\overline{A}^T = A$

$$\therefore \quad \overline{X}^T A = \overline{\lambda X}^T$$

Multiplying by X from right we get

$$\overline{X}^T AX = \overline{\lambda X}^T X$$

$$\text{or,} \quad \overline{X}^T \lambda X = \overline{\lambda X}^T X$$

$$\text{or,} \quad \overline{X}^T X\left(\lambda - \overline{\lambda}\right) = 0$$

Since X is non-null, $\overline{X}^T X \neq 0$, hence $\lambda = \overline{\lambda}$

$\therefore$ λ is real

Hence the theorem is proved.

We note futher if the matrix A is real but skew symmetric, then $\overline{A}^T = -A$ and hence from equation (1) we get

$$\overline{X}^T A = \overline{-\lambda X}^T$$

Right multiplication by X yields

$$\overline{X}^T AX = \overline{-\lambda X}^T X$$

or, $\overline{X}^T \lambda X = \overline{-\lambda X}^T X$

or, $\left(\lambda + \overline{\lambda}\right) = \overline{X}^T X = 0$

Since $\overline{X}^T X \neq 0, \quad \lambda + \overline{\lambda} = 0$

Hence λ is purely imaginary

Thus the eigen values of real skew symmetric matrix are purely imaginary.

Theorem : The eigenvector corresponding to two distinct eigen values of a real symmetric matrix are orthogonal.

Proof : Let X_1, X_2 be two distinct eigenvectors corresponding to distinct eigenvalues λ_1, λ_2 of a real symmetric matrix A. Then λ_1, λ_2 are real and

$$AX_1 = \lambda_1 X_1, \; AX_2 = \lambda_2 X_2$$

$\therefore \quad (AX_1)^T = (\lambda_1 X_1)^T \Rightarrow X_1^T A^T = \lambda_1 X_1^T$

But $A^T = A$, since A is symmetric.

$\therefore \quad X_1^T A = \lambda_1 X_1^T$

$\therefore \quad X_1^T AX_2 = \lambda_1 X_1^T X_2$

or, $\quad X_1^T \lambda_2 X_2 = \lambda_1 X_1^T X_2$

or, $\quad (\lambda_2 - \lambda_1) X_1^T X_2 = 0$

Since $\quad \lambda_1 \neq \lambda_2$, we have $X_1^T X_2 = 0$

Hence, $\quad X_1, X_2$ are orthogonal

Theorem Each eigen value of real othogonal matrix has unit modulus.

Proof : Let X be an eigen ovetor corresponding to an eigen value λ of a real othogonal matrix

Then $\quad AX = \lambda X$

$$\left(\overline{AX}\right)^T = \left(\overline{\lambda X}\right)^T$$

$$\Rightarrow \quad \overline{X}^T \overline{A}^T = \overline{\lambda X}^T$$

$$\Rightarrow \quad \overline{X}^T A^T = \overline{\lambda X}^T, \quad \text{Q} \quad A \text{ is real}$$

$$\text{or,} \quad \left(\overline{X}A^T\right)(AX) = \overline{\lambda X}(AX)$$

$$\text{or,} \quad \overline{X}\left(A^T A\right)X = \overline{\lambda X}\lambda X$$

$$\text{or,} \quad \overline{X}X = \overline{\lambda}\lambda\overline{X}X, \quad \text{Q} \quad A^T A = I$$

$$\text{or,} \quad \overline{X}X = \left(\overline{\lambda}\lambda - 1\right) = 0$$

since $\overline{X}X \neq 0$, $\lambda\overline{\lambda} = 1$ or, $|\lambda| = 1$

This proves the theorem.

Diagonalisation of matrices :

An $n \times n$ matrix A is said to be diagonalisable if A is similar to an $n \times n$ diagonal matrix.

If A is similar to a diagonal matrix $D = \text{diag}\ (\lambda_1, \lambda_2 \ldots\ldots\ldots \lambda_n)$ then $\lambda_1, \lambda_2 \ldots\ldots\ldots \lambda_n$ are the eigen values of A.

Diagonalisation of a real symmetric matrices :

Theorem : Let A be an $(n \times n)$ real symmetric matinx with distinct eigenvalues λ_1, $\lambda_2 \ldots\ldots\ldots \lambda_n$. Let X_i be the normalised column eigenvector correspondting to the eigen value λ_i. Then the matrix

$P = [X_1, X_2 \ldots\ldots\ldots X_n]$

is such that $P^T AP = \begin{bmatrix} \lambda_1 & 0 \ldots\ldots 0 \\ 0 & \lambda_2 \ldots\ldots 0 \\ \ldots\ldots & \ldots\ldots \\ 0 & 0 \ldots\ldots \lambda_n \end{bmatrix}$

Proof : We have

$$\begin{aligned} AP &= A[X_1, X_2, \ldots\ldots\ldots X_n] \\ &= [AX_1, AX_2, \ldots\ldots\ldots AX_n] \\ &= [\lambda_1 X_1, \lambda_2 X_2, \ldots\ldots\ldots \lambda_n X_n] \\ &= [X_1\lambda_1, X_2\lambda_2, \ldots\ldots\ldots X_n\lambda_n] \end{aligned}$$

$$= \quad [X_1, X_2, \ldots\ldots\ldots\ldots X_n \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_n \end{bmatrix}$$

$$= \quad P\begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_n \end{bmatrix}$$

Since the eigen vectors corresponding to distinct eigen values of a real symmetric matrix are orthogonal and since the eigen vectors are normalised it follows that the matrix P is orthogonal matrix ; hence

$$P^{-1}AP = P^{T}AP = \begin{bmatrix} \lambda_1 & 0\ldots\ldots0 \\ 0 & \lambda_2\ldots.0 \\ \ldots\ldots\ldots\ldots & \\ 0 & 0\ldots\ldots\lambda_n \end{bmatrix}$$

Ex. 2. Diagonalise the matrix

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

The characteristic equation is $|A\,\lambda I| = 0$

$$\Rightarrow \quad \begin{vmatrix} 2-\lambda & \sqrt{2} \\ \sqrt{2} & 1-\lambda \end{vmatrix} = 0 \Rightarrow \lambda^2 - 3\lambda = 0$$

$\therefore$ Eigenvalues are $\lambda_1 = 0$, $\lambda_2 = 3$

Corresponding to $\lambda_1 = 0$, we have

$$2x_1 + \sqrt{2}x_3 = 0$$

$$\sqrt{2}x_1 + x_2 = 0$$

$$\Rightarrow \quad x_1 = -\frac{1}{\sqrt{2}}x_2$$

If we wish to find an eigen vector of unit length (normalised), we must require that $x_1^2 + x_2^2 = 1$, This gives

$$\frac{3}{2}x_2^2 = 1$$

Choosing the positive square root we obtain

$$x_2 = \sqrt{\frac{2}{3}}, x_1 = -\frac{1}{\sqrt{3}}$$

and the eigen vector of unit length corresponding to λ_1 is

$$X_1 = \left(-\frac{1}{\sqrt{3}}, \sqrt{\frac{2}{3}}\right)^T$$

Proceeding similarly, corresponding to λ_2, the linearly independent eigen vector (normalised) will be

$$X_2 = \left(\sqrt{\frac{2}{3}}, \frac{1}{\sqrt{3}}\right)^T$$

We observe here that $X_1^T X_2 = 0$, X_1, X_2 are orthogonal

$$\text{We have P} = \begin{pmatrix} -\frac{1}{\sqrt{3}} & \sqrt{\frac{2}{3}} \\ \sqrt{\frac{2}{3}} & \frac{1}{\sqrt{3}} \end{pmatrix} = \frac{1}{\sqrt{3}}\begin{pmatrix} -1 & \sqrt{2} \\ \sqrt{2} & 1 \end{pmatrix}$$

$$P^T AP = \frac{1}{3}\begin{bmatrix} -1 & \sqrt{2} \\ \sqrt{2} & 1 \end{bmatrix}\begin{bmatrix} 2 & \sqrt{2} \\ \sqrt{2} & 1 \end{bmatrix}\begin{bmatrix} -1 & \sqrt{2} \\ \sqrt{2} & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$

Next we shall show that if an nth order matrix A has n linearly independent eigenvetors, then there exists a similarity transformation which diagonalises A.

If fact, if $P = [X_1, X_2, \ldots\ldots\ldots X_n]$ is a matrix whose columms are a set of n linearly independent eigenvectors, $P^{-1}AP$ is a diagonal matrix whose diagonal elements are the eigenvalues of A.

To prove this let X_i be an eigenvector with eigenvalue λ_i (not all λ_i are necessarily different). Also we write $D = \begin{pmatrix} \lambda_1 & 0 & 0 \ldots 0 \\ 0 & \lambda_2 & 0 \ldots 0 \\ 0 & 0 & \ldots \lambda_n \end{pmatrix}$

$$\text{Then } PD = (X_1, X_2, \ldots\ldots\ldots X_n)\begin{pmatrix} \lambda_1 & 0 & 0 \ldots 0 \\ 0 & \lambda_2 & 0 \ldots 0 \\ 0 & 0 & \ldots \lambda_n \end{pmatrix}$$

$$= (\lambda_1 X_1, \lambda_2 X_2, \ldots\ldots\ldots \lambda_n X_n)$$

and $AP = (AX_1, AX_2, \ldots\ldots AX_n) = (\lambda_1 X_1, \lambda_2 X_2, \ldots\ldots\ldots \lambda_n X_n)$

Hence $AP = PD$

or, $D = P^{-1}AP$

Thus the matrix A is similar to the diagonal matrix D whose diagonal elements are the eigen values of A.

If A is not symmetric, the matrix P is not in general an orthogonal matrix.

The above result shows that if the eigenvalues of A are all different, A can always be diagonalised by a similarity transformation. If the eigen values of A are not all different, A can be diagonalised if it has n linearly independent eigenvectors. If a does not have n linearly independent eigenvectors, A cannot be diagonalised by a similarlity transformation. However, by a similarity transoformation, any square matrix A can always be converted into a matrix with the following properties :

(1) All elements below the main diagonal vanish.

(2) The elements on the main diagonal are the eigenvalues of A, equal eigenvalues appear in adjacent positions on the diagonal.

(3) The only elements above the main diagonal which do not vanish are those column index j equal to $i + 1$, where i is the row index. Any such nonvanishing element has the value unity. Howeve, it can have the value unity only if the diagonal elements in positions i and $i + 1$ are equal.

The a 5th order nonsymmetric matrix with $\lambda_1 = \lambda_2 = \lambda_3, \lambda_4 = \lambda_5$, could to reduced to the unique form

$$\begin{bmatrix} \lambda_1 & \beta_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & \beta_2 & 0 & 0 \\ 0 & 0 & \lambda_3 & 0 & 0 \\ 0 & 0 & 0 & \lambda_4 & \beta_3 \\ 0 & 0 & 0 & 0 & \lambda_5 \end{bmatrix}$$

where the value of the β_i is either 0 or 1. This is called Jordan canonical form for the matrix A.

Elementary operations :

An elementary operation on a matrix A over a field F is an operation of the following three types :

1. Interchange of any two rows (or columns) of A

2. Multiplication of a row (or column) by a non zero scalar $c \in F$.

3. Addition of a scalar muliple of one row (or column) to another row (or column)

We call elementary operation an elementary row operation when applied to row and similarly an elementary column operation when applied to column.

The operation of interchange of ith & jth rows will be dentoed by R_{ij}.

Addition of c times of jth row to ith row will be denoted by $R_i + cR_j$

Consider for example, the matrix

$$A = \begin{bmatrix} 5 & 2 \\ 4 & 6 \\ 3 & 1 \end{bmatrix}$$

Applying R_{23} on a we get the matrix

$$B = \begin{bmatrix} 5 & 2 \\ 3 & 1 \\ 4 & 6 \end{bmatrix}$$

Applying $2R_2$ on B we get

$$C = \begin{bmatrix} 5 & 2 \\ 6 & 2 \\ 4 & 6 \end{bmatrix}$$

Finally applying $-2R_1 + R_3$ we get

$$D = \begin{bmatrix} 5 & 2 \\ 6 & 2 \\ -6 & 2 \end{bmatrix}$$

Thus we have

$$A \xrightarrow{R_2} B \xrightarrow{2R_2} C \xrightarrow{-2R_1+R_3} D$$

Let S be the set of the $m \times n$ matrices over a field F. A matrix $B \in S$ is said to be row (column) equivalent to a matrix $A \in S$ if B can be obtained by successive application of a finite number of elementary row (column) operations on A.

Thus in the above example, the matrix D is row equivalent to the matrix A.

Definition : An $m \times n$ matrix is called **row** reduced if

(a) the first nonzero element in a non zero row is 1 (called the leading 1), and

(b) each column containing the leading 1 of some row has all other elements zero.

Thus the matrices below

$$\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix},\begin{pmatrix}0&1&4&0\\1&0&1&0\\0&0&0&1\end{pmatrix},\begin{pmatrix}0&1&0&0\\0&0&1&2\\1&0&0&1\\0&0&0&0\end{pmatrix}$$

are row reduced

Definition : An $m \times n$ matrix A is said to be a row reduced echelon matrix if

(a) A is row reduced

(b) there is an integer r $(0 \leq r \leq m)$ such that the first r rows of A are nonzero rows and the remaining rows are all zero rows

(c) for each nonzero row, if the leading element of the row i occurs in column k_i, then

$$k_1 < k_2, < .... < k_r.$$

The matrix $\begin{pmatrix}0&1&1&0&1\\0&0&0&1&0\\0&0&0&0&0\end{pmatrix}$ is row reduced echelon matrix.

Definition : An elementary row (column) transformation matrix is a matrix that can be obtained from an identify matrix I_n by a single elementary row (column) operaton. A matrix that is either an elementary row transformation matrix or an elementary column transformation matrix is called an elementary matrix.

Form the above definition it is obvious that elementary matices are of three types, of which the first type, $\begin{bmatrix}1&0&0\\0&0&1\\0&1&0\end{bmatrix}$ the second type, $\begin{bmatrix}1&0&0\\0&c&0\\0&0&1\end{bmatrix}$ where $c \neq 0$ and the third type, $\begin{bmatrix}1&0&0\\0&1&0\\0&c&1\end{bmatrix}$ are third order examples.

Theorem : The effect of the elementary row opeations R_{ij}, cR_i $(c \neq 0)$ and $R_i + cR_j$ on an $m \times n$ matrix A is obtained by premultiplying A by E_{ij}, $E_i(c)$ and $E_{ij}(c)$ respectively, where each elementary matrix is of order m and E_{ij}, E_i (c), E_{ij} (c) are the three elementry operations on the identity matrix of order m.

Proof : Let Q be an $r \times m$ matrix and Q_1, Q_2 .... Q_r be the rows of Q and let $A^{(1)}, A^{(2)}$, ... $A^{(n)}$ be the columns of A. Since $Q = \begin{bmatrix} Q_1 \\ Q_2 \\ Q_r \end{bmatrix}$ and

$A = \left[A^{(1)} A^{(2)} ......(A)^{(n)} \right]$ the product QA is difined and QA will be $r \times n$ matrix, such that

$$QA = \begin{bmatrix} Q_1A^{(1)} & Q_1A^{(2)} & \text{L} & Q_1A^{(n)} \\ Q_2A^{(1)} & Q_2A^{(2)} & \text{L} & Q_2A^{(n)} \\ \text{M} & & & \text{M} \\ Q_rA^{(1)} & Q_rA^{(2)} & \text{L} & Q_rA^{(n)} \end{bmatrix}$$

It is elear that if two rows of Q, are interchanged, then the corresponding two rows are also interchanged in the product QA.

This means that $[R_{ij}(Q)]A = R_{ij}(QA)$

If a row of Q, say Q_i be multiplied by a non zero scalar say c, then the row or QA is also multiplied by the same scalar c.

$$\therefore [R_i(c)\,(Q)]A = R_i(c)\,(QA)$$

If c times the jth row of Q be added to the ith row of Q, then the ith row of QA is added by c times the jth row of it

$$\therefore [R_{ij}(c)\; Q]A = R_{ij}(c)\,(QA)$$

Since A is an $m \times n$ matrix, we can write

$A = I_m A$ where I_m is the $m \times m$ identily matrix

$$\begin{aligned} \therefore \quad & R_{ij}(A) = R_{ij}(I_mA) = [R_{ij}\,(I_m)A] = E_{ij}A \\ & R_i(c)(A) = R_i(c)\,(I_mA) = [R_i(c)(I_m)]A = E_i(c)A \\ & R_{ij}(c)(A) = R_{ij}(c)\,(I_mA) = [R_{ij}(c)\,(I_m)]A = E_{ij}(c)A \end{aligned}$$

Theorem : Each elementary matrix is non singular. The inverse of an elementary matrix is an elementary matrix of the same type.

Proof : Since the rank of a matrix remains same under elementary transformation and since elementry matrices are obtained by giving elementary transformation to the identity matrix I_n of order n, it follows that the rank of an elementary matrix is n and hence it is non singular.

Let A be an $n \times p$ matrix and let it be premultiplied by E_{ij}.

$$E_{ij}A = R_{ij}(A)$$

$$\therefore \quad E_{ij}(E_{ij}A) = R_{ij}(R_{ij}(A)) = A$$

$$\therefore \quad (E_{ij}E_{ij})A = A$$

$$\Rightarrow \quad E_{ij}E_{ij} = I_n \text{ or, } E_{ij}^{-1} = E_{ij}$$

Similarly we can also prove that

$$\left[E_i(c)\right]^{-1} = E_i\left(\frac{1}{c}\right) \text{ and } \left[E_{ij}(c)\right]^{-1} = E_{ij}(-c)$$

Theorem : A matrix is non-singular if and only if it can be expressed as the product of a finite number of elementary matrices.

Proof : Let A be a non-singular matrix of order n. Since an $n \times n$ matrix A is non-singular if and only if A is row equivalent to the indentity matrix I_n, so I_n can be obtained by a finite number of elementary row operations.

Since the effect of an elementary row operation on A is expressed by the product PA where P is an $n \times n$ elementary matrix, I_n can be expressed as

$I_n = P_1 P_2 ......P_rA$ where $P_1, P_2 ......P_r$ are

elementary matrices of order n.

Since each elementary matrix P_i is non singular, P_i^{-1} exists.

$$\begin{aligned} \therefore \quad A &= (P_1P_2 ......P_r)^{-1}I_n \\ &= P_r^{-1}P^{-1}_{r-1}......P_2^{-1}P_1^{-1}I_n \\ &= P_r^{-1}P^{-1}_{r-1}......P_2^{-1}P_1^{-1} \\ &= Q_rQ_{r-1}......Q_2Q_1 \end{aligned}$$

where $Q_i = P_i^{-1}$ is an elemntary matrix. Conversely, if an $n \times n$ matrix is the product of elementary matrices, we have to show that A is non-singular.

Let $A = E_1, E_2 ......E_r$, where E_i is an elementary matrix. Since each E_i is non-singular. A is non-singular.

Definition : When a non zero matrix has been reduced by elementary operations to one of the forms

$$\left[\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right] or\ [I_r \mid 0]\ or \left[\begin{array}{c} I_r \\ \hline 0 \end{array}\right] or\ I_r$$

we say that it has been reduced to **normal form** or **canonical form for equivalence.**

The normal form of a zero matrix is that zero matrix.

EXAMPLES

Ex. 1. We apple elementary row and column transformations to the matrix

$A = \begin{bmatrix} 1 & 2 & 0 \\ 4 & 6 & 9 \\ 0 & -2 & 9 \end{bmatrix}$ to reduce it to its normal form.

we have $\begin{bmatrix} 1 & 2 & 0 \\ 4 & 6 & 9 \\ 0 & -2 & 9 \end{bmatrix} \xrightarrow{-4R_1+R_2} \begin{bmatrix} 1 & 2 & 0 \\ 0 & -2 & 9 \\ 0 & -2 & 9 \end{bmatrix}$

$$\xrightarrow{-2C_1+C_2} \begin{bmatrix} 1 & 0 & 0 \\ 0 & -2 & 9 \\ 0 & -2 & 9 \end{bmatrix} \xrightarrow{-\frac{1}{2}R_2} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -\frac{9}{2} \\ 0 & -2 & 9 \end{bmatrix}$$

$$\xrightarrow{2R_2+R_3} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -\frac{9}{2} \\ 0 & 0 & 0 \end{bmatrix} \xrightarrow{\frac{9}{2}C_2+C_3} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} I_2 & 0 \\ 0 & 0 \end{bmatrix}$$

Theorem : Any $m \times n$ matrix A can be reduced to normal form by elementary operations.

Proof : Since A is nonzero, there exists some element a_{ij} of A which is nonzero. If necessary the operations $R_i \leftrightarrow R_1$ and $C_j \leftrightarrow C_1$ are applied to move a_{ij} to the (1, 1) position. Hence we can assume without loss of generality that $a_{11} \neq 0$. The following sequence of elementary row and column operations

$$\frac{1}{a_{11}} R_1, (-a_{21} R_1 + R_2), \ldots (a_{m1}R_1 + R_m)$$

$$(-a_{12}C_1 + C_2), \ldots\ldots.(-a_n C_1 + C_n)$$

will transform A to an equivalent matrix of the form

$\left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & B \end{array}\right]$ if $m > 1$ and $n > 1$, and B is $(m-1) \times (n-1)$. If either $m = 1$ or $n = 1$, or both $m = 1$ and $n - 1$, then A is equivalent to $[1, 0, 0 \ldots\ldots\ldots\ldots 0]$ or $\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$ I_1, respectively and the proof is complete.

If B is a null matrix, then the proof is complete ; otherwise repeat the above procedure for the non-zero matrix B ; if $m > 2$ and $n > 2$, then A will be equivalent to a matrix of the form

$$\left[\begin{array}{c|c} I_2 & 0 \\ \hline 0 & C \end{array}\right]$$

where C in $(m-2) \times (m-2)$ matrix. If $m = 2$ or $n = 2$ or if both $m = 2$ and $n = 2$, then A is equivalent to $[I_2 ; 0]$ a $\begin{bmatrix} I_2 \\ 0 \end{bmatrix}$ or I_2, respectively and the proof is complete. If otherwise occurs we continue the above procedure until a null matrix appears in the lower right hand corner or until all of the rows or columns are exhausted. In either event, the final matrix is in normal form and since only elementary row and column operations have been used, A is equivalent to the final matrix.

Ex. 2. Show that the matrix $\begin{bmatrix} 2 & 0 & 1 \\ 3 & 3 & 0 \\ 6 & 2 & 3 \end{bmatrix}$ is non-singular and express it as a product of elemenrary matrices.

Let the given matrix be denoted by A we apply elementary row operations on A to reduec it to a row-reduced echelon matrix.

$$A \xrightarrow{\frac{1}{2}R_1} \begin{bmatrix} 1 & 0 & \frac{1}{2} \\ 3 & 3 & 0 \\ 6 & 2 & 3 \end{bmatrix} \xrightarrow[R_3-6R_1]{R_2-3R_1} \begin{bmatrix} 1 & 0 & \frac{1}{2} \\ 0 & 3 & -\frac{3}{2} \\ 0 & 2 & 0 \end{bmatrix}$$

$$\xrightarrow{\frac{1}{3}R_2} \begin{bmatrix} 1 & 0 & \frac{1}{2} \\ 0 & 1 & -\frac{1}{2} \\ 6 & 2 & 0 \end{bmatrix} \xrightarrow[R_3-6R_1]{R_3-2R_2} \begin{bmatrix} 1 & 0 & \frac{1}{2} \\ 0 & 1 & -\frac{1}{2} \\ 0 & 1 & 1 \end{bmatrix} \xrightarrow[R_2+\frac{1}{2}R_3]{R_1-\frac{1}{2}R_3} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Since A is row equivalent to I_3, A is non-singular. We observe that

$$\left(R_2+\frac{1}{2}R_3\right)\left(R_1-\frac{1}{2}R_3\right)(R_3-2R_2)\left(\frac{1}{3}R_2\right)(R_3-6R_1)(R_2-3R_1)\left(\frac{1}{2}R_1\right)A=I_3$$

or, $E_{23}\left(\frac{1}{2}\right)E_{13}\left(-\frac{1}{2}\right)E_{32}(-2)E_2\left(\frac{1}{3}\right)E_{31}(-6)E_{21}(-3)E_1\left(\frac{1}{2}\right)A=I_3$

$$\therefore\ A=\left\{E_1\left(\frac{1}{2}\right)\right\}^{-1}\{E_{21}(-3)\}^{-1}\{E_{31}(-6)\}^{-1}\left\{E_2\left(\frac{1}{3}\right)\right\}^{-1}\{E_{32}(-2)\}^{-1}\times$$

$$\left\{E_{13}\left(-\frac{1}{2}\right)\right\}^{-1}\left\{E_{23}\left(\frac{1}{2}\right)\right\}^{-1}$$

$$=E_1(2)E_{21}(3)E_{31}(6)E_2(3)E_{32}(2)E_{13}\left(\frac{1}{2}\right)E_{23}\left(-\frac{1}{3}\right)$$

Minimum Polynomial : Let A be a square matrix of order n over a field F. Observe that there are non-zero polynomials $f(t)$ for which $f(A)=0$; for example if f(t) is the characteristic polynomial of A, then by Cayley Hamilton's theorem $f(A)=0$. Among these polynomials for which A is a root we consider those of lowest degree and form them we select one whose leading. coefficient is 1, i.e. the polynomial is a monic. Such a polynomial $m(t)$ exists and is unique ; we call it **minimum polynomial** of A.

Theorem : The minimum polynomial $m(t)$ of a matrix A divides every polynomial which has A as zero. In particular, $m(t)$ divides the characteristic polynomial of A.

Proof : Suppose $f(t)$ is a polynomial for which $f(A)=0$. By the division algorithm there exist polynomials $q(t)$ and $r(t)$ for which $f(t)=m(t)\,q(t)+r(t)$ .... (1), where either $r(t)=0$ or deg $r(t)<$ deg $m(t)$. Substituting $t=A$ in equation (1) and unsing the fact that $f(A)=0$ and $m(A)=0$, we get $r(A)=0$

If $r(t)\neq 0$, then by the division algorithm deg $r(t)<$ deg $m(t)$

This means that there is a polynomial $r(t)$ of degree less than that of $m(t)$ such that $r(A)=0$. [from (2)] which is a contradiction, since $m(t)$ is a polynomial of least degree such that $m(A)=0$

Hence $r(t)=0$, and so

$f(t)=m(t)\,q(t)$, and $m(t)$ divides $f(t)$

As a particular case, since A satisfies its own characteristic equation by Cayley Hamilton's theorem $m(t)$ divides the characteristic polynomial.

Theorem : Let $m(t)$ be the minimum polynomial of an n square matrix A. Then the characteristic polynomical of A divides $m(t)^n$.

Proof : Suppose $m(t) = t^r + c_1t^{r-1} + ...+ c_{r-1}t + c_r$.

Consider the following matrices

$$\begin{aligned}
&B_0 = I\\
&B_1 = A + c_1I\\
&B_2 = A^2 + c_1A + c_2I\\
&\vdots\\
&B_{r-1} = A^{r-1} + c_1A^{r-2} + \ldots + c_{r-1}I
\end{aligned}$$

Then
$$\begin{aligned}
&B_0 = I\\
&B_1 - AB_0 = c_1I\\
&B_2 - AB_1 = c_2I\\
&\ldots\ldots\ldots\\
&B_{r-1} - AB_{r-2} = c_{r-1}I
\end{aligned}$$

Also
$$\begin{aligned}
-AB_{r-1} &= c_rI - (A^r + c^1A^{r-1} + ...+ c_{r-1}A + c_rI)\\
&= c_rI - m(A)\\
&= c_rI
\end{aligned}$$

Set $B(t) = t^{r-1}B_0 + t^{r-2}B_1 + ...+ tB_{r-2} + B_{r-1}$.

Then
$$\begin{aligned}
(tI - A)\,B(t) &= (t^r - B_0 + t^{r-1}B_1 + ...+ tB_{r-1})\\
&\quad - (t^{r-1}AB_0 + t^{r-2}AB_1 + ...+ AB_{r-1})\\
&= t^rB_0 + t^{r-1}(B_1 + AB_0) + t^{r-2}(B_2 + AB_1) + \ldots + t(B_{r-1} - AB_{r-2}) - AB_{r-1}\\
&= t^rI + c_1t^{r-1}I + c_2t^{r-2}I + \ldots. + c_{r-1}tI + c_rI\\
&= m(t)I
\end{aligned}$$

Taking determinant of both sides we get

$$|tI - A|\,|B(t)| = |m(t)\,I| = (m(t))^n.$$

Since $|B(t)|$ is a polynomial, $|tI - A|$ divides $(m(t))^n$.

Hence the characteristic polynomial divides $(m(t))^n$.

Theorem : The characteristic polynomial and the minimum polynomial of a matrix A have the same irreducible factors.

Proof : Suppose $f(t)$ is an irreducible polynomial. If $f(t)$ divides the minimum polynomial $m(t)$, then since $m(t)$ divides the characteristic polynomial, $f(t)$ must divide the characteristic polynomial. On the other hand, if $f(t)$ divides the characteristic polynomial of A, by the last theorem, $f(t)$ must divide $m(t)^n$. But $f(t)$ is irreducible, hence $f(t)$ also divides $m(t)$.

Thus $m(t)$ and the characteristic polynomial have the same irreducible factors.

Ex. 3. Find the minimum polynomial $m(t)$ of the matrix

$$A=\begin{pmatrix}2&1&0&0\\0&2&0&0\\0&0&2&0\\0&0&0&5\end{pmatrix}$$

The characteristic polynomial of A is $(t-2)^3$ $(t-5)$. Since the characteristic polynomial and the minimum polynomial have the same irreducible factors, if follows that both $t-2$ and $t-5$ must be factors of $m(t)$. also $m(t)$ must divide the characteristic polynomial. Hence it follows that $m(t)$ must be one of the following three polynomials ;

(i) $m(t) = (t-2)\,(t-5)$

(ii) $m(t) = (t-2)^2\,(t-5)$

(iii) $m(t) = (t-2)^3\,(t-5)$

we can verify that for the type (i), $m(A) \neq 0$, but for types (ii) and (iii), $m(A) = 0$

Since $m(t)$ is minimum polynomial,

$$m(t) = (t-2)^2\,(t-5)$$

Characteristic and minimum polynomials of linear operators :

Let $T : V \rightarrow V$ be a linear operator on a vector space with finite dimension.

We define the characteristic polynomial of T to be the characteristic polynomial of any matrix represstation of T.

On the otherhand, the minimum polynomial of the operator T is defined independently of the theory of matrices, as the polynomial of lowest degree and leading coeffieient 1 which has T as a zero. However, for any polynomial $f(t)$,

$$f(T) = 0 \text{ if and only if } f(A) = 0$$

where A is any matrix representation of T. Accordingly, T and A have the same minimum polynomial.

Jordan Normal form :

Let us consider $n \times n$ matrices with elements in the field F. We will isolate a special type called Jordan matrices and it may be shown that these matrices serve as a normal form for a very broad class of matrices. Namely, matrices, all the caracteristic roots of which lie in the base field F (and only such matrices) are similar to certain Jordan matrices; we say that they can be reduced to a Jordan normal form. It will then follow, if for the field F we take the field of complex numbers, then ay matrix with complex elements can be reduced to a Jordan normal form in the field of complex numbers.

We will need some difinitions. A kth order Jordan submatrix referring to the number λ_0 is a matrix of order k, $1 \le k \le n$, of the form

$$\begin{bmatrix} \lambda_0 & 1 \ldots\ldots 0 \\ & \lambda_0 1 \ldots\ldots \\ & 0 \ldots\ldots \lambda_0 \end{bmatrix}$$

In otherwords, one and the same number λ_0 from the field F occupies the principal diagonal, with unity along the diagonal immediately above and zero else where. Thus,

$$[\lambda_0], \begin{bmatrix} \lambda_0 & 1 \\ 0 & \lambda_0 \end{bmatrix}, \begin{bmatrix} \lambda_0 & 1 & 0 \\ 0 & \lambda_0 & 1 \\ 0 & 0 & \lambda_0 \end{bmatrix}$$

are respectively Jordan submatrices of first, second and third order.

A Jordan matrix of order n is a matrix of order n having the form

$$J = \begin{bmatrix} \boxed{J_1} & & & 0 & \\ & \boxed{J_2} & & & \\ & & \boxed{J_3} & & \\ & & & \mathrm{O} & \\ 0 & & & & \boxed{J_s} \end{bmatrix}$$

The elements along the principal diagonal are **Jordan submatrices** or **Jordan blocks** $J_1, J_2, \ldots\ldots J_s$ of certain orders, not necessarily distinct, refering to certain numbers (not necessarily distinct either) lying in the field F.

Thus, a matrix is a Jordan matrix if and only if it has form

$$\begin{bmatrix} \lambda_1 & \varepsilon_1 \ldots\ldots 0 & & \\ & \lambda_2 \; \varepsilon_2 \ldots\ldots & & \\ & \ldots\ldots \lambda_{n-1} \varepsilon_{n-1} & & \\ & & & \lambda_n \end{bmatrix}$$

where λ_i, $i = 1, 2$ .... n are arbitary numbers in F and every $\in_j$, $j = 1, 2$ .... $n - 1$ is equal to unity or zero ; note that if $\in_j = 1$, then $\lambda_j = \lambda_{i+1}$.

Diagonal matrices are a special case of Jordan matrices. These are Jordan matrices whose submatrices are of order 1.

The theorem below describes all the eigenvalues and eigenvectors of a Jordan block.

Theorem : Let J be a Jordan block of order k. Then J has exactly one eigenvalue, which is equal to the scalar on the main diagonal. The corresponding eigen vectors are the non zero scalar multiples of the k-dimensional unit coordinate vector [1, 0, ... 0]

Proof : Suppose the diagonal entries of J are equal to λ. A column vector $X = [x_1, x_2, \ldots\ldots x_k]^T$ satisfies the equation $JX = \lambda X$ if and only if its components satisfy the following k scalar equations.

$$\lambda x_1 + x_2 = \lambda x_1$$

$$\lambda x_2 + x_3 = \lambda x_2$$

$$[$$

$$\lambda x_{k-1} + x_k = \lambda x_{k-1}$$

$$\lambda x_k = \lambda x_k$$

From the first $(k - 1)$ equations we obtain

$$x_2 = x_3 = \ldots = x_k = 0$$

So λ is an eigenvalue for J and all eigenvectors have the same form $x_1[1, 0, \ldots 0]$ with $x_1 \neq 0$.

To show that λ is the only eigenvalue for J, assume that $JX = \mu X$ for sume scalar $\mu \neq \lambda$.

Then the components satisfy the following k scalar equtions

$$\lambda x_1 + x_2 = \mu x_1$$

$$\lambda x_2 + x_3 = \mu x_2$$

....

$$\lambda x_{k-1} + x_k = \mu x_{k-1}$$

$$\lambda x_k = \mu x_k$$

Because $\lambda \neq \mu$, the last relation gives $x_k = 0$ and from the other equations we get $x_{k-1} = x_{k-2} = ... = x_2 = x_1 = 0$. Hence only zero vector satisfies $JX = \mu X$, so no scalar different from λ can be an eigen value for J.

We now state a very important theorem

Theorem : Let V be an n-dimensional liner space with complex scalars, and let $T : V \to V$ be a linear transformation of V into itself. Then there is a basis for V relative to which T has a block-diagonal matrix representation diag $(J_1, J_2, ... J_m)$, with each J_k being a jordan block.

The proof of the theorem may be found in Linear Algebra by *TM* Apostol.

Ex. 3. Verify that the matrix $A = \begin{bmatrix} -1 & 3 & 0 \\ 0 & 2 & 0 \\ 2 & 1 & -1 \end{bmatrix}$ has eigenvalues 2, –1, –1.

Find a nonsigngular matrix C with initial entry $C_{11} = 1$ that transforms A to the following Jordan normal form

$$C^{-1}AC = \begin{bmatrix} 2 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{bmatrix}$$

Clearly the eigen values are the roots of the equation

$$\det \begin{bmatrix} -1-\lambda & 3 & 0 \\ 0 & 2-\lambda & 0 \\ 2 & 1 & -1-\lambda \end{bmatrix} = 0$$

$\therefore (2 - \lambda)(1 + \lambda)^2 = 0 \quad \therefore \lambda = -1, -1, 2$

The eigen vector corresponding to $\lambda = 2$ is obtained by solving equations

$$-3x + 3y \qquad = 0$$

$$2x + y - 3z = 0$$

$$\Rightarrow x = y = z$$

$\therefore$ The eigen vector corresponding to $\lambda = 2$ is $k(1, 1, 1)$, k is a constant

Corresponding to $\lambda = -1$, the equations are

$$3y = 0$$

$$2x + y = 0$$

$$\therefore \quad x = 0, y = 0,$$

$\therefore$ Eigen vector is $(0, 0, a)$, a is arbitrary

We construct the matrix C whose first two columns are the eigen vectors corresponding to $\lambda = 2$ and $\lambda = -1$. Since $C_{11} = 1$, we msut have $k = 1$. The third column is chosen in such a way that

$AC = CB$, where B is the Jordan normal form

$$\therefore \quad C = \begin{bmatrix} 1 & 0 & b \\ 1 & 0 & c \\ 1 & a & d \end{bmatrix}$$

$AC = CB \Rightarrow$

$$\begin{bmatrix} -1 & 3 & 0 \\ 0 & 2 & 0 \\ 2 & 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 & b \\ 1 & 0 & c \\ 1 & a & d \end{bmatrix} = \begin{bmatrix} 1 & 0 & b \\ 1 & 0 & c \\ 1 & a & b \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{bmatrix}$$

$$\Rightarrow \quad -b + 3c = -b$$

$$2c = -c$$

$$2b + c - d = a - d$$

$$\Rightarrow \quad c = 0, a = 2b$$

Hence $C = \begin{bmatrix} 1 & 0 & b \\ 1 & 0 & 0 \\ 1 & 2b & d \end{bmatrix}$ where $b \neq 0$ and d is arbitrary.

EXERCISES

1. Prove that the matrices $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ and $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ have the same eigenvalues but are not similar.

2. Find a nonsingular matrix C such that $C^{-1}AC$ is a diagonal matrix, where

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

3. Detemine the eigenvalues and eigenvectors of the matrix

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

and thereby show that it is not similar to a diagonal matrix.

4. For the matrix $\begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$ express A^3 in terms of I, A and A^2.

5. Verify that the matrix $A = \begin{bmatrix} -11 & -7 & -5 \\ 16 & 11 & 6 \\ 12 & 6 & 7 \end{bmatrix}$ has eigenvalues 1, 3, 3.

Find a nonsingular matrix C with initial entry $C_{11} = 1$ that transforms A to Jordan normal from :

$$C^{-1}AC = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{bmatrix}$$

6. Determine all possible canonical forms J for a matrix of order 5 whose minimal polynomial is $m(t) = (t-2)^2$.

7. Find the minimum polynomial m(t) of the matrix

$$\begin{bmatrix} \lambda & a & 0 \\ 0 & \lambda & a \\ 0 & 0 & \lambda \end{bmatrix}$$

8. Show that a matrix A and its transpose A^T have the same characteristic polynomial.

9. Suppose $M = \begin{bmatrix} A_1 & B \\ 0 & A_2 \end{bmatrix}$ where A_1 and A_2 are square matrices. Show that the characteristic polynomial of M is the product of the characteristic polynomials of A_1 and A_2.

10. Show that the matrix $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ is not diagonalizable.

11. For the matrix $\begin{bmatrix} 1 & -3i \\ i & -1 \end{bmatrix}$ find all the eigen values and linearly independent vectors.

12. Use Cayley-Hamilton theorem to find A^{100} where

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

13. Diagonalise the matrix $A=\begin{bmatrix} -1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{bmatrix}$

14. Find an orthogonal matrix P such that P^TAP is a diagonal matrix, where

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

Unit : 5 □ Quadratic Forms, Canonical Forms, Classification of Quadrics

Quadratic forms :

Def Let $x = [x_1, x_2 .....x_n]^T$ (i. e. x is n component column vector) be an n-vector in the vector space V over a field F and let $A = (a_{ij})$ be an n-square matrix over F. Then any polynomial of the form

$$q(x_1, x_2 .....x_n) = x^T \ Ax = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}\, x_i\, x_j \qquad ...\ (1)$$

is called a quadratic form of order n over F in the variables $x_1, x_2 .. x_n$. A is called the matrix of the form and rank of A is the rank of the quadratric form. A quadratric form is said to be real (or complex) if the matrix A is over a real field R(or over a complex field C) and the variable x is in the vector space R^n (or C^n). Unless stated to the contrary, we shall confine our discussion to quadratic forms in which the matrix A as well as the variables are considered to be real.

EXAMPLES

Ex. 1. Let $q(x_1, x_2, x_3) = 2x_1^2 - 3x_1x_2 + 3x_2^2 + 2x_1x_3 + 4x_2x_3 + x_3^2$ be a quadratic form in the variables x_1, x_2, x_3.

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

can be verified by computing $x^T Ax$, where $x^T = [x_1, x_2, x_3]$ as a matrix of the quadratic form.

There are numerous ways in which this matrix A may be expressed. But there is only one way in which the matrix A can always be put into a symmetric form. This is done by writing the eoefficients of the product $x_ix_j\, (i \neq j)$, as the sum of two equal terms of the form

$$q(x_1, x_2, x_3) = 2x_1^2 - \frac{3}{2}x_1x_2 + x_1x_3 - \frac{3}{2}x_2x_1 + 3x_2^2 + 2x_2x_3 + x_3x_1 + 2x_3x_2 + x_3^2$$

and noting that the coefficient of the product $x_i x_j$ is a_{ij} consequently the matrix of the quadratic form in this example can also be writen as

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

which is a symmetric matrix.

The same matrix may be obtained by finding $\frac{1}{2}\frac{\partial q}{\partial x_1}, \frac{1}{2}\frac{\partial q}{\partial x_2}, \frac{1}{2}\frac{\partial q}{\partial x_3}$ and writing the coefficient matrix, which is exactly equal to A.

Thus, in our present example,

$$\frac{1}{2}\frac{\partial q}{\partial x_1} = 2x_1 - \frac{3}{2}x_2 + x_3$$

$$\frac{1}{2}\frac{\partial q}{\partial x_2} = -\frac{3}{2}x_1 + 3x_2 + 2x_3$$

$$\frac{1}{2}\frac{\partial q}{\partial x_3} = x_1 + 2x_2 + x_3$$

$$\therefore A = \text{coefficient matrix} = \begin{bmatrix} 2 & -\frac{3}{2} & 1 \\ -\frac{3}{2} & 3 & 2 \\ 1 & 2 & 1 \end{bmatrix}$$

Next we shall show that there is only one way in which a quadratic form (1) can be expressed uniquely by a real symmetric matrix. For, suppose a quadratric form (1) is given.

We can write it as

$$q(x_1, x_2, .. x_n) = x^T A x = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}\, x_i\, x_j$$

$$= \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}\, x_i\, x_j + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}\, x_i\, x_j$$

$$= \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}\, x_i\, x_j + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}\, x_i\, x_j\,,$$

(interchanging i & j in the second sum)

$$= \sum_{i=1}^{n}\sum_{j=1}^{n}\left(a_{ij}\, x_{ji}\right) x_i x_j$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij}\, x_i x_j = x^T Bx$$

where $b_{ij} = \frac{1}{2}\left(a_{ij} + a_{ji}\right) = \frac{1}{2}\left(a_{ji} + a_{ij}\right) = b_{ji}$

and so $B = \frac{1}{2}\left(A + A^T\right) = B^T$

which is a symmetric matrix of the quadratic form (1). Henceforth, unless otherwise stated, the matrix A of the quadratic form x^TAx will be considered to be real and symmetric.

Definition : The determinant det A of the matrix A of a quadratic form x^TAx is called the discriminant of the quadratic form. The quadratic form is said to be nonsingular if its discriminant is non zero, otherwise, it is said to be singular.

The substitution $x = Py$ in the quadratic form x^TAx is called a linear transformation of the quadratic form. The linear transformation is referred to as real (or complex) if P is real (or complex) and nonsingular (or singular) if the matrix P is nonsingular (or singular).

If a new coordinate system is introduced in the vector space V so that the old variable x is related to the new variable y by $x = Py$, where P is nonsingular transformation matrix, then (1) is changed into another quadratic form, viz.

$$\begin{aligned} x^TAx &= (Py)^T A(Py) \\ &= y^T (P^TAP)^T \\ &= y^TBy \qquad \text{.... (2)} \end{aligned}$$

where $B = P^TAP$ .... (3)

The matrix A of a quadratic form under non-singular transformation carried out in (2) changes to P^TAP which is symmetric if A is symmetric and the discriminant changes from det A to (det P)^2 det A, but its rank, i. e. rank A = rank P^TAP remains invariant. When such a nonsingular transformation exists over F such that (2) holds, these quadratic forms x^TAx and y^TBy are said to be equivalent over F. Since by definition, two matrices A and B are said to be congruent if (3) holds, it follows that two quadratic forms are equivalent if their corresponding matrices are congruent over F. Conversely, if (3) holds, then (2) is true for a nonsingular tranformation $x = Py$.

Thus x^TAx and y^TBy are equivalent. Formally, we have,

Definition : Two quadratic forms x^TAx and y^TBy are said to be equivalent over F if one can be obtanied from the other by a nonsingular transformation over F defined by $x = Py$. such that (2) and (3) are satisfied.

EXAMPLE

Ex. 1. The quadratic form defined by

$$\begin{aligned} q(x_1, x_2, x_3) &= 5x_1^2 + x_2^2 + 21x_3^2 + 2x_1^2 + 4x_1x_2 - 20x_1x_3 - 8x_2x_3 \\ &= x^TAx \end{aligned}$$

and $\quad q(y_1, y_2, y_3) = y_1^2 + y_2^2 + y_3^2 = y^TBy$

where $A = \begin{bmatrix} 5 & 2 & -10 \\ 2 & 1 & -4 \\ -10 & -4 & 21 \end{bmatrix}, \qquad B = \text{diag}\,(1, 1, 1)$

are equivalent. This is so because there exits a non singular matrix.

$$P = \begin{bmatrix} 2 & -1 & 0 \\ 0 & 2 & -1 \\ 1 & 0 & 0 \end{bmatrix}$$

such that $P^TAP = B$

The above quadratic forms are nonsingular because det A = det B = 1, is non zero.

Reduction of quadratic forms :

Definition : Diagonal and unit quadratic forms. If the matrix of a quadratic form is diagonal then it is called a diagonal quadratic form. In a special case when the matrix A is I_n, the quadratic form is called a unit quadratric form. For example the quadratric form $x^T I_n x = x_1^2 + x_2^2 + \ldots + x_n^2$ is a unit and x^T diag $(a_1, a_2, \ldots a_n)x = a_1 x_1^2 + a_2 x_2^2 + \ldots + a_n x_n^2$ is a diagonal quadratic form. In diagonal form some of the a_i's may be zero.

These two forms are sometimes called the **cononical representation** of a quadratic form.

Next we shall consider reduction of a quadratic form to a diagonal form.

Orthogond transformation : Reduction of a quadratic from to a diagonal form by means of an orthogonal transformation is referred to as orthogonal reduction.

Theorem : Every quadratic form $x^T Ax$ can be reduced to a diagonal quadratic form $y^T Dy$ by means of an orthogonal transformation P, where the diagonal elements of D are the eigenvalues of A.

Proof : Since A is symmetric and since for every real symmetric matrix A, there exixts an orthogonal transformation P such that at $P^T AP = \text{diag}\,[\lambda_1, \lambda_2, \ldots \lambda_n]$, where $\lambda_1, \lambda_2, \ldots \lambda_n$ are the eigenvalues of A not necessarily distinct and non zero, we have $P^T AP = D$ and $P^T P = I_n$, where D is a diagonal matrix having the eigen velues λ_i's, i = 1, 2 ... n, of A as diagonal elements.

Let $x = Py$. Then

$$x^T Ax = y^T (P^T AP) = y^T Dy$$

$$= \lambda_1 y_1^2 + \lambda_2 y_2^2 + \ldots + \lambda_n y_n^2$$

An easy consequence of the above theorem is the following corollary :

Corollary : Every quadratic form $x^T Ax$ of rank r can be reduced by an orthogonal transformation P to

$$\lambda_1 y_1^2 + \lambda_2 y_2^2 + \ldots + \lambda_r y_r^2$$

where $x = Py$ and $\lambda_1, \lambda_2, \ldots \lambda_n$ are the nonzero eigenvalues of A.

EXAMPLE

Ex. 1. Let $q(x_1, x_2, x_3) = x^T Ax$ be a real quadratic form, where

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

The characteristic equation of A is

$$\lambda^3 - \lambda^2 - 5\lambda - 3 = (\lambda + 1)^2(\lambda - 3) = 0$$

which shows that the eigencalues of A are $\lambda_1 = -1$ of multiplicity 2 and $\lambda_2 = 3$, simple.

By solving $\lambda_1 x = Ax$ we see that the eigen vectors associated with $\lambda_1 = -1$ can be expressed by $[a, -a, b]^T$ where a and b are real and hence the two independent eigen vectors corresponding to -1 are

$$x_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \text{ and } x_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

which are orthogonal.

By solving $\lambda_2 x = Ax$ we get $x_3 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ as an eigen vector associated with $\lambda_2 = 3$,

which is orthogonal to both x_1 and x_2. Hence the orthogonal matrix that diagonlises the matrix A is

$$P = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \end{bmatrix}$$

Hence using the transformation $x = Py$ the given quadratic form becomes

$q = y^T (P^T AP)y$

We may check that $P^T AP = \text{diag}(-1, -1, 3)$

$$\therefore\ q = y^T \operatorname{diag}(-1, -1, 3)y$$
$$= -1.y_1^2 - 1.y_2^2 + 3y_3^2 = -y_1^2 - y_2^2 + 3y_3^2$$

Definition : A real quadratic form $q(x_1, x_2 .. x_n)$ assumes the values 0 when $x = 0$. But q takes up different real values for different nonzero x.

A real quadratic form $q(x_1, x_2 .. x_n) = x^T Ax$ is said to be

(i) positive definite if $q > 0$ for all $x \neq 0$

(ii) positive semidefinite if $q \geq 0$ for all x and $q = 0$ for $x \neq 0$.

(iii) negative definite if $q < 0$ for all $x \neq 0$.

(iv) negative semidefinite if $q \leq 0$ for all x and $q = 0$ for $x \neq 0$.

(v) indefinite $q \geq 0$ for some $x \neq 0$ and $q \leq 0$ for some other $x \neq 0$.

In the cooresponding cases the associated real symmetric matrix A is said to be positive definite, positive semidefinite, negative definite, negative semidefinite and indefinite, respectively.

EXAMPLES

Ex. 1. Let us consider the quadratic form $q(x_1, x_2, x_3)$

$= x_1^2 + 2x_2^2 + 4x_3^2 + 2x_1x_2 - 4x_2x_3 - 2x_3x_1$

Clearly, $q = (x_1 + x_2 - x_3)^2 + (x_2 - x_3)^2 + 2x_3^2$

$\therefore\ q \geq 0$ for all (x_1, x_2, x_3) and $q = 0$ only when $x_1 = 0, x_2 = 0, x_3 = 0$

Therefore q is positive definite. The associated matrix $\begin{bmatrix} 1 & 1 & -1 \\ 1 & 2 & -2 \\ -1 & -2 & 4 \end{bmatrix}$ is therefore positive definite

Ex. 2. Show that the quadratic form $x_1^2 + 2x_2^2 + 3x_3^2 - 2x_1x_2 + 4x_2x_3$ is indefinite.

The associated matrix A of the quadratic form is

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

We apply congruence operations on A. We have

$$A \xrightarrow{R_2+R_1} \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 2 \\ 0 & 2 & 3 \end{bmatrix} \xrightarrow{C_2+C_1} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 2 & 3 \end{bmatrix}$$

$$\xrightarrow{R_3+2R_2} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & -1 \end{bmatrix} \xrightarrow{C_3-2C_2} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

Hence the reduced normal form is

$$q = x_1^2 + x_2^2 - x_3^2$$

Since $q > 0$ for $x_1 \neq 0$, $x_2 = x_3 = 0$ and $q < 0$ for $x_1 = x_2 = 0$, $x_3 \neq 0$. If follows that q is indefinite.

It may be shown that any real quadratic form q may be reduced to sum of squres of several unknows with coefficients +1 or –1 via a non singular linear transformation with real coefficients. Such a form is called **normal form** of the quadratic form.

A real quadratic form may be reduced to normal form by many different transformation ; however, to within the numbering of the unknowns, it can be reduced only to one normal form. This is demonstrated by the following important theorem, which is called the **law of inertia** of real quadratic forms.

Theorem : The number of positive and the number of negative squares int the normal form to which a given quadratic form with real coefficients can be reduced by a real nonsingular linear transformation is independent of the choice of the transformation.

Proof : Suppose that the statement of the theorem be not true. Let the quadratic form $q = q(x_1, x_2 .. x_n)$ be transformed to

$y_1^2 + y_2^2 + ... + y_m^2 - y_{m+1}^2 - ... - y_r^2$ by the transformation $x = Py$

and to $z_1^2 + z_2^2 + ... + z_k^2 - z_{k+1}^2 - ... - z_r^2$ by the transformation $x = Qz$, where P and Q are non singular and $m \neq k$.

Without loss of generality we may assume $m < k$.

Now, $y = P^{-1}x$ and $z = Q^{-1}x$

and let $P^{-1} = (p_{ij})$ and $Q^{-1} = (q_{ij})$

Let us consider m + n – k equations in n unknowns

$$\left.\begin{array}{l} p_{11}x_1 + p_{12}x_2 + \square + p_{1n}x_n = 0 \\ \square \qquad \square \qquad \square \\ p_{m1}x_1 + p_{m2}x_2 + \square + p_{mn}x_n = 0 \\ q_{k+11}x_1 + q_{k+12}x_2 + \square + q_{k+nn}x_n = 0 \\ \square \qquad \square \qquad \square \\ q_{n1}x_1 + q_{n2}x_2 + \square + q_{nn}x_n = 0 \end{array}\right\} \qquad (1)$$

Since $m < k$, there exists a non-zero solution $(x'_1, x'_2 .. x'_n)$ of the homogencous system.

When $x = (x'_1, x'_2 .. x'_n)$, let $y = (y'_1, y'_2 .. y'_n)$ and $z = (z'_1, z'_2 .. z'_n)$

Since $y = P^{-1}x = (p_{ij})x$, it follows from the first equation of (1) that $y'_1 = 0$.

Similarly, $y'_2 = ... = y'_m = 0$ and $z'_{k+1} = ... = z'_k = 0$.

$$\begin{aligned} \therefore q(x'_1, x'_2 .. x'_n) &= -y'^2_{m+1} - y'^2_{m+2} - \cdots - y'^2_r \\ &= z'^2_1 + z'^2_2 + .. + z'^2_k \end{aligned}$$

or, $\quad y'^2_{m+1} + y'^2_{m+2} + ... + y'^2_r + z'^2_1 + ... + z'^2_k = 0$

$\Rightarrow \quad y'_{m+1} = y'_{m+2} = ... = y'_r = 0$

and $\quad z'_1 = z'_2 = ... = z'_k = 0$

Since P and Q are non-singular matrices, we have

$$(x'_1, x'_2 .. x'_n) = (0,\ 0,\ ...\ 0),$$

which is a contradiction.

Therefore, the assumption that $m \neq k$ is wrong and hence m is invariant. This proves the theorem.

Since the rank of a quadratic form remains invaniant under a non-singular transformation, it follows that under all non-singular transformations the rank of a real quadratic form temains inivariant.

This is **Sylvester law of quadratic forms.**

The number of positive squates in the normal form to which a given real quadratic form is reduced, is called the positive index of inertia of this form ; the number of negative squares is termed as negative index of inertia. The number of positive indices minus the number of negative indices in the normal form of q is called the signature of q.

Ex. 3. Obtain a nonsingular transformation that will reduce the quadratic form

$$q = x_1^2 + 2x_2^2 + 3x_3^2 - 2x_1x_2 - 4x_2x_3$$

to normal form and hence find its signature.

The associated matrix is

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

We consider the block matrix (A, I)

$$(A. I) = \begin{bmatrix} 1 & -1 & 0 & 1 & 0 & 0 \\ -1 & 2 & 2 & 0 & 1 & 0 \\ 0 & 2 & 3 & 0 & 0 & 1 \end{bmatrix}$$

We shall apply congruence operations on (A. I)

We have

$$(A. I) \xrightarrow{R_2+R_1} \begin{bmatrix} 1 & -1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 2 & 1 & 1 & 0 \\ 0 & 2 & 3 & 0 & 0 & 1 \end{bmatrix} \xrightarrow{C_2+C_1} \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 2 & 1 & 1 & 0 \\ 0 & 2 & 3 & 0 & 0 & 1 \end{bmatrix}$$

$$\xrightarrow{R_3+2R_2} \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 2 & 1 & 1 & 0 \\ 0 & 0 & -1 & -2 & -2 & 1 \end{bmatrix} \xrightarrow{C_3-2C_2} \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & -1 & -2 & -2 & 1 \end{bmatrix}$$

The transformations $x = Py$, where $P = \begin{bmatrix} 1 & 1 & -2 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix}$

will transform the given quadratic form to normal form.

The transformation is

$$\begin{aligned} x_1 &= y_1 + y_2 - 2y_3 \\ x_2 &= \phantom{y_1 + {}} y_2 - 2y_3 \\ x_3 &= \phantom{y_1 + y_2 - 2{}} y_3 \end{aligned}$$

and the normal form is $y_1^2 + y_2^2 - y_3^2$ and hence the signature is $2 - 1 = 1$.

Suppose we have a quadratic form q in n unknowns with the matrix $A = (a_{ij})$. The minors of order 1, 2, ... n of this matrix situated in the upper left corner are the principal minors of A. Thus the principal minors of order 1, 2, 3, ... n are.

$$a_{11}, \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}, \mathrm{L} \begin{vmatrix} a_{11} & a_{12} & \mathrm{L} & a_{1n} \\ a_{21} & a_{22} & & a_{2n} \\ \mathrm{L} & \mathrm{L} & & \\ a_{n1}^{*} & a_{n2} & \mathrm{L} & a_{nn} \end{vmatrix}$$

We now consider the following theorem.

Theorem : A quadratic form q in n unknowns with real coefficients is positive definite if and only if its principal minors are strictly positive.

Proof : For $n = 1$, the theorem is true since the form is then ax^2 and therefore is positive definite if and only if $a > 0$. Thus the theorem holds for $n = 1$. We assume that the theorem holds for a quadratic form with $n - 1$ variables $x_1, x_2 .. x_{n-1}$ and with real coefficients and we shall prove the theorem for n variables $x_1, x_2 .. x_n$.

We note that if a non-singular linear transformation $x = Py$ is given to quadratic form, then the reduced quadratic form is $y^T(P^TAP)y$ and the determinant of the transformed matrix P^TAP is

$$\left|P^T AP\right| = \left|P^T\right|\left|A\right|\left|P\right| = \left|A\right|\left|P\right|^2$$

This shows that the determinant of the matrix of a quadratic form maintains the same sign after nonsingular linear transformation.

Now, let the real quadratic form in n variables $x_1, x_2 .. x_n$ be

$$q = \sum_{i,j=1}^{n} a_{ij} x_i x_j$$

It can be written as

$$q = \phi\left(x_1 x_2 \mathrm{L}\ x_{n-1}\right) + 2\sum_{i=1}^{n-1} a_{in} x_i x_n + a_{nn} x_n^2 \qquad ...(1)$$

where ϕ is a quadratic form in $(n - 1)$ variables $x_1, x_2 .. x_{n-1}$. The principal minors of the form ϕ evidently coincide with all principal minors of the form q except the last.

Let the form q be positive definite. Then clearly the form ϕ will be positive definite. For, if ϕ is not positive definite, there is a set of values of $x_1, x_2 .. x_{n-1}$ not all zero such that ϕ is not strictly positive. Then taking these values and $x_n = 0$, we find that q is not positive definite, which is a contradiction. By the induction hypothesis all the principal minors of ϕ, that is of q are positive. The last principal minor of q is also positive, because q is reduced by a nonsingular linear transformation to a normal form consiting of n positive squares, the determinant of this normal form is strictly positive and so the determinant of the form q is strictly positive.

Now let all the principal minors of the form q be strictly positive. So all the principal minors of ϕ are positive and hence by the induction hypothsis ϕ is positive definite. Therefore, there is a nonsingular linear transformation of the unknowns $x_1, x_2 .. x_{n-1}$ which reduces ϕ to a sum of $(n-1)$ positive squares in the new unknowns $y_1, y_2 .. y_{n-1}$. By setting $x_n = y_n$, this linear tranformation may be completed to form a (non-singular) linear transformation of all the unknews $x_1, x_2 .. x_n$. By (1), the form q is reduced by the transformation to

$$q = \sum_{i=1}^{n-1} y_i^2 + 2\sum_{i=1}^{n-1} b_{in}\, y_i\, y_n + b_{nn}\, y_n^2 \qquad ...(2)$$

Since $\quad y_i^2 + 2b_{in}y_iy_n = (y_i + b_{in}y_n)^2 - b_{in}^2 y_n^2$

it follows that the non-singular linear transformation

$$z_i = y_i + b_{in}y_n, \; i = 1, 2, ... n-1$$

$$z_n = y_n$$

reduces the form q by (2) to the canonical form

$$q = \sum_{i=1}^{n-1} z_i^2 + cz_n^2 \qquad ...(3)$$

To prove positive definiteness of q we have to show that c is positive. Clearly, the determinant of the form on the right hand side of (3) is c and since the form (3) is obtained by applying two non-singular linear transformations from the form (1) whose determinat is positive, it follows that c is positive.

This completes the proof.

EXAMPLES

Ex. 1. The quadratic form

$$q = 5x_1^2 + x_2^2 + 5x_3^2 + 4x_1x_2 - 8x_1x_3 - 4x_2x_3$$

is positive definite, since its principal minors

$$5(>0), \quad \begin{vmatrix} 5 & 2 \\ 2 & 1 \end{vmatrix} = 1(>0) \quad \text{and} \quad \begin{vmatrix} 5 & 2 & -4 \\ 2 & 1 & -2 \\ -4 & -2 & 5 \end{vmatrix} = 1(>0)$$

are all positive.

Ex. 2. For the quadratic form

$$q = 3x_1^2 + x_2^2 + 5x_3^2 + 4x_1x_2 - 8x_1x_3 - 4x_2x_3$$

the matrix is $\begin{bmatrix} 3 & 2 & -4 \\ 2 & 1 & -2 \\ -4 & -2 & 5 \end{bmatrix}$

The principal minors are

$$3(>0), \quad \begin{vmatrix} 3 & 2 \\ 2 & 1 \end{vmatrix} = -1(<0) \quad \text{and} \quad \begin{vmatrix} 3 & 2 & -4 \\ 2 & 1 & -2 \\ -4 & -2 & 5 \end{vmatrix} = 1(<0)$$

The quadratic form is, therefore, not positive definite, it is in fact indefinite.

Simultaneous reduction of two quadratic forms

Let there be a pair of real quadratic forms q_1 and q_2 in n unknowns $x_1, x_2 .. x_n$. We try to investigate whether there is a non-singular linear transformation of the unknown $x_1, x_2 .. x_n$ which will simultaneously reduce both the forms q_1 and q_2 to canomical form.

In the general case the answer is no. But under some restriction on the nature of forms q_1 and q_2 the answer is yes. So we consider the following theorem.

Theorem : If q_1 and q_2 form a pair of real quadratic forms in n unknowns, and the second one is positive definite. then there exists a nonsingular linear transformation which simultaneously reduces q_2 to normal form and q_1 to canonical form.

Proof : Let us first perform the nonsingular linear transformation of the unknowns $x_1, x_2 .. x_n$.

$$x = Ty$$

which reduces the positive definite form q_2 to normal form,

$$q_2(x_1, x_2 \ldots. x_n) = y_1^2 + y_2^2 + \ldots.. + y_n^2$$

Then the form q_1 will be transformed into some form ϕ in new unknowns.

$$q_1(x_1, x_2 \ldots. x_n) = \phi\ (y_1, y_2 \ldots. y_n)$$

Now perform an orthogonal transformation of the unknowns $y_1, y_2 \ldots. y_n = Qz$ which reduces ϕ to principal axes,

$$\phi(y_1, y_2 \ldots. y_n) = \lambda_1 z_1^2 + \lambda_2 z_2^2 + \ldots. + \lambda_n z_n^2$$

This transformation carries the sum of the squares of the unknowns $y_1, y_2 \ldots. y_n$ into the sum of the squares of the unknowns $z_1, z_2 \ldots z_n$. As a result we get.

$$q_1(x_1, x_2 \ldots. x_n) = \lambda_1 z_1^2 + \lambda_2 z_2^2 + \ldots. + \lambda_n z_n^2$$

$$q_2(x_1, x_2 \ldots. x_n) = z_1^2 + z_2^2 + \ldots. + z_n^2$$

That is the linear transformation,

$$x = (Ty)z$$

is the required transformation.

Classification of quadrics :

The general equations of second degree in x, y, z is

$$ax^2 + by^2 + cz^2 + 2hxy + 2fyz + 2gzx + 2ux + 2vy + 2wz + d = 0$$

In matrix notation the equation may be expressed as

$$X^T AX + BX + dI = 0$$

where $A = \begin{bmatrix} a & h & g \\ h & b & f \\ g & f & c \end{bmatrix}$, $X = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ $B = [\ 2u\ 2v\ 2w]$

and I is the identity matrix of order 3.

Since A is a real symmetric matrix, there exists an orthogonal matrix P such that

$P^T AP$ is the diagonal matrix $D = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}$

where $\lambda_1, \lambda_2, \lambda_3$ are the eigen values of A.

Therefore, by the orthogonal transformation $X = PX'$,

where $X' = \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix}$, the equation transforms to

$$X'^T DX' + (BP)X' + dI = 0$$

or, $$\lambda_1 x'^2 + \lambda_2 y'^2 + \lambda_3 z'^2 + 2u_1 x' + 2v_1 y' + 2w_1 z' + d = 0 \ldots (1)$$

We shall consider the following cases :

Case I Rank of $A = 3$.

In this case $\lambda_1, \lambda_2, \lambda_3$ are all non-zero. Let us transfer the origin to the point $\left(\frac{-u_1}{\lambda_1}, \frac{-v_1}{\lambda_2}, \frac{-w_1}{\lambda_3}\right)$. Then the equation (1) takes the form

$$\lambda_1 x^2 + \lambda_2 y^2 + \lambda_3 z^2 + d_1 = 0$$

If $d_1 \neq 0$, the equation represent a central quadric and if $d_1 = 0$, it represents a cone.

Case II : Rank of $A = 2$.

In this case one $\lambda_1, \lambda_2, \lambda_3$ is zero. Let $\lambda_3 = 0$. Let us transfer the origin to $\left(\frac{-u_1}{\lambda_1}, \frac{-v_1}{\lambda_2}, 0\right)$. Then (1) takes the form.

$$\lambda_1 x^2 + \lambda_2 y^2 + 2\,w_1 z + d_1 = 0$$

If $w_1 \neq 0$, we transfer the origin to the point $\left(0, 0, \frac{-d_1}{2w_1}\right)$ so that the equation takes the form

$$\lambda_1 x^2 + \lambda_2 y^2 + 2w_1 z = 0$$

If $w_1 = 0$, the equation represents a pair of planes if $d_1 = 0$, and a hyperbolic or elliptic cylinder if $d_1 \neq 0$,

Case III : In this case two of $\lambda_1, \lambda_2, \lambda_3$ are zero. Let $\lambda_2 = \lambda_3 = 0$. Let us transfer the origin to $\left(\frac{-u_1}{\lambda_1}, 0, 0\right)$. Then the equation (1) takes the form

$$\lambda_1 x^2 + 2v_1 y + 2\,w_1 z + d_1 = 0 \qquad \ldots (2)$$

Let at least one of v_1 and w_1 be nonzero. Let $v_1 \neq 0$. Then we transfer the origin to $\left(0, \frac{-d_1}{2v_1}, 0\right)$ so that ther equation (2) takes the form

$$\lambda_1 x^2 + 2v_1 y + 2w_1 z = 0.$$

By the orthogonal transformation $X = PX'$, where

$$P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{v_1}{k} & \frac{-w_1}{k} \\ 0 & \frac{w_1}{k} & \frac{v_1}{k} \end{bmatrix}, \qquad \mathrm{k} = \sqrt{v_1^2 - w_1^2}$$

the equation reduces to $\lambda_1 x^2 + 2ky = 0$. This represents a parabolic cylinder.

When $v_1 = w_1 = 0$, the equation represents a pair of coincident planes if $d_1 = 0$ and a pair of parallel planes if $d_1 \neq 0$.

EXAMPLE

Ex. 1. Reduce the equation $x^2 + y^2 + z^2 - 2xy - 2yz + 2zx + x - 4y + z + 1 = 0$ into canonical form and determine the nature of the quadric.

Let $A = \begin{bmatrix} 1 & -1 & 1 \\ -1 & 1 & -1 \\ 1 & -1 & 1 \end{bmatrix}$, $\quad B = [1, -4\ 1]$, $\quad X = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$

Then the equation is written in matrix form as

$$X^T AX + BX + I = 0$$

The eigenvalues of A are 3, 0, 0.

The eigenvectors correrponding to the eigen value 3 are

$$k\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, k \neq 0$$

The eigen vectors corrsponding to the eigen value 0 are

$$c\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + d\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \qquad c, d \neq 0$$

Taking $c = 1$, $d = 0$, one eigen vector is $\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$.

If the other eigen vector orthogonal to $\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ is

$$c\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + d\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} c \\ c+d \\ d \end{bmatrix}, \text{ then}$$

$$c + c + d = 0 \qquad\qquad \text{or, } 2c + d = 0$$

We can take $c = 1$, $d = -2$, so the three multually orthogonal eigen vectors are

$$\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \text{ and } \begin{bmatrix} 1 \\ -1 \\ -2 \end{bmatrix}$$

and the corresponding orthonormal vectors are

$$\begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{-1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}, \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \\ \end{bmatrix}, \begin{bmatrix} \frac{1}{\sqrt{6}} \\ \frac{-1}{\sqrt{6}} \\ \frac{-2}{\sqrt{6}} \end{bmatrix}$$

Let $\qquad P = \dfrac{1}{\sqrt{6}}\begin{bmatrix} \sqrt{2} & \sqrt{3} & 1 \\ -\sqrt{2} & \sqrt{3} & -1 \\ \sqrt{2} & 0 & -2 \end{bmatrix}$

Then $\qquad P^T AP = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ and $BP = \begin{bmatrix} \frac{6}{\sqrt{3}} & \frac{-3}{\sqrt{2}} & \frac{3}{\sqrt{6}} \end{bmatrix}$

By the orthogonal transformation $X = PX'$. where $X' = \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix}$ the given equation reduces to

$$3x'^2+\frac{6}{\sqrt{3}}x'-\frac{3}{\sqrt{2}}y'+\frac{3}{\sqrt{6}}z'+1=0$$

or, $$3\left(x'+\frac{1}{\sqrt{3}}\right)^2-\frac{3}{\sqrt{2}}y'+\frac{3}{\sqrt{6}}z'=0$$

Now, using the transformation

$$x''=x'+\frac{1}{\sqrt{3}}$$
$$y''=y'$$
$$z''=z'$$

The equation becomes $$3x''^2-\frac{3}{\sqrt{2}}y''+\frac{13}{\sqrt{6}}z''=0$$

Let us apply orthogonal transformation

$$x''=x_1$$
$$y''=y_1\ cos\ \theta-z_1\ sin\ \theta$$
$$z''=y_1\ sin\ \theta+z_1\ cos\ \theta$$

such that $-\frac{3}{\sqrt{2}}\cos\theta+\frac{3}{\sqrt{6}}\sin\theta=0$. This gives $\theta=\frac{\pi}{3}$

Then the equation is trinsformed to $3x_1^2+\sqrt{6z_1}=0$

This is the canonical form of the quadric, which represents a parablic cylinder,

EXERCISES

1. Find the symmetric matrix which corresponds to each of the following quadratic form :

   (i) $q(x, y) = 4x^2 - 6xy - 7y^2$

   (ii) $q(x, y, z) = 3x^2 + 4xy - y^2 + 8xz - 6yz + z^2$

   (iii) $q(x, y) = xy + y^2$

2. Reduce the following quadratic forms to their normal forms. Find the rank and signature of each.

   (i) $2x^2 + 5y^2 + 10z^2 + 4xy + 12yz + 6zx$

   (ii) $2x^2 + 3y^2 + 4z^2 - 4xy + 4yz$

3. Show that the following quadratic form is positive definite.

$5x^2 + y^2 + 5z^2 + 4xy - 8xz - 4yz$

4. Show that the quadratic form

$x_1^2 + x_2^2 + x_3^2 + 3x_2x_3$ is indefinite.

5. Reduce the equation

$$7x^2 - 2xy + 7y^2 - 16x + 16y - 8 = 0$$

into canonical form and determine the nature of the conic.

6. Reduce the equation

$$x^2 + 2yz + 4x + 2y + 2z + 5 = 0$$

into canonical form and determine the nature of the quadric.

7. Reduce the equation $2x^2 - 2xy - 2yz + 2zx - x - z = 0$ into canonical form and determine the nature of the quadic.

8. Determine an orthogonal matrix P that reduces the quadratic form $2x_1^2 + 4x_1x_2 + 5x_2^2$ to a diagonal form.

BIBLIOGRAPHY

1. Finkbeiner, D. T. : Introduction to Matrices and Linear Transformations, D, B. Taraporevala sons & Co Pvt Ltd.
2. Kurosh, A : Higher Algebra, MIR publishers.
3. Apostol, T. M. : Linear Algebra, John Wiley & sons
4. Lipschutz, S : Linear Algebra, Schaum's Series
5. Mapa, S. K. : Higher Algebra (Abstract and Linear) Asoke Prakasan.
6. Ramachandra Rao, A & Bhimasankaram, P : Linear Algebra, McGraw Hill (New Delhi)
7. Krishnamurthy, V, Mainra, VP & Arora, J. L. : An Introduction to Linear Algebra, Affiliated East-West press Pvt Ltd.
8. Campbell, H G : Linear Algebra with Applications, Appleton-century-crofts Educational Division.
9. Hoffman, K & Kunze, R : Linear Algebra, Prentice Hall of India Pvt Ltd.
10. Hadly, G : Linear Algebra, Narosa
11. Datta, K B : Matrix and Linear Algebra
12. Chatterjee, B C : Linear Algebra, Das Gupta & Co.

মানুষের জ্ঞান ও ভাবকে বইয়ের মধ্যে সঞ্চিত করিবার যে একটা প্রচুর সুবিধা আছে, সে কথা কেহই অস্বীকার করিতে পারে না। কিন্তু সেই সুবিধার দ্বারা মনের স্বাভাবিক শক্তিকে একেবারে আচ্ছন্ন করিয়া ফেলিলে বুদ্ধিকে বাবু করিয়া তোলা হয়।

— *রবীন্দ্রনাথ ঠাকুর*

ভারতের একটা mission আছে, একটা গৌরবময় ভবিষ্যৎ আছে, সেই ভবিষ্যৎ ভারতের উত্তরাধিকারী আমরাই। নূতন ভারতের মুক্তির ইতিহাস আমরাই রচনা করছি এবং করব। এই বিশ্বাস আছে বলেই আমরা সব দুঃখ কষ্ট সহ্য করতে পারি, অন্ধকারময় বর্তমানকে অগ্রাহ্য করতে পারি, বাস্তবের নিষ্ঠুর সত্যগুলি আদর্শের কঠিন আঘাতে ধূলিসাৎ করতে পারি।

— *সুভাষচন্দ্র বসু*

Any system of education which ignores Indian conditions, requirements, history and sociology is too unscientific to commend itself to any rational support.

— *Subhas Chandra Bose*

Price : ₹ 225.00

(Not for sale to the Students of NSOU)

Published by : Netaji Subhas Open University, DD-26, Sector-I, Salt Lake, Kolkata-700 064 &
Printed at : The Saraswati Printing Works, 2, Guru Prosad Chowdhury Lane, Kolkata 700 006