RINGS AND LINEAR ALGEBRA

(DSMAT31/DBMAT31)

(BSC, BA MATHS - III)

ACHARYA NAGARJUNA UNIVERSITY

CENTRE FOR DISTANCE EDUCATION

NAGARJUNA NAGAR,

GUNTUR

ANDHRA PRADESH

LESSON - 1

RINGS AND INTEGRAL DOMAINS

1.1 Objectives of the Lesson:

To learn the difinitions of algebraic structures such as a ring, a field and an integral domain and to study some examples of these structures and thier basic properties.

1.2 Structure

1.3 Introduction:

In this lesson we define an algebraic structure called a ring. We also derive the concepts of a field and an internal domain. We learn some basic properties of a ring. Using the basic properties, we prove some theorems on rings and fields.

1.4.1 Definition of a Ring :

A ring is a nonempty set R together with two binary operations + and . called addition and multiplication respectively such that

i) $a,b \in R \Rightarrow a+b \in R$

ii) $(a+b)+c = a+(b+c), \forall a,b,c \in R$

iii) $\exists \; o \in R$ such that $a+o = o+a = a, \forall a \in R$

iv) $a \in R \Rightarrow \exists -a \in R$ such that $a+(-a) = 0 = (-a)+a$

v) $a+b=b+a, \forall a,b \in R$.

vi) $a,b \in R \Rightarrow a.b \in R$

vii) $(a.b).c = a.(b.c), \forall a,b,c \in R$

viii) $a.(b+c) = a.b + a.c$ and $(b+c).a = b.a + c.a \; \forall a,b,c \in R$

Note:

1.4.1

1) In a Ring R the binary operation '+' is called addition and '.' is called multiplication.

2) We usually write a b instead of a. b

3) R is a ring if i) (R,+) is an abelian group

ii) (R, .) is a semigroup

iii) Multiplication is both left and right distributive over addition.

4) The element O in R is called the zero element of R

5) Sometimes ring R is denoted by (R, +, .) The additive inverse of an element 'a' is denoted by '-a'.

1.4.2 Examples

1) Let $R=\{0\}$ and +, . be the operations defined by O + O = O and O . O = O, then (R, +,.) is a ring. This ring is called zero Ring or Null Ring .

2) Z is the set of all integers and + and . are usual addition and multiplication respectively. Then (Q, +, .) is a ring.

3) Q is the set of all rational numbers and + and . are usual addition and multiplication respectively. Then (Q, +, .) is a ring.

4) The set of all real numbers $\mathbb{R}$ is a ring w.r.t usual addition and multiplication.

5) The set of all complex numbers C is a ring w.r.t usual addition and multiplication.

6) Let $n > O$ be an integer. If $a,b \in Z$ and n/(a-b) (i.e., n divides (a-b)) then a is said to be congruent to b modulo n. This is denoted by $a \equiv b$ (mod n). Congruence modulo n is an equivalence relation on Z. Denote the equivalence class of an integer a by $\overline{a}$. Note that $\overline{a} = \overline{b}$ iff $a \equiv b(\bmod n)$. Given $a \in Z$, there are integers q and r, with $o \le r < n$, such that a = nq + r. Hence a - r = nq and $a \equiv r(\bmod n)$. Therefore $\overline{a} = \overline{r}$. Since a was arbitrary and $o \le r < n$, it follows that every equivalence class must be one of $\overline{O}, \overline{1}, \overline{2}, ..\overline{(n-1)}$. However these n equivalence

classes are distinct. For if $O \le i < \bar{r} < n$, then $O < \bar{r} - i < n$ and $n \times (\bar{r} - i)$. Thus $i \not\equiv \dot{r} (\bmod n)$ and hence $\bar{i} \not\equiv \bar{r}$. Therefore, there are exactly n equivalence classes. Let Z_n denote the set of all equivalence classes. Then $Z_n = \left\{ \bar{O}, \bar{1}, \bar{2}, .. \overline{(n-1)} \right\}$. Define addition and multiplication by $\bar{a} + \bar{b} = \overline{a+b}$ and $\bar{a}.\bar{b} = \overline{ab}$. Then Z_n together with these operations is a ring. Z_n is called the ring of integers modulo n. It is customary to denote the elements in Z_n as $0,1,.,...,n-1$ rather than $\bar{O}, \bar{1}, ..., \overline{n-1}$. This notation will be used whenever convenient. Addition and multiplication in Z_n are sometimes written as $+_n$ and $\times_n$ respectively.

1.4.3 Ring with Unity: A ring $(R,+,.)$ is said to be a ring with unity if there exists $1 \in R$ such that $a.1 = 1.a = a \forall a \in R$

1.4.4 The ring of intiger $(Z,+,.)$ is a ring with unity.

1.4.5 Note: If R is a ring with unity 1, then 1 is called unity or multiplicative identity or smply an identity element.

1.4.6. commutative Ring: A ring $(R,+,.)$ is said to be a commutative ring if a.b = b.a, $\forall a,b \in R$, i.e. Multiplication is commntative in R.

1.4.7 Example : The ring of integer $(Z,+,.)$ is a commntative ring.

1.4.8 Note:

i. If a Ring R has no multiplicative identity then the ring R is said to be a ring without unity.

ii. A Ring R is said to be non-commutative ring if R is not commutative.

1.4.9 Example:

i. The set of all Even intigers is a commutative ring without unity under usual addition and multiplication.

ii.

1.4.10 Note: If $(R,+,.)$ is ring then (R,+) is an abelian group. Therefore we have.

i) The zero element (additive identity) of R is unique and $a + 0 = a, \forall a \in R$

ii) For $a \in R$ the additive inverse -a is unique and a + (-a) = 0

iii) For $a \in R$, - (- a) = a

iv) For $O \in R$, - 0 = 0

v) For $a,b \in R, -(a+b) = -a-b$

vi) For $a,b,c \in R, a+b = a+c \Rightarrow b = c$

and $b+a = c+a \Rightarrow b = c$

vii) The unity element 1 is unique, if R has unity element 1.

1.4.11 Elementary Properties of Rings:

Theorem: If R is a ring then for $a,b,c \in R$

i) $oa = ao = o$

ii) $a(-b) = (-a)b = -(ab)$

iii) $(-a)(-b) = ab$

iv) $a(b-c) = ab - ac$

Proof: i) $oa = (o+o)a$ $\because o+o = o$

$= oa + oa$ By right distributive law

$\Rightarrow$ $oa + o = oa + oa$ Since 'o' is the additive Indentity

$\Rightarrow$ $o = oa$ By Left cancellation Law.

III ly $ao = a(o+o)$ $\because o+o = o$

$= ao + ao$ By left distributive law

$\Rightarrow$ $ao + o = ao + ao$ $\because o$ is the additive identity

$\Rightarrow$ $o = ao$ By left cancellation law.

$\therefore oa = ao = o$

ii) To show that $a(-b) = -(ab) = (-a)b$

We have $a(-b) + ab = a(-b+b)$ by left distributive law

$= ao$

$= o$ by (i)

$\therefore a(-b) + ab = o \Rightarrow a(-b) = -ab$

III ly $(-a)b + ab = (-a+a)b$ by right distributive law

$= ob$

$= o$ by (i)

$\therefore (-a)b + ab = o \Rightarrow (-a)b = -(ab)$

iii) $(-a)(-b) = -((-a)b)$ by (ii)

$= -(-(ab))$ by ii

$= ab$

iv) $a(b-c) = a[b+(-c)]$

= $ab + a(-c)$ by left distributive law.

= $ab + (-ac)$ by (ii)

= $ab - ac$

1.1.12 Note: If R is a ring with unity element and $a \in R$, then

i) $(-1)a = -a = a(-1)$

ii) $(-1)(-1) = 1$

1.1.13 Definition: Idempotent Element :- Let R be a ring. An element $a \in R$ is said to be an idempotent element if $a^2 = a$.

1.4.14 Example :

1) In the ring of Integers (Z, +, .), o and 1 are idempotent elements.

2) In the ring $(Z_6, +_6, \times_6), 0, 1, 3, 4$ are idempotent elements.

1.4.15 Boolean Ring: A ring R is said to be a Boolean ring if every element of R is an idempotent element i.e., $a^2 = a, \forall a \in R$.

1.4.16 Theorem: If R is a Boolean ring then

i) $a + a = 0, \forall a \in R$

ii) $a + b = 0 \Rightarrow a = b$

iii) R is commutative i.e. $ab = ba, \forall a, b \in R$.

Proof: Given that R is a Boolean Ring

i) Let $a \in R \Rightarrow a + a \in R$ $\because$ R is closed w.r.t '+'

$\Rightarrow (a+a)^2 = a+a$ Since $a^2 = a, \forall a \in R$

$\Rightarrow (a+a)(a+a) = a+a$

$\Rightarrow (a+a)a + (a+a)a = a+a$ by left distributive law

$\Rightarrow (aa+aa) + (aa+aa) = a+a$ by right distributive law

$\Rightarrow (a^2+a^2) + (a^2+a^2) = a+a$

$\Rightarrow (a+a) + (a+a) = a+a$ Since $a^2 = a, \forall a \in R$

$\Rightarrow (a+a) + (a+a) = (a+a) + 0$ Since O is the zero element of R.

$\Rightarrow a+a = 0$ by left cancellation law.

ii) Let $a,b \in R$ and $a+b=0$

$\Rightarrow a+b = b+b$ by i $b+b=0$

$\Rightarrow a=b$ by right cancellation law

iii) Let $a,b \in R \Rightarrow a+b \in R$ $\because$ R is closed under '+'

$\Rightarrow (a+b)^2 = a+b$ $\because$ R is Boolean Ring

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

$\Rightarrow (a+b)a + (a+b)b = a+b$ by left distributive law

$\Rightarrow (a^2+ba) + (ab+b^2) = a+b$ by right distributive law

$\Rightarrow a+ba+ab+b = a+b$ $\because$ R is a Boolean Ring

$\Rightarrow ba+ab = 0$ by left and right cancellation laws

$\Rightarrow ba = ab$ by ii.

$\therefore$ R is a commutative ring.

1.4.17 Definition : Nilpotent Element : Let R be a ring. An element $a \in R$ is said to be nilpotent element if there exists a positive integer 'n' such that $a^n = 0$.

1.4.18 Example: In the ring $(Z_9, +_9, \times_9)$, 3 is a nilpotent element.

1.5 Zero Divisors of a Ring:

Let R be a ring. An element $0 \neq a \in R$ is said to be a zero divisor if there exists a $b \neq 0 \in R$ such that ab = 0.

1.5.1 Note: 1) A ring R is said to have zero divisors if there exist $a, b \in R, a \neq 0, b \neq 0$ but $ab = 0$.

2) A ring R is said to have no zero divisors if $a, b \in R$ and $ab = 0 \Rightarrow a = 0$ or $b = 0$.

1.5.2 Examples: i) The ring $(Z_6, +_6, \times_6)$ has zero divisors 2, 3 and 4. For $2 \times_6 3 = 0, 3 \times_6 4 = 0$.

ii) The set R of all 2 x 2 matrices with real numbers as entries is a ring with zero divisors w.r. to addition and multiplication of matrices.

For $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$ are two non-zero elements of R but AB = O.

iii) The ring of integers Z is without zero divisors for $a, b \in Z$ and ab = 0 $\Rightarrow$ a = 0 or b = 0.

1.5.3 Cancellation Laws in a Ring: We say that cancellation Laws hold in a ring R if $a \neq 0, ab = ac \Rightarrow b = c$ and $a \neq 0, ba = ca \Rightarrow b = c$ for $a, b, c \in R$.

1.5.4 Theorem: A ring R is without zero divisors if and only if the cancellation laws hold in R.

Proof: Suppose R is a ring without zero divisors.

i.e. $a, b \in R$ and $ab = 0 \Rightarrow a = 0$ or $b = 0$

To prove that cancellation laws hold in R.

Let $a \neq 0, b, c \in R$ and $ab = ac$

Now $ab = ac \Rightarrow ab - ac = 0$

$\Rightarrow a(b - c) = 0$

$\Rightarrow b - c = 0$ $\because a \neq 0$ and R is without zero divisors

$\Rightarrow b = c$

Similarly we can show that $ba = ca \Rightarrow b = c$

$\therefore$ cancellation laws hold in R.

Suppose cancellation laws hold in R.

To prove that R has no zero divisors.

Let $a, b \in R$ and $ab = 0$

If possible suppose $a \neq 0$

Now $ab = 0 \Rightarrow ab = a0$

$\Rightarrow b = 0$ by cancellation law.

Similarly if $b \neq 0$, then

$ab = 0 \Rightarrow ab = 0b$

$\Rightarrow a = 0$ by cancellation law

$\therefore ab = 0 \Rightarrow$ either $a = 0$ or $b = 0$

Hence R has no zero devisors.

1.6.1 Definition: Integral Domain : A ring $(R,+,.)$ is said to be an integral domain if i) R is a commutative Ring with unity 1 $\neq 0$, ii) R is without zero divisors.

1.6.2. Example: The ring of integers $(Z,+,\times)$ is an integral domain.

1.6.3 Theorem: A commutative ring R with unity $1 \neq 0$ is an Integral domain if and only if cancellation laws hold in R.

Proof of this theorem follows from the previous theorem.

1.6.4 Definition: Invertible element (or) Unit: Let R be a ring with unity 1. A non zero element $a \in R$ is said to be inversible if there exists $b \in R$ such that $ab = ba = 1$. Here b is called multiplicative inverse of a and is denoted by a^{-1}.

1.6.5 Division Ring: A ring $(R,+,.)$ is said to be a division ring if i) R has unity $1 \neq 0$, ii) Every non zero element has multiplicative inverse.

1.6.6. Example: The ring of rational numbers $(Q,+,\times)$ is a division ring.

1.6.7 Theorem: A division ring has no zero divisors.

Proof: Let R be a division ring.

Let $a, b \in R$ and $ab = 0$

If possible let $a \neq 0$

$\Rightarrow \exists \ a^{-1} \in R$ such that $aa^{-1} = a^{-1}a = 1$ ($\because R$ is a division ring)

Now $ab = 0$

$\Rightarrow \quad a^{-1}(ab) = a^{-1}0$

$\Rightarrow \quad (a^{-1}a)b = 0$

$\Rightarrow \quad 1b = 0$

$\Rightarrow \quad b = 0$

Similarly if $b \neq 0$ we can prove that $a = 0$

$\therefore ab = 0 \Rightarrow$ either $a = 0$ or $b = 0$.

Hence R has no zero divisors.

1.6.8 Definition: Field : A ring R with atleast two elements is called a field if i) R is commutative, ii) R has unity, iii) Every non-zero element of R has multiplicative inverse.

1.6.9 Example: The rings of rational numbers, real numbers and complex numbers are fields.

1.6.10 Note: 1. A ring R is a field if $\left(R-\{0\},.\right)$ is an abelian group.

2. A Commutative division ring is a field.

1.6.11 Theorem: A field has no zero divisors.

Proof: A field is a commutative devision ring. Hence from 1.6.7 it follows that a field has no zero divisors.

1.6.12 Theorem: Every field is an integral domain.

Proof: Suppose $(F,+,.)$ is a field.

$\Rightarrow$ F is a commutative ring.

To show that F is an integral domain, it is enough to show that F has no zero divisors.

Let $a, b \in F$ and $ab = 0$

If possible let $a \neq 0$

$\Rightarrow \exists \ a^{-1} \in F$ such that $aa^{-1} = a^{-1}a = 1 \qquad \because F$ is a field

Now $ab = 0 \quad \Rightarrow \quad a^{-1}(ab) = a^{-1}0$

$\Rightarrow \quad (a^{-1}a)b = 0$

$\Rightarrow \quad 1b = 0$

$\Rightarrow \quad b = 0$

Similarly if $b \neq 0$ we can show that $a = 0$

Hence $ab = 0 \Rightarrow$ either $a = 0$ or $b = 0$

$\because F$ has no zero divisors

Now F is a commutative ring without zero divisors.

i.e. F is an integral domain.

1.6.13 Note: The converse of the above theorem is not true. i.e. an integral domain need not be a field.

Example: Ring of integers is an integral domain, but not a field.

Theorem: A finite integral domain is a field.

Proof: Let R be a finite integral domain with n elements. Now R is a commutative ring without zero divisors. To prove that R is a field it is enough to prove that R has i) Unity element, ii) Every non zero element of R has multiplicative inverse.

Let $R = \{a_1, a_2, ...a_n\}$ where a_i's are n distinct elements let $0 \neq a \in R$

Now the n products $aa_1, aa_\alpha ......aa_n \in R$

If possible let $aa_i = aa_j$ for $1 \leq i \leq n, 1 \leq j \leq n$ and $i \neq j$

$\Rightarrow aa_i - aa_j = 0$

$\Rightarrow a(a_i - a_j) = 0$

$\Rightarrow a_i - a_j = 0$ $\because a \neq 0$ and R is integral domain.

$\Rightarrow a_i = a_j$

This is a contradiction to the fact that a_i's are distinct.

$\therefore$ The n products $aa_1, aa_2 ......aa_n$ are distinct elements of R.

Existence of Identity:

Since $a \in R = \{aa_1, aa_2, ...aa_n\}$ it follows that $a = aa_i$ for some $a_i \in R$.

We claim that a_i is the identity element of R.

Let $b \in R \Rightarrow b = aa_j$ for some $a_j \in R$

Now $ba_i = (aa_j)a_i$

$= a(a_j a_i)$

$= a(a_i a_j)$ $\quad \therefore R$ is commutative.

$= (aa_i)a_j$

$= aa_j \quad \therefore \quad aa_i = a$

$= b$

$\therefore \quad a_i$ is the identity element of R.

Existence of Inverse: Since the identity $1 \in R$ we have $1 = aa_k$ for some $a_k \in R$.

$\therefore$ For $0 \neq a \in R$ there exists $a_k \in R$ such that $aa_k = 1$.

$\Rightarrow$ Every non zero element of R is inversible. Hence R is a field.

1.7.1 Integral Multiples: In a ring R we define $0a = 0$, where left hand zero is integer O and right hand side zero is the zero element of the ring R, where $a \in R$.

We define $na = a + a + ... + a$, n times where n is a positive integer.

$(-n)a = (-a) + (-a) + ... + (-a)$, n times

note that $(-n)a = n(-a) = -(na)$

1.7.2. Note: If m, n are integers and a,b are elements of a ring R then

i) $(m+n)a = ma + na$

ii) $m(na) = (mn)a$

iii) $m(a+b) = ma + mb$

iv) $m(ab) = (ma)b = a(mb)$

v) $(ma)(nb) = mn(ab)$

1.7.3 Integral Powers : If m, n are positive integers and a, b are elements of a ring R then

i) $a^m = a.a....a, m$ times

ii) $a^m.a^n = a^{m+n}$

iii) $(a^m)^n = a^{mn}$

1.7.4 Charecteristic of a Ring: The characteristic of a ring R is said to be 'n' if there is a positive integer n such that $na = 0, \forall a \in R$.

If there is no such positive integer 'n' then we say that the characteristic of the ring R is zero or infinite.

1.7.5 Example: 1. The characteristic of the ring $(Z_5, +_5, \times_5)$ is 5

2. The characteristic of the ring $(Z, +, .)$ is zero.

1.7.6 Theorem: The characteristic of a ring with unity is zero or n > 0, according as the order of the unity element is zero or n > 0 respectively regarded as member of the additive group of the ring.

Proof: Suppose R is a ring with unity element 1. Suppose order of 1 is zero, regarded as element of the group (R, +)

$\Rightarrow$ There is no positive integer n such that $n.1 = 0$

$\Rightarrow$ There is no positive integer n such that $na = 0 \forall a \in R$

$\Rightarrow$ Charecteristic of R is zero.

Suppose order of 1 = n > 0 regarded as element of the group (R, +)

$\Rightarrow$ n is the least positive integer such that $n.1 = 0$

Let $a \in R$

Now $na = a + a + .. + a$ n times

$= a(1 + 1 + .. + 1)$ n times

$= a(n.1)$

$= a0 \qquad \because n1 = 0$

$= 0$

$\therefore na = 0, \forall a \in R$

Further n is the least positive integer such that $na = 0, \forall a \in R$. Since if m < n then $m.1 \neq 0$. Hence charecteristic of a ring R is n.

1.7.7. Theorem: The characteristic of an integral domain is either zero or prime.

Proof: Suppose R is an Integral domain

Let P be the charecteristic of R

If P = 0 there is nothing to prove.

Suppose $P \neq 0$.

We claim that P is a prime number.

If possible assume that P is not a prime number.

$\Rightarrow$ P = mn where 1 < m, n < P.

Let $0 \neq a \in R$.

$\Rightarrow pa = 0$ $\because P$ is the charecteristic of R.

$\Rightarrow pa^2 = 0$

$\Rightarrow mna^2 = 0$

$\Rightarrow (ma)(na) = 0$

$\Rightarrow ma = 0$ or $na = 0$ $\because R$ has no zero divisors

Suppose $ma = 0$

Let $b \in R$

Now $ma = 0 \Rightarrow (ma)b = 0$

$\Rightarrow (mb)a = 0$

$\Rightarrow mb = 0$ $\because a \neq 0$ and R is without zero divisors.

$\therefore mb = 0, \forall b \in R$

A contradiction since m < p and Ch of R is P.

Similarly we can arrive at a contradiction if $na = 0$

$\therefore$ our assumption that P is not a prime is wrong

Hence P is a prime number.

1.7.8 Corollary : The charecteristic of a field is either zero or Prime.

Proof: Every field is an Integral domain.

$\therefore$ By theorem: 1.7.7. it follows that the charecteristic of a field is either zero or prime.

1.8.1 Definition: Subring: Let R be a ring. A non-empty subset S of R is said to be a subring of R if S it self is a ring relative to the same operations in R.

1.8.2 Example: 1. $(z,+,\times)$ is a subring $(Q,+,\times)$

1.8.3 Note: If R is a ring then $S=\{0\}$ where 0 is the zero element of R and $S=R$ are subrings of R. These subrings are called trivial or improper subrings of R. A subring other than the above two is called a non-trivial or proper subring of R.

1.8.4 Theorem: Let R be a ring. A non-empty subset S of a ring R is a subring iff $a,b\in S \Rightarrow a-b\in S$ and $ab\in S$.

Proof: Suppose R is a ring and S a non-empty subset of R. Suppose S is a subring of R.

Let $a,b\in S$

$b\in S \Rightarrow -b\in S$ $\because$ additive inverse exists.

Now $a,-b\in S \Rightarrow a-b\in S$ by closure law for addition. Also S is closed w.r. to multiplication.

$\therefore a,b\in S \Rightarrow ab\in S$.

Hence $a,b\in S \Rightarrow a-b\in S$ and $ab\in S$

Conversily suppose S is a non-empty subset of R such that $a,b\in S \Rightarrow i) a-b\in S$ and $ii) ab\in S$.

To show that S is a subring.

Existence of zero element: Since S is non-empty S has atleast one element say $a\in S$.

$a\in S \Rightarrow a-a\in S$ by the given condition.

$\Rightarrow 0\in S$

Existence of additive inverse:

$0\in S$ and $a\in S \Rightarrow 0-a\in S$

$\Rightarrow -a\in S$

Closure w.r. to addition:

Let $a,b\in S$

$b\in S \Rightarrow -b\in S$

Now $a,-b\in S \Rightarrow a-(-b)\in S$ by cndition (i).

$\Rightarrow a+b\in S$

Since elements of S are elements of R, associative law for addition and multiplication, commutative law for addition and distributive law holds good in S.

Also by condition (ii) S is closed w.r.to multiplication

$\therefore$ S is a ring and hence a subring of R.

1.8.5 Theorem: The intersection of two subrings of a ring R is a subring of R.

Proof: Let R be a ring and S_1, S_2 be two subrings of R.

Let $S = S_1 \cap S_2$. To show that S is a subring of R.

Since every subring contains the zero element we have $0 \in S_1$ and $0 \in S_2$.

$$\Rightarrow 0 \in S_1 \cap S_2 = S$$

$\therefore S$ is a non-empty subset of R.

Let $a, b \in S \Rightarrow a, b \in S_1 \cap S_2$

$\Rightarrow a, b \in S_1$ and $a, b \in S_2$

$a, b \in S_1 \Rightarrow a - b \in S_1$ and $ab \in S_1$ $\quad \because S_1$ is a subring

$a, b \in S_2 \Rightarrow a - b \in S_2$ and $ab \in S_2$ $\quad \because S_2$ is a subring

$\therefore a - b \in S_1 \cap S_2$ and $ab \in S_1 \cap S_2$

i.e. $a - b \in S$ and $ab \in S$

Hence S is a subring of R by Theorem 1.8.4.

1.8.6 Note: 1. The intersection of an arbitrary family of subrings is a subring.

2. Union of two subrings need not be a subring.

Let $S_1 = \{2n / n \in z\}$ and $S_2 = \{3n / n \in z\}$

S_1 and S_2 are two subrings of the ring of integers $(z, +, \times)$.

$$S_1 \cup S_2 = \{-6, -4, -3, -2, 0, 2, 3, 4, 6, 8, 9, ...\}$$

Clearly $2, 3 \in S_1 \cup S_2$

But $3 - 2 = 1 \notin S_1 \cup S_2$

Hence $S_1 \cup S_2$ is not a subring.

1.8.7 Theorem: Let S_1, S_2 be two subrings of a ring R.

then $S_1 \cup S_2$ is a subring iff $S_1 \subseteq S_2$ or $S_2 \subseteq S_1$

(OR)

Union of two subrings of a ring R is again a subring iff one is contained in the other.

Proof: Let R be a ring and S_1, S_2 be two subrings of a ring R.

Suppose $S_1 \leq S_2$ or $S_2 \leq S_1$

If possible assume that $S_1 \not\leq S_2$ or $S_2 \not\leq S_1$

$S_1 \not\leq S_2 \Rightarrow$ there is an element $a \in S_1$ and $a \notin S_2$ $\rightarrow 1$

$S_2 \not\leq S_1 \Rightarrow$ there is an element $b \in S_2$ and $b \notin S_1$ $\rightarrow 2$

$a \in S_1 \Rightarrow a \in S_1 \cup S_2$

$b \in S_2 \Rightarrow b \in S_1 \cup S_2$

Now $a, b \in S_1 \cup S_2$ and $S_1 \cup S_2$ is a subring

$\Rightarrow a + b \in S_1 \cup S_2$

$\Rightarrow a + b \in S_1$ or $a + b \in S_2$

If $a + b \in S_1$ then $a + b - a \in S_1$ Since $a \in S_1$ and S_1 is a subring.

$\Rightarrow b \in S_1$

A contradiction to (2).

If $a + b \in S_2$ then $a + b - b \in S_2$ Since $b \in S_2$ and S_2 is a subring.

$\Rightarrow a \in S_2$

A contradiction to (1)

$\therefore a + b \notin S_1$ and $a + b \notin S_2$

$\Rightarrow a+b \notin S_1 \cup S_2$

$\therefore S_1 \cup S_2$ is not a subring.

Which is a contradiction to the hypothesis.

$\therefore$ Our assumption $S_1 \nleq S_2$ and $S_2 \nleq S_1$ is wrong.

Hence either $S_1 \nleq S_2$ or $S_2 \nleq S_1$

Conversily suppose $S_1 \nleq S_2$ or $S_2 \nleq S_1$

Suppose $S_1 \nleq S_2 \Rightarrow S_1 \cup S_2 = S_2$ which is a subring.

If $S_2 \nleq S_1 \Rightarrow S_1 \cup S_2 = S_1$ which is a subring.

$\therefore$ either $S_1 \leq S_2$ or $S_2 \leq S_1$ we have $S_1 \cup S_2$ is a subring.

1.8.8 Definition: Subfield: Let $(F,+,.)$ be a field. A non-empty subset K of F is said to be a subfield of F if K itself is a field w.r.to the same operations in F.

1.8.9 Example: The set of all rational numbers $(Q,+,.)$ is a subfield of $(R,+,.)$

1.8.10 Theorem: Let $(F,+,.)$ be a field. A non-empty subset K of F is a subfield of F iff

$i) a,b \in K \Rightarrow a-b \in K$

$ii) a \in K, 0 \neq b \in K \Rightarrow ab^{-1} \in K$ and

$iii)\ 1 \in K$

Proof: Let F be a field and

K be a non-empty subset of F.

Suppose K is a subfield of F.

Let $a,b \in K$

$b \in K \Rightarrow -b \in K$ since additive inverse exists in a field.

Now $a,-b \in K \Rightarrow a+(-b) \in K$ by closure axiom for addition.

$\Rightarrow a-b \in K$

Let $a, 0 \neq b \in K$

$0 \neq b \in K \Rightarrow b^{-1} \in K$ since multiplicative inverse exists in a field.

Now $a, b^{-1} \in K \Rightarrow ab^{-1} \in K$ by closure axiom for multiplication.

Clearly $1 \in K$

$\therefore$ conditions (i), (ii) and (iii) hold good.

Suppose the condition (i) $a, b \in K \Rightarrow a - b \in K$

(ii) $a, 0 \neq b \in K \Rightarrow ab^{-1} \in K$ hold good.

From condition (i) it follows that $(K, +)$ is a sub group of $(F, +)$ from (ii) and (iii) it follows that $(K - \{0\}, .)$ is a sub group of $(F - \{0\}, .)$.

Since $K \not\subseteq F$, commutative law for addition and multiplication and distributive laws hold for elements of K.

Hence k is a subfield of F.

1.9. Summary :

In this lesson we learnt the definitions of Ring, Integral domain, Field, Subring and Subfield. We proved some theorems on Rings, Integral domains and fields.

1.10 Technical Terms:

i) Ring

ii) Commutative Ring

iii) Ring with Unity

iv) Idempotent Element

v) Boolean Ring

vi) Nilpotent Element

vii) Zero Divisors

viii) Integral Domain

ix) Unit or Invertible Element.

x) Division Ring or Skew Field

xi) Field

xii) Subring

xiii) Subfield

Xiv) Charecteristic of a Ring.

1.11 Model Questions:

1.11.1 The set M of all n X n matrices with entries as real numbers is a non-commutative ring with unity with respect to addition and multiplicatin of matrices.

Solution: Let M be the set of all n x n matrices with entries as real numbers.

i) We know that sum of the n x n matrices and product of two n x n matrices is again a n x n matrix.

$\therefore$ M is closed w.r. to addition and multiplication of matrics. i.e. $A, B \in M \Rightarrow A + B \in M$ and $AB \in M$.

ii) We know that addition of matrices is associative and commutative.

$$\therefore (A+B)+C = A+(B+C), \forall A, B, C \in M$$

$$A+B = B+A, \forall A, B, \in M$$

iii) The null matrix $O_{n\times n} \in M$ and $A+0 = A, \forall A \in M$

iv) To each $A \in M$ there is $-A \in M$ and $A+(-A) = O_{n\times n}$

v) Multiplication of matrices is Associative.

$$\therefore (AB)C = A(BC) \forall A, B, C \in M$$

vi) We know that multiplication of matrices is distributive over addition.

$$\therefore A(B+C) = AB + AC$$

$$(B+C)A = BA + CA, \forall A, B, C \in M$$

$\therefore M$ is aring.

The unit matrix $I_{n\times n} \in M$ and $IA = AI = A, \forall A \in M$

$\therefore M$ is a ring with unity I.

But we know that multiplication of matrices is not commutative.

Hence M is a non-commutative ring with unity.

1.11.2 The set $z(i) = \{x + iy / x, y \in z\}$ of Gaussian integers is a commutative ring with unity w.r. to addition and multiplication of complex numbers.

Solution: Given $z(i) = \{x + iy / x, y \in z\}$

Let $a = x_1 + iy_1$ and $b = x_2 + iy_2 \in z(i)$

$\Rightarrow x_1, y_1, x_2, y_2 \in Z$

Now $a + b = (x_1 + iy_1) + (x_2 + iy_2)$

$= (x_1 + x_2) + i(y_1 + y_2)$

$\therefore a + b \in z(i)$ since $x_1 + x_2$ and $y_1 + y_2 \in Z$

$ab = (x_1 + iy_1)(x_2 + iy_2)$

$= (x_1x_2 - y_1y_2) + i(x_1y_2 + x_2y_1)$

$\therefore ab \in z(i)$

i.e. $z(i)$ is closed under addition and multiplication. We know that addition of complex numbers is associative and commutative.

$\therefore$ Associative law and commutative law for addition hold good in $z(i)$ since $z(i)$ is a subset of the set of complex numbers C.

Clearly $0 = 0 + i0 \in z(i)$ and

$a + 0 = (x_1 + iy_1) + (0 + i0) = x_1 + iy_1 = a$

$\therefore O$ is the zero element of $z(i)$

$a \in z(i) \Rightarrow -a = -x_1 - iy_1 \in z(i)$ and

$a + (-a) = (x_1 + iy_1) + (-x_1 - iy_1) = 0$

$\therefore -a$ is the additive inverse of a.

We know that multiplication of complex numbers is associative and commutative.

$\therefore a(bc) = (ab)c$ and $ab = ba \ \forall \ a, b, c \in z(i)$ since $Z(i) \subseteq C$

Also multiplication of complex numbers is distributive over addition.

$\therefore a(b + c) = ab + ac \ \forall \ a, b, c \in z(i)$ since $Z(i) \subseteq C$

clearly $1 = 1 + 0i \in z(i)$ and $a1 = (x_1 + iy_1)(1 + 0i) = a$

$\therefore Z(i)$ is a commutative ring with unity.

1.11.3 If $Q(\sqrt{2})=\left\{a+b\sqrt{2}:a,b\in Q\right\}$ Then $Q(\sqrt{2})$ is a field with respect to addition and multiplication of real numbers.

Solution: $Q(\sqrt{2})=\left\{a+b\sqrt{2}:a,b\in Q\right\}$

Let $x=a+b\sqrt{2}$ and $y=c+d\sqrt{2}\in Q(\sqrt{2})$

$\Rightarrow a,b,c,d\in Q$

Now $x+y=a+b\sqrt{2}+c+d\sqrt{2}$

$=(a+c)+(b+d)\sqrt{2}$

$\therefore x+y\in Q(\sqrt{2})\quad \because a+c,b+d\in Q$

$xy=(a+b\sqrt{2})(c+d\sqrt{2})$

$=(ac+2bd)+\left(ad+bc\right)\sqrt{2}$

$\therefore xy\in Q(\sqrt{2})\ \because ac+2bd,ad+bc\in Q$

$\therefore Q(\sqrt{2})$ is closed under addition and multiplication.

We know that addition of real numbers is associative and commutative. Since $Q(\sqrt{2})\subseteq R$. We have $(x+y)+z=x+(y+z)$ and $x+y=y+z\ \forall\ x,y,z\in Q(\sqrt{2})$

clearly $0=0+0\sqrt{2}\in Q(\sqrt{2})$ and $x+0=x,\forall x\in Q(\sqrt{2})$

$x\in Q(\sqrt{2})\Rightarrow -x=-a-b\sqrt{2}\in Q(\sqrt{2})$ and $x+(-x)=0$

$\therefore -x$ is the additive invirse of x.

Also multiplication of real numbers is associative, commutative and is distributive over addition.

$\therefore\ x(yz)=(xy)z,xy=yz,x(y+z)=xy+xz$, for all $x,y,z\in Q(\sqrt{2})$ since $Q(\sqrt{2})\subseteq R$

clearly $1=1+0\sqrt{2}\in Q(\sqrt{2})$ and $x.1=(a+ib)(1+0\sqrt{2})=x$

$\therefore$ 1 is the multiplicative identity.

Let $0 \neq x = a + b\sqrt{2} \in Q(\sqrt{2})$

$\Rightarrow a \neq 0$ or $b \neq 0$.

Now $\frac{1}{x} = \frac{1}{a+b\sqrt{2}} = \frac{a-b\sqrt{2}}{a^2-2b^2} = \frac{a}{a^2-2b^2} - \frac{b}{a^2-2b^2}\sqrt{2} \in Q(\sqrt{2})$

Since $a, b \in Q$ and $a^2 - 2b^2 \neq c$

$\{a^2 - 2b^2 = 0 \Leftrightarrow a^2 = 2b^2 \Leftrightarrow a = \sqrt{2}b$ but $a \in Q \therefore a^2 - 2b^2 = 0 \Leftrightarrow a = 0$ and $b = 0\}$

$\therefore$ every non-zero element of $Q(\sqrt{2})$ has multiplicative inverse. Hence $Q(\sqrt{2})$ is a field.

1.11.4 $Z_p = \{0,1,2,3,...........p-1\}$ where P is a prime. Z_p is a field w.r.t the addition modulo P and multiplication modulo P.

Solution: $Z_p = \{0,1,2,3,...........p-1\}$

Let $a, b \in Z_p$

$a \oplus b =$ remainder, where $a + b$ is divided by P.

$a \odot b =$ remainder, where ab is divided by P.

Clearly $a \oplus b$ and $a \odot b \in z_p$

$\therefore z_p$ is closed under addition and multiplication modulo P.

Let $a, b, c \in z_p$

$(a \oplus b) \oplus c = r_1 \oplus c$ where $a + b = q_1 p + r_1, 0 \leq r_1 \leq p$

$= r_2$ where $r_1 + c = q_2 p + r_2, 0 \leq r_2 \leq p$

Now $a + b + c = q_1 p + r_1 + c = q_1 p + q_2 p + r_2 = (q_1 + q_2) p + r_2$

$\therefore r_2$ is the remainder where $a + b + c$ is divided by P

$a \oplus (b \oplus c) = a \oplus s_1$ where $b + c = m_1 p + s_1, 0 \leq s_1 \leq p$

$= s_2$ where $a + s_1 = m_2 p + s_2, 0 \leq s_2 \leq p$

Now $a+b+c = a+m_1 p+s_1 = m_1 p+a+s_1 = m_1 p+m_2 p+s_2$

$$= (m_1+m_2)p+s_2$$

$\therefore s_2$ is the remainder when $a+b+c$ is divided by P.

$$\therefore r_1 = s_2 \Rightarrow (a\oplus b)\oplus c = a\oplus(b\oplus c)$$

ie addition modulo P is associative.

Clearly $0 \in z_p$

$$a\oplus 0 = a$$

$\therefore 0$ is the zero element of z_p

Let $a \in z_p$

if $a=0$ then additive inverse of a is itself.

If $a \neq 0$ then $0 \le a \le p \Rightarrow 0 \ge -a \ge -p$

$$\Rightarrow p \ge p-a \ge p-p \Rightarrow p \ge p-a \ge 0$$

$$\therefore p-a \in z_p$$

Now $a\oplus(p-a) =$ Remainder when $a+(p-a)$ is divided by P.

$$=0$$

$\therefore p-a$ is the additive inverse of $a \neq 0$

Clearly $a\oplus b =$ remainder when $a+b$ is divided by P.

= remainder when $b+a$ is divided by P. $p\{\because a+b=b+a\}$

$$= b\oplus a$$

$\therefore$ Addition modulo P is commutative.

$(a\odot b)\odot c = r_1 \odot c$ where $ab = q_1 p + r_1 \quad 0 \le r_1 \le p$

$= r_2$ where $r_1 c = q_2 p + r_2 \quad 0 \le r_2 \le p$

Now $abc = (q_1 p + r_1)c = q_1 pc + r_1 c = q_1 pc + q_2 p + r_2$

$= (q_1 c + q_2) p + r_2$

$\therefore$ r_2 is the remainder when abc is divided by P.

$a \odot (b \odot c) = a \odot s_1$ where $bc = m_1 p + s_1 \quad 0 \le s_1 \le p$

$= s_2$ where $as_1 = m_2 p + s_2 \quad 0 \le s_2 \le p$

Now $abc = a(m_1 p + s_1 = am_1 p + as_1 = am_1 p + m_2 p + s_2$

$= (am_1 + m_2) p + s_2$

$\therefore s_2$ is the remainder when abc is divided by P.

$\therefore r_2 = s_2 \Rightarrow (a \odot b) \odot c = a \odot (b \odot c)$ i.e., multiplication modulo P is associative.

clearly $1 \in z_p$ and $a \odot 1 = 1 \odot a = a, \forall a \in z_p$

$\therefore$ 1 is the multiplicative identity.

Clearly $a \odot b = b \odot a, \forall a, b \in z_p$ since $ab = ba$

$\therefore$ Multiplication modulo P is commutative.

Hence z_p is a commutative ring with unity.

Suppose $a \odot b = 0$ where $a, b \in z_p$

$\Rightarrow 0$ is the remainder when ab is divided by P.

$\Rightarrow p$ divides ab

$\Rightarrow \frac{p}{a}$ or $\frac{p}{b}$ since P is prime.

$\Rightarrow a = 0$ or $b = 0$ since $0 \le a \le p$ and $0 \le b \le p$

$\therefore z_p$ has no zero divisors.

$\therefore z_p$ is an integral domain.

Since z_p is a finite integral domain z_p is a field.

Remark: the ring z_p described above is the same as the ring defined at 1 & 2 exampled. (upto isomorphism).

1.11.5 Give example of a division ring which is not a field.

Solution: Let $A = \left\{ \begin{pmatrix} a+ib & c+id \\ -c+id & a-id \end{pmatrix} / a,b,c,d \in R \right\}$

We know that the set M of all matrices with complex numbers as entries is a non-commutative ring with unity.

Clearly A is a non-empty subset of M.

Let $X, Y \in A \Rightarrow X = \begin{pmatrix} a_1+ib_1 & c_1+id_1 \\ -c+id_1 & a_1-ib_1 \end{pmatrix}$ and

$$Y = \begin{pmatrix} a_2+ib_2 & c_2+id_2 \\ -c_2+id_2 & a_2-ib_2 \end{pmatrix}$$

Now $X - Y = \begin{pmatrix} (a_1-a_2)+i(b_1-b_2) & (c_1-c_2)+i(d_1-d_2) \\ -(c_1-c_2)+i(d_1-d_2) & (a_1-a_2)-i(b_1-b_2) \end{pmatrix} \in A$

$$XY = \begin{pmatrix} a_1+ib_1 & c_1+id_1 \\ -c_1+id_1 & a_1-ib_1 \end{pmatrix}\begin{pmatrix} a_2+ib_2 & c_2+id_2 \\ -c_2+id_2 & a_2-ib_2 \end{pmatrix}$$

$$= \begin{pmatrix} a_1a_2-b_1b_2-c_1c_2-d_1d_2+i(a_1b_2+a_2b_2+c_1d_2-c_2d_1) & a_1c_2-b_1d_2+a_2c_1+b_2d_1+i(a_1d_2+b_1c_2+d_1a_2-b_2c_1) \\ -c_1a_2-d_1b_2-a_1c_2+b_1d_2+i(-c_1b_2+a_2d_1+a_1d_2+b_1c_2) & -c_1c_2-d_1d_2+a_1a_2-b_1b_2+i(c_1d_2+c_2d_1-a_1b_2-a_2b_1) \end{pmatrix}$$

$$\therefore \; XY \in A$$

Hence $X, Y \in A \Rightarrow X+Y$ and $XY \in A$

$\therefore A$ is a subring of M and hence a ring.

Since M is a non commutative ring with unity $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

For $0 \neq X \in A$ where $X = \begin{pmatrix} a+ib & c+id \\ -c+id & a-ib \end{pmatrix}$

$X \neq 0 \Rightarrow$ one of a, b, c, d must be not equal to zero

$\therefore |X| = a^2 + b^2 + c^2 + d^2 \neq 0$ and hence X is invertible.

Hence every non-zero element of A has a multiplicative inverse.

$\therefore A$ is a division ring.

Since A is non-commutative it is not a field.

1.11.6 Show that the charecteristic of a Boolean ring R is 2.

Solution: Let R be a Boolean ring

$\Rightarrow a^2 = a, \forall a \in R$

Let $a \in R \Rightarrow a + a \in R$

$\Rightarrow (a+a)^2 = a + a$ Since R is a Boolean ring

$\Rightarrow (a+a)(a+a) = a + a$

$\Rightarrow a^2 + a^2 + a^2 + a^2 = a + a$

$\Rightarrow a + a + a + a = a + a$

$\Rightarrow a + a = 0$

$\Rightarrow 2a = 0$

$\therefore$ Charecteristics of a Boolean ring is 2.

1.11.7 If the charecteristic of a ring R is 2 and $a, b \in R$ commute then $(a+b)^2 = a^2 + b^2$.

Solution: Let R be a ring with charecteristic 2.

Let $a, b \in R$ such that $ab = ba$

Now $(a+b)^2 = (a+b)(a+b)$

$= a^2 + ab + ba + b^2$

$= a^2 + ab + ab + b^2$

$= a^2 + 2ab + b^2$

$= a^2 + 0 + b^2$ Since ch. of R is 2, 2a = 0 $\forall a \in R$

$= a^2 + b^2$

1.11.8 If R is a ring and $C(R) = \{x \in R : ax = xa \forall a \in R\}$ then show that C(R) is a subring of R. C(R) is called the centre of R.

Solution: Let R be a ring.

$C(R) = \{x \in R / ax = xa \forall a \in R\}$

clearly $0 \in R$ and $ao = oa = a \forall a \in R$

$\therefore 0 \in C(R)$

$\therefore C(R)$ is a non-empty subset of R.

Let $x, y \in C(R)$

$\Rightarrow ax = xa$ and $ay = ya \forall a \in R$

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

$= xa - ya$

$= (x-y)a \quad \forall a \in R$

$a(xy) = (ax)y = (xa)y = x(ay) = x(ya)$

$= (xy)a, \forall a \in R$

$\therefore x - y, xy \in C(R)$

Hence $C(R)$ is a subring of R. by 1.8.4.

1.11.9 The set of all those integers which are multiples of a given integer say m is a subring of the ring of integers.

Solution: Let $S = \{mx / x \in z\}$ where m is a given integer.

Clearly $o \in z \Rightarrow m.o = o \in s$

$\therefore s$ is a non-empty subset of Z.

Let $a, b \in S$

$\Rightarrow a = mx$ and $b = my$ for $x, y \in z$

Now $a-b=mx-my$

$=m(x-y)$

$\therefore a-b \in s \because x-y \in z$

also $ab = mx.my$

$= m.mxy$

$\therefore ab \in s$ since $mxy \in z$

Hence S is a subring of Z by 1.8.4.

1.11.10 Show that 0 and 1 are the only idempotent elements of an integral domain.

Solution: Let R be an integral domain.

Suppose $a \in R$ is an idempotent element.

$\Rightarrow a^2 = a$

$\Rightarrow a^2 - a = 0$

$\Rightarrow a(a-1) = 0$

$\Rightarrow a = 0$ or $a-1=0$ since R is an Integral domain.

$\Rightarrow a = 0$ or $a=1$

$\therefore$ 0 and 1 are the only idempotent elements of an integral domain.

2. Exercise:

1. Prove that the set of even integer is a ring with respect to usual addition and multiplication.

2. In the set of integers addition $\oplus$, multiplication $\otimes$ are defined as $a \oplus b = a+b-1$ and $a \otimes b = a+b-ab \quad \forall a,b \in z$. Prove that $(z,\oplus,\otimes)$ is a commutative ring.

3. Is $R=\left\{a\sqrt{2} / a \in Q\right\}$ a ring under ordinary addition and multiplication of real numbers?

4. Is the set of all pure imaginary numbers $\{iy / y \in R\}$ a ring with respect to addition and multiplication of complex numbers?

5. Let $Q=\left\{\alpha_0+\alpha_1\bar{i}+\alpha_2\bar{j}+\alpha_3\bar{k} / \alpha_0,\alpha_1,\alpha_2,\alpha_3 \in R\right\}$ where $\bar{i},\bar{j},\bar{k}$ are the quoternion units $(i^2=j^2=k^2=ijk=-1, ij=-ji=k, jk=-kj=i, ki=-ik=j)$. Show that Q is a division ring with respect to addition and multiplication defined as follows. For $X,Y \in Q$ where

$X = \alpha_0 + \alpha_1 \bar{i} + \alpha_2 \bar{j} + \alpha_3 \bar{k}$, $Y = \beta_0 + \beta_1 i + \beta_2 j + \beta_3 k$.

$X + Y = (\alpha_0 + \beta_0) + (\alpha_1 + \beta_1)\bar{i} + (\alpha_2 + \beta_2)\bar{j} + (\alpha_3 + \beta_3)\bar{k}$

$XY = (\alpha_0\beta_0 - \alpha_1\beta_1 - \alpha_2\beta_2 - \alpha_3\beta_3) + (\alpha_0\beta_1 - \alpha_1\beta_0 - \alpha_2\beta_3 - \alpha_3\beta_2)\bar{i} + (\alpha_0\beta_1 + \alpha_2\beta_0 + \alpha_3\beta_1 - \alpha_1\beta_3)\bar{j} + (\alpha_0\beta_3 + \alpha_1\beta_2 + \alpha_1\beta_2 - \alpha_2\beta_1)\bar{k}$

This ring is called ring of quaternions. This ring is not commutative.

6. Find the Zero divisors of the ring $(z_{12}, +_{12}, \times_{12})$

7. Prove that $\pm 1, \pm i$ are the only four units of the ring of Gaussian Integers $z(i)$.

8. Show that the set of matrices $\begin{pmatrix} a & b \\ 0 & c \end{pmatrix}$ where $a, b, c \in z$ is a subring of the ring of 2 X 2 matrices whose elements are integers.

9. If R is a division ring show that $C(R) = \{x \in R / xa = ax, \forall a \in R\}$ is a field.

10. Is $s = \{\bar{1}, \bar{3}, \bar{5}\}$ a subring of the ring z_6 of resedue classes modulo 6 under addition and multiplication of resedue classes.

11. Find the centre of the ring of 3 X 3 matrices $M_3(R)$, where R is the field of reals.

- Smt. K. Ruth

LESSON - 2

IDEALS AND HOMOMORPHISM OF RINGS

2.1 Objective of the Lesson:

To learn the definition of ideals, types of ideals, homomorphism of rings and theorems related to them.

2.2 Structure

2.3 Introduction:

In this lesson we define Right Ideal, Left Ideal Ideal, Principal Ideal, Prime Ideal, Maximal Ideal and Quotient rings. We also define homomorphism of rings and Kernal of a homomorphism.

2.4.1 Right Ideal: Let R be a ring. A nonempty subset I of R is said to be a right ideal of R if

i) $a, b \in I \Rightarrow a - b \in I$

ii) $a \in I, r \in R \Rightarrow ar \in I$

2.4.2 Left Ideal: Let R be a ring : A nonempty subset I of R is said to be a left ideal of R if

i) $a, b \in I \Rightarrow a - b \in I$

ii) $a \in I, r \in R \Rightarrow ra \in I$

2.4.3 Example: 1. The set $I = \left\{ \begin{bmatrix} 0 & a \\ 0 & b \end{bmatrix} \middle/ a, b \in z \right\}$ is a left ideal of the ring of 2 x 2 matrices with integers as entries but not a right ideal.

2. The set $I = \left\{ \begin{bmatrix} a & b \\ 0 & 0 \end{bmatrix} \middle/ a, b \in z \right\}$ is a right ideal of the ring of 2 x 2 matrices with integers as entries but not a left ideal.

2.4.4 Ideal: Let R be a ring. A non-empty subset I of R is said to be an ideal of R if

i) $a, b \in I \Rightarrow a - b \in I$

ii) $a \in I, r \in R \Rightarrow ar$ and $ra \in I$

2.4.5 Example: 1. Let R be a ring. Then $I = \{0\}$ where 0 is the zero element of the ring R is an ideal of R.

For $0 - 0 \in I$ and $0.r = r.0 = 0 \in I$, $\forall r \in R$

This ideal is called the Null ideal or zero ideal.

2. The ring R itself is an ideal of R.

For $a, b \in R \Rightarrow a - b \in R$ and $a \in R, r \in R \Rightarrow ar$ and $ra \in R$.

This ideal is called the unit ideal.

The above two ideals are called trivial or improper ideals of R.

3. The set of even integers $I = \{2n / n \in z\}$ is an ideal of the ring of integers.

2.4.6 Note: 1. Every ideal of a ring is a subring of the ring.

Let I be an ideal of a ring R.

Let $a, b \in I \Rightarrow a - b \in I$ since I is an ideal.

$b \in I \Rightarrow b \in R$

$\therefore a \in I, b \in R$ and I is an ideal $\Rightarrow ab \in I$.

Hence by 1.8.4 I is a subring.

2. The converse of the above note is note true. i.e. every subring of a ring need not be an ideal of the ring.

We know that $(z,+,\times)$ is subring of $(Q,+,\times)$ .....

We have $2 \in z$ and $\frac{1}{4} \in Q$

But $2 \times \frac{1}{4} = \frac{1}{2} \notin Z$

$\therefore Z$ is not an ideal of Q.

2.4.7 Corollary: Let I be an ideal of a ring R with unity element 1. If $1 \in I$ then $I = R$.

Proof: Let I be an ideal of a ring R with unity element 1. Suppose $1 \in I$.

Clearly $I \subseteq R$ ..... (1)

Let $x \in R$

Now $1 \in I, x \in R$ and I is an ideal.

$\Rightarrow 1.x \in I$

$\Rightarrow x \in I$

$\therefore R \subseteq I$ ..... (2)

From (1) and (2) $I = R$

2.4.8 Theorem: A field has no proper ideals.

(OR)

The only ideals of a field are $\{0\}$ and itself

Proof: Let F be a field.

We know that $I = \{0\}$ and $I = F$ are ideals of F by 2.4.5

Let $I \neq \{0\}$ be an ideal of F.

Since $I \neq \{0\}$ there is $0 \neq a \in I$

$\Rightarrow a \in F$ since $I \subseteq F$

$\Rightarrow \exists\ a^{-1} \in F$ such that $aa^{-1} = a^{-1}a = 1$ since F is a field.

Now $a \in I, a^{-1} \in F$ and I is an ideal

$\Rightarrow aa^{-1} \in I$

$\Rightarrow 1 \in I$

Now by 2.4.7 $I = F$

$\therefore$ Every non-zero ideal of F is equal to F itself.

Hence the only ideals of a field are $\{0\}$ and itself.

2.4.9 Algebra of Ideals:

Theorem: The intersection of two ideals of a ring R is an ideal of R.

Proof: Let R be a ring and I_1, I_2 be two ideals of R

Take $I = I_1 \cap I_2$

Clearly $0 \in I_1$ and $0 \in I_2 \Rightarrow 0 \in I_1 \cap I_2 = I$

$\Rightarrow$ I is a non-empty subset of R.

Let $a, b \in I$ and $r \in R$.

$a, b \in I \Rightarrow a, b \in I_1 \cap I_2$

$\Rightarrow a, b \in I_1$ and $a, b \in I_2$

$\Rightarrow a - b \in I_1$, ra and $ar \in I_1$ since I_1 is an ideal of R and

$a - b \in I_2$, ra and $ar \in I_2$ since I_2 is an ideal of R.

Now $a - b \in I_1$, $a - b \in I_2$ $ra \in I_1$, $ra \in I_2$, $ar \in I_1$, $ar \in I_2$

$\therefore$ $a - b \in I_1 \cap I_2 = I$, $ra \in I_1 \cap I_2 = I$, $ar \in I_1 \cap I_2 = I$.

Hence I is an ideal of R.

2.4.10. Note: Union of two ideals need not be an ideal.

2.4.11 Theorem: Union of two ideals of a ring R is an ideal if and only if one is contained in the other.

Proof: Let R be a ring and I_1, I_2 be two ideals of a ring R.

Suppose $I_1 \cup I_2$ is an ideal of R.

To prove that $I_1 \subseteq I_2$ or $I_2 \subseteq I_1$.

If possible assume that $I_1 \not\subseteq I_2$ or $I_2 \not\subseteq I_1$

$I_1 \not\subseteq I_2 \Rightarrow$ There is $a \in I_1$ and $a \notin I_2$ .... (1)

$I_2 \not\subseteq I_1 \Rightarrow$ There is $b \in I_2$ and $b \notin I_1$ .... (2)

$a \in I_1 \Rightarrow a \in I_1 \cup I_2$

$b \in I_2 \Rightarrow b \in I_1 \cup I_2$

Now $a, b \in I_1 \cup I_2 \Rightarrow a - b \in I_1 \cup I_2$ since $I_1 \cup I_2$ is an ideal.

$a - b \in I_1 \cup I_2 \Rightarrow a - b \in I_1$ or $a - b \in I_2$

If $a - b \in I_1$ Then $a - (a - b) \in I_1$ since $a \in I_1$ and I_1 is an ideal

$\Rightarrow b \in I_1$

A contradiction to (2) $\Rightarrow a - b \notin I_1$ ...... (3)

If $a - b \in I_2$ Then $b + a - b \in I_2$ since $b \in I_2$ and I_2 is an ideal

$\Rightarrow a \in I_2$

A contradiction to (1) $\therefore$ $a - b \notin I_2$ .... (4)

$\therefore a - b \notin I_1 \cup I_2$ from (3) and (4)

$\Rightarrow I_1 \cup I_2$ is not an ideal.

This is a contradiction.

Hence our assumption is wrong.

$\therefore$ either $I_1 \subseteq I_2$ or $I_2 \subseteq I_1$

Conversily suppose $I_1 \subseteq I_2$ or $I_2 \subseteq I_1$

If $I_1 \subseteq I_2$ Then $I_1 \cup I_2 = I_2$ which is an ideal

If $I_2 \subseteq I_1$ Then $I_1 \cup I_2 = I_1$ which is an ideal

$\therefore$ Either $I_1 \subseteq I_2$ or $I_2 \subseteq I_1$, $\Rightarrow I_1 \cup I_2$ is an ideal

2.4.12 Theorem: If I_1 and I_2 are two ideals of a ring R then $I_2 + I_2 = \{a+b/a \in I_1$ and $b \in I_2\}$ is an ideal of R.

Proof: Let R be a ring and I_2, I_2 be two ideals of R.

Suppose $I_1 + I_2 = \{a+b/a \in I_1$ and $b \in I_2\}$

Clearly $0 \in I_1$ and $0 \in I_2$

$\therefore 0 = 0 + 0 \in I_1 + I_2$

$\Rightarrow I_2 + I_2$ is a non-empty subset of R.

Let $x, y \in I_1 + I_2$ and $r \in R$

Then $x = a_1 + b_1$ where $a_1 \in I_1$ and $b_1 \in I_2$

$y = a_2 + b_2$ where $a_2 \in I_1$ and $b_2 \in I_2$

Now $x - y = (a_1 + b_1) - (a_2 + b_2)$

$= (a_1 - a_2) + (b_1 - b_2)$

Since $a_1, a_2 \in I_1, b_1, b_2 \in I_2$ and I_1, I_2 are ideals

We have $a_1 - a_2 \in I_1$ and $b_1 - b_2 \in I_2$

$\therefore x - y \in I_1 + I_2$ .... (1)

$xr = (a_1 + b_1)r = a_1 r + b_1 r \in I_1 + I_2$

Now $rx = r(a_1 + b_1) = ra_1 + rb_1 \in I_1 + I_2$ since $a_1 \in I_1, b_1 \in I_2, r \in R$ and I_1, I_2 are ideals.

$\therefore rx$ and $xr \in I_1 + I_2$ .... (2)

From (1) and (2) $I_1 + I_2$ is an ideal.

2.4.13 Note: $I_1 + I_2$ is an ideal containing both I_1 and I_2.

By the above theroem $I_1 + I_2$ is an ideal.

Clearly $x \in I \Rightarrow x+0 \in I_1 + I_2$ since $0 \in I_2$

$\Rightarrow x \in I_1 + I_2$

$\therefore I_1 \subseteq I_1 + I_2$

$y \in I_2 \Rightarrow 0 + y \in I_1 + I_2$ since $0 \in I_1$

$\Rightarrow y \in I_1 + I_2$

$\therefore I_2 \subseteq I_1 + I_2$

$I_1 \subseteq I_1 + I_2, I_2 \subseteq I_1 + I_2 \Rightarrow I_1 \cup I_2 \subseteq I_1 + I_2$

2.4.14 Definition: Ideal Generated by a subset:

Let R be a ring and S a subset of R. An ideal I of R is said to be generated by S.

If i) $S \subseteq I$ ii) I^1 is an ideal of R and $S \subseteq I^1 \Rightarrow I \subseteq I^1$

Ideal generated by S is denoted by $\langle S \rangle$

2.4.15 Note: Ideal generated by S is the smallest ideal containing S.

2.4.16 Theorem: If I_1 and I_2 are two ideals of ring R then $I_1 + I_2$ is the ideal generated by $I_1 \cup I_2$ i.e. $I_1 + I_2 = \langle I_1 \cup I_2 \rangle$

Proof: Let R be a ring and I_1, I_2 be two ideals of a ring R.

By 2.4.12 $I_1 + I_2$ is an ideal of R.

By 2.4.13 $I_1 \cup I_2 \subseteq I_1 + I_2$ .... (1)

Let I^1 be an ideal of R and $I_1 \cup I_2 \subseteq I^1$

To prove that $I_1 + I_2 \subseteq I^1$

Let $x \in I_1 + I_2$

$\Rightarrow x = a + b$ where $a \in I_1$ and $b \in I_2$

$a \in I_1 \Rightarrow a \in I_1 \cup I_2 \Rightarrow a \in I^1$ since $I_1 \cup I_2 \subseteq I^1$

$b \in I_2 \Rightarrow b \in I_1 \cup I_2 \Rightarrow b \in I^1$

$\therefore a, b \in I^1$ and I^1 is an ideal.

$\Rightarrow a + b \in I^1$

$\Rightarrow x \in I^1$

Hence $I_1 + I_2 \subseteq I^1$ .... (2)

From (1) and (2) $I_1 + I_2 = \langle I_1 \cup I_2 \rangle$

2.5.1 Definition: Principal Ideal : Let R be a ring. An ideal I of R is said to be Principal ideal of R if I is generated by a single element of R. i.e. $I = \langle a \rangle$ where $a \in R$.

2.5.2 Example: (1) The null ideal $\{0\}$ of a ring R is a principal ideal.

(2) In a ring R with unity element, R itself is a principal ideal.

Definition: If a and b are nonzero elements of a ring R such that ab = 0, then a and b are called divisors of zero (or zero divisors)

Definition: A commutative ring R with identity $1 \neq 0$ and no zero divisors is called an integral domain.

2.5.3 Principal Ideal Domain: An integral domain R with unity is said to be a principal ideal domain if every ideal of R is a principal ideal.

2.5.4 Note: Every field is a principal ideal domain. We know that a field has only two ideal $\{0\}$ and itself by 2.4.8.

Clearly $\{0\}$ is generated by 0 i.e. $\{0\} = \langle 0 \rangle$

and F is generated by 1 i.e. $F = \langle 1 \rangle$

$\therefore$ The ideals of a field are principal ideals and hence a field is a principal ideal domain.

2.5.5. Theorem: The ring of integers is a principal ideal domain

Proof: Let $(z, +, .)$ be the ring of integers and I be an ideal of Z.

If $I = \{0\}$ then I is a principal ideal.

Suppose $I \neq \{0\}$

$\Rightarrow \exists 0 \neq a \in I$

$\Rightarrow -a \in I$ since I is an ideal.

Since $a, -a \in I, I$ contains atleast one postive integer.

Let I^+ be the set of all positive integers of I.

Then I^+ is non-empty.

$\therefore$ By well ordering principal I^+ has a least element say “b”.

We claim that $I = \langle b \rangle$

Let $x \in I$

Now x, b are integers and $b \neq 0$.

By division algorithm there exists $q, r \in z$

Such that $x = bq + r$ where $0 \leq r < b$.

Now $b \in I, q \in z$ and I is an ideal.

$\Rightarrow bq \in I$

$\therefore x - bq \in I$ since $x \in I$

$\Rightarrow r \in I$

Since b is the least positive integer in I and $0 \leq r < b.$ r can not be positive.

$\therefore r = 0$

$\Rightarrow x = bq, q \in z$

$\therefore I = \{bq / q \in z\} = \langle b \rangle$

Hence Z is a principal ideal domain.

2.6.1 Definition: Prime ideal: A proper ideal I of a commutative ring R is said to be a prime ideal if $a, b \in R$ and $a.b \in I \Rightarrow a \in I$ or $b \in I$

2.6.2 Example : In an integral domain, null ideal is prime.

Let R be an integral domain.

Suppose $a, b \in R$ and $ab \in \{0\}$

$\Rightarrow ab = 0$

$\Rightarrow a = 0$ or $b = 0$ $\because$ R has no zero divisors

$\Rightarrow a \in \{0\}$ or $b \in \{0\}$

$\therefore \{0\}$ is a prime ideal.

2.6.3 Definition: Maximal ideal: Let R be a ring and M is an ideal of R such that $M \neq R$. M is said to be a maximal ideal of R if for any ideal I of R such that $M \subseteq I \subseteq R$ then $I = M$ or $I = R$.

2.6.4 Theroem: An ideal M of the ring of integers Z is maximal iff M is generated by some prime number.

Proof: Let Z be the ring of integers.

We know that Z is a principal ideal domain (by 2.5.5)

Suppose M is a maximal ideal of Z. Then $M = \langle n \rangle$, where $n > 0$.

To prove that n is a prime number.

If possible assume that n is not prime.

$\Rightarrow n = pq$ where p and q are integers such that $1 < p < n$ and $1 < q < n$.

Now $I = \langle p \rangle$ is an ideal of Z and $M \subseteq I \subseteq Z$

M is maximal $\Rightarrow I = M$ or $I = Z$

If $I = M$ then $\langle p \rangle = \langle n \rangle$

$\Rightarrow p = nm$ for some $m \in z$

$\Rightarrow p = pqm,$ since $n = pq$.

$\Rightarrow mq = 1$

$\Rightarrow m = 1$ and $q = 1$

This is a contradiction.

If $I = Z$ then $\langle p \rangle = \langle 1 \rangle$

$\Rightarrow p = 1$. Again this is a contradiction.

$\therefore$ Our assumption that n is not prime is wrong.

Hence $M = \langle n \rangle$ where n is a prime number.

Conversely suppose M is generated by a prime number

i.e. $M = \langle n \rangle$, where n is a prime number.

To prove that M is a maximal ideal of Z.

Let I be an ideal of Z such that $M \subseteq I \subseteq Z$.

I is an ideal of Z $\Rightarrow I = \langle m \rangle$ for since $m > 0$.

$M \subseteq I \subseteq Z \Rightarrow \langle n \rangle \subseteq \langle m \rangle \subseteq Z$

$\Rightarrow n \in \langle m \rangle$

$\Rightarrow n = mq, q \in z$

$\Rightarrow m = 1$ or $q = 1$, since n is prime.

If m = 1 then $\langle m \rangle = \langle 1 \rangle = Z$

$\Rightarrow I = Z$

If $q = 1$ then $n = m$

$\Rightarrow \langle n \rangle = \langle m \rangle$

$\Rightarrow M = I$

$\therefore$ Either $I = M$ or $I = Z$

$\therefore$ M is a maximal ideal of Z.

2.6.5 Note: In the ring of integers an ideal generated by a composite number is not maximal.

Eq: Let $M = \langle 6 \rangle$

Clearly $\langle 6 \rangle \subseteq \langle 3 \rangle \subseteq Z$ and $\langle 3 \rangle \neq \langle 6 \rangle$ and $\langle 3 \rangle \neq Z$

$\therefore M$ is not maximal.

2.7.1 Quotient Rings:

Coset: Let R be a ring and I an ideal of R. Then $r+I=\{r+a/a\in I\}$ is called a coset of I generated by r.

Since $(R,+)$ is an abelian group $r+I=I+r, \forall r\in R$, $r+I$ is also called the resedue class containing r.

2.7.2 Theorem: Let R be a ring and I an ideal of R.

Then $R/I=\{r+I/r\in R\}$ is a ring with respect to the addition and multiplication of cosets defined by $(a+I)+(b+I)=(a+b)+I$

$(a+I)(b+I)=ab+I$ for $a+I, b+I\in R/I$

Proof: Let R be a ring and I an ideal of R.

$R/I=\{r+I/r\in R\}$

Addition of resedue classes is well defined:

Suppose $a_1+I=a_2+I$ and $b_1+I=b_2+I$

$\Rightarrow a_1-a_2\in I$ and $b_1-b_2\in I$. Since from groups $H+a=H+b\Leftrightarrow a-b\in H$.

$\Rightarrow a_1-a_2+b_1-b_2\in I$. Since I is an ideal.

$\Rightarrow (a_1+b_1)-(a_2+b_2)\in I$

$\Rightarrow (a_1+b_1)+I=(a_2+b_2)+I$

$\Rightarrow (a_1+I)+(b_1+I)=(a_2+I)+(b_2+I)$

$\therefore$ Addition is well defined.

Multiplication is well defined:

Suppose $a_1+I=a_2+I$ and $b_1+I=b_2+I$

$\Rightarrow a_1\in a_2+I$ and $b_1\in b_2+I$

$\Rightarrow a_1=a_2+u_1$ and $b_1=b_2+u_2$ for some $u_1,u_2\in I$

Now $a_1b_1=(a_2+u_1)(b_2+u_2)$

$= a_2b_2 + a_2u_2 + u_1b_2 + u_1u_2$

$\Rightarrow a_1b_1 - a_2b_2 = a_2u_2 + u_1u_2$

$\Rightarrow a_1b_1 - a_2b_2 \in I$ since $u_1, u_2 \in I, a_2, b_2 \in R$ and I is an ideal.

$\Rightarrow a_1b_1 + I = a_2b_2 + I$

$\Rightarrow (a_1 + I)(b_1 + I) = (a_2 + I)(b_2 + I)$

$\therefore$ Multiplication is well defined.

Closure Law : Let $a + I, b + I \in R/I$

$\Rightarrow a, b \in R$

$\Rightarrow a + b \in R$

$\Rightarrow (a+b) + I \in R/I$

$\Rightarrow (a+I) + (b+I) \in R/I$

Associative law : Let $a+I, b+I, c+I \in R/I$

Now $[(a+I)+(b+I)]+(c+I) = ((a+b)+I)+(c+I)$

$= [(a+b)+c]+I$

$= [(a+(b+c)]+I$ since $a,b,c, \in R$

$= (a+I)+[(b+c)+I]$

$= (a+I)+[(b+I)+(c+I)]$

Existence of Identity:

Now $(0+I)+(a+I) = (0+a)+I = a+I, \forall a+I \in R/I$

$\therefore a+I = I$ is the zero element of R/I

Existence of Inverse:

Let $a+I \in R/I \Rightarrow a \in R$

$\Rightarrow -a \in R$

$\Rightarrow -a+I \in R/I$

Now $(a+I)+(-a+I)=a+(-a)+I$

$=0+I$

$\therefore -a+I$ is the additive inverse of $a+I$

Commutative Law : Let $a+I, b+I \in R/I$

$(a+I)+(b+I)=(a+b)+I=(b+a)+I$ since $a,b \in R$

$=(b+I)+(a+I)$

Closure with respect to multiplication:

Let $a+I, b+I \in R/I \Rightarrow a,b \in R$

$\Rightarrow ab \in R$ since R is a ring.

$\Rightarrow ab+I \in R/I$

$\Rightarrow (a+I)(b+I) \in R/I$

Associative Law for multiplication.

Let $a+I, b+I, c+I \in R/I$

$\left[(a+I)(b+I)(c+I\right]=(ab+I)(c+I)$

$=(ab)c+I$

$=a(bc)+I$ since $a,b,c \in R$

$=(a+I)(bc+I)$

$=(a+I)\left[(b+I)(c+I)\right]$

Distributive Law:

Let $a+I, b+I, c+I \in R/I$

$(a+I)\left[(b+I)+(c+I)\right]=(a+I)\left[(b+c)+I\right]$

$=a(b+c)+I$

$=(ab+ac)+I$ since $a,b,c \in R$

$= (ab+I)+(ac+I)$

$= (a+I)(b+I)+(a+I)(c+I)$

Similarly $\left[(b+I)+(c+I)\right](a+I)=(b+I)(a+I)+(c+I)(a+I)$

$\therefore R/I$ is a ring.

The ring R/I is called the quotient ring of R modulo I.

2.7.3 Note: If R is a commutative ring and I is an ideal of R then the Quotient ring R/I is also commutative.

Proof: Let R be a commutative ring and I an ideal of R.

Let $a+I, b+I \in R/I$

$(a+I)(b+I)=ab+I$

$=ba+I$ Since R is commutative.

$=(b+I)(a+I)$

$\therefore$ R/I is commutative.

2. If R is a ring with unity then the Quotient ring R/I is also a ring with unity.

Proof: Let R be a ring with unity and I an ideal of R.

$1 \in R \Rightarrow 1+I \in R/I$

Now $(1+I)(a+I)=1a+I=a+I$

$(a+I)(1+I)=a1+I=a+I$

$\therefore 1+I$ is the unity element of R/I

Hence R/I is a ring with unity.

2.7.4 Theorem: An ideal I of a commutative ring R with unity is prime iff the Quotient ring R/I is an integral domain.

Proof: Let R be a commutative ring with unity and I an ideal of R

$\Rightarrow$ The quotient ring R/I is a commutative ring with unity.

Suppose I is a prime ideal.

To show that the Quotient ring R/I is an integral domain.

Since R/I is a commutative ring it is enough to show that R/I has no zero divisors.

Let $a+I, b+I \in R/I$ and

$(a+I)(b+I) = 0+I$

$\Rightarrow ab+I = I$

$\Rightarrow ab \in I$

$\Rightarrow a \in I$ or $b \in I$. Since I is prime.

$\Rightarrow a+I = I$ or $b+I = I$

$\therefore R/I$ has no zero divisors.

Hence R/I is an integral domain.

Conversily suppose R/I is an integral domain.

To prove that I is a prime ideal.

Let $a, b \in R$ and $ab \in I$

$\Rightarrow ab+I = I$

$\Rightarrow (a+I)(b+I) = I$

$\Rightarrow a+I = I$ or $b+I = I$ Since R/I is an integral domain.

$\Rightarrow a \in I$ or $b \in I$

$\therefore I$ is a prime ideal.

2.7.5 Theorem: An ideal I of a commutative ring R with unity is maximal iff the quotient ring R/I is a field.

Proof: Let R be a commutative ring with unity and I an ideal of R.

$\Rightarrow$ The Quotient ring R/I is a commutative ring with unity.

Suppose I is a maximal ideal of R.

To prove that R/I is a field.

Since R/I is a commutative ring with unity it is enough to prove every non-zero element of R/I has multiplicative inverse.

Let $0+I \neq a+I \in R/I$

$a+I \neq 0+I \Rightarrow a \notin I$

We have $\langle a \rangle$ is a principal ideal of R.

$\therefore \langle a \rangle + I$ is an ideal of R and $I \subseteq \langle a \rangle + I \subseteq R$ Since $a \notin I$

Since I is a maximal ideal either $\langle a \rangle + I = I$ or $\langle a \rangle + I = R$

But $a \notin I \Rightarrow \langle a \rangle + I \neq I$ since $a \in \langle a \rangle + I$

$\therefore \langle a \rangle + I = R$

$1 \in R \Rightarrow 1 \in \langle a \rangle + I$

$\Rightarrow 1 = ar + x$ for some $r \in R$ and $x \in I$

$\Rightarrow 1 - ar = x \in I$

$\Rightarrow 1 + I = ar + I$ since $a + H = b + H \Leftrightarrow a - b \in H$

$\Rightarrow 1 + I = (a+I)(r+I)$

$\therefore 0 + I \neq a + I \in R/I \Rightarrow \exists$ $r + I \in R/I$ such that $(a+I)(r+I) = 1 + I$

Hence every non zero element of R/I has multiplicative inverse.

$\therefore$ R/I is a field.

Conversely suppose R/I is a field.

To prove that I is a maximal ideal.

Let I^1 be an ideal of R such that $I \subseteq I^1 \subseteq R$.

If possible assume that $I^1 \neq I$

$\Rightarrow$ there is $a \in I^1$ and $a \notin I$

$\Rightarrow a + I \neq I$

i.e. $a + I$ is a non-zero element of R/I.

$\Rightarrow$ There exists $b+I \in R/I$ such that $(a+I)(b+I)=1+I$ since R/I is a field.

$\Rightarrow ab+I=1+I$

$\Rightarrow 1-ab \in I$

$\Rightarrow 1-ab \in I^1 \quad \because \; I \subseteq I^1$

Now $a \in I^1$ and $b \in R, I^1$ is an ideal $\Rightarrow ab \in I^1$

$\therefore 1-ab+ab \in I^1$

$\Rightarrow 1 \in I^1$

$\Rightarrow I^1 = R$ by corollary 2.4.7

$\therefore I$ is a maximal ideal of R.

2.7.6 Note: (1) If R is a commutative ring with unity, then every maximal ideal is a prime ideal.

Let M be a maximal ideal of R.

$\Rightarrow R/M$ is a field by theorem 2.7.5

$\Rightarrow R/M$ is an integral domain

$\Rightarrow M$ is a prime ideal by theorem 2.7.4

2. The converse of the above need not be true i.e., every prime ideal of a commutative ring with unity need not be a maximal ideal.

For example in the ring of integers null ideal is prime but not maximal

For $\{0\} \subseteq \langle 2 \rangle \subseteq Z$ and $\langle 2 \rangle \neq \{0\}$ and $\langle 2 \rangle \neq Z$

2.8.1 Homomorphism of rings:

Definition: Let R and R' be two rings. A mapping $f: R \to R^1$ is said to be a homomorphism if i) $f(a+b) = f(a) + f(b)$ ii) $f(ab) = f(a)f(b)$ for all $a, b \in R$

2.8.2 Definition: Let R and R' be two rings. A mapping $f: R \to R^1$ is said to be

i) Monomorphism if *f* is a homomorphism and one-one

ii) Epimorphism if *f* is a homomorphism and onto

iii) Isomorphism if *f* is a homomorphism one-one and onto.

2.8.3 Definition: A homomorphism $f:R\to R$ of a ring R into itself is called an endomorphism.

2.8.4 Definition: An isomorphism $f:R\to R$ of a ring R into itself is called an automorphism.

2.8.5 Example: 1. Let R and R' be two rings $f:R\to R^1$ defined by $f(x)=0^1, \forall x\in R$ where 0^1 is the zero element of R is a homomorphism.

Solution: Let $a,b\in R$.

$$f(a+b)=0^1=0^1+0^1=f(a)+f(b)$$

$$f(ab)=0^1=0^1.0^1=f(a)f(b)$$

$\therefore f$ is a homomorphism.

This homomorphism is called Zero homomorphism.

2. Let R be a ring. Then the identity mapping $I:R\to R$ is an automorphism.

Solution : Let $a,b\in R$

$$I(a+b)=a+b=I(a)+I(b)$$

$$I(ab)=ab=I(a).I(b)$$

$\therefore I$ is a homomorphism.

We know that identity mapping is a bijection.

$\therefore I$ is an automorphism.

This is called identity homomorphism.

2.8.6 Note: If $f:R\to R^1$ is an isomorphism then we say R is isomorphic to R^1 and we write $R\cong R^1$.

2.8.7 Elementary Properties of Homomorphism:

Theorem: Let $f:R\to R^1$ be a homorphism of a ring R into a ring R^1 and $0,0^1$ be the zero elements of R and R' respectively then for $a,b\in R$.

i) $f(0)=0^1$ ii) $f(-a)=-f(a)$ iii) $f(a-b)=f(a)-f(b)$

Proof: $f:R\to R^1$ is a homomorphism.

i) $0\in R\Rightarrow f(0)R^1$

$f(0) = f(0+0)$ since $0+0=0$

$\Rightarrow f(0) = f(0) + f(0)$ since f is a homomorphism.

$\Rightarrow f(0) + 0^1 = f(0) + f(0)$ since 0^1 is the zero element of R^1.

$\Rightarrow 0^1 = f(0)$ by left cancellation law.

ii) $a \in R \Rightarrow \exists -a \in R$ such that $a + (-a) = 0$

$\therefore f[a + (-a)] = f(0)$

$\Rightarrow f(a) + f(-a) = 0^1$ by (i) and f is a homomorphism.

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

iii) Let $a, b \in R$

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

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

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

2.8.8 Homomorphic Image: Let $f : R \to R^1$ be a homomorphism of a ring R into a ring R^1. Then the image set $\{ f(x) / x \in R \}$ is called the homomorphic image of R and is denoted by $f(R)$.

2.8.9 Theorem: A homomorphic image of a ring is a ring.

Proof: Let $f : R \to R^1$ be a homomorphism and $f(R) = \{ f(x) / \ x \in R \}$ be the homomorphic image of R.

To prove that $f(R)$ is a ring.

Clearly $f(R) \subseteq R^1$ and $f(0) = 0^1 \in f(R)$

$\therefore f(R)$ is a non-empty subset of R^1.

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

$\Rightarrow a = f(x)$ and $b = f(y)$, where $x, y \in R$

$x, y \in R \Rightarrow x - y \in R$ and $xy \in R$

$\Rightarrow f(x-y) \in f(R)$ and $f(xy) \in f(R)$

$\Rightarrow f(x) - f(y) \in f(R)$ and $f(xy) \in f(R)$

$\Rightarrow a-b \in f(R)$ and $ab \in f(R)$

$\therefore f(R)$ is a subring of R^1

Hence $f(R)$ is a ring.

2.8.10 Note: A homomorphic image of a commutative ring is commutative.

Proof: Let $f;R \rightarrow R^1$ be a homomorphism from a commutative ring R into a ring R^1.

$f(R) = \{f(x)/ \; x \in R\}$ is the homomorphic image of R.

Let $a,b \in f(R) \Rightarrow a = f(x), b = f(y)$ where $x, y \in R$.

Now $ab = f(x)f(y)$

$= f(xy)$ since f is a homomorphism.

$= f(yx)$ since R is commutative.

$= f(y)f(x)$

$= ba$

$\therefore f(R)$ is commutative.

From 2.8.9 $f(R)$ is a ring $\Rightarrow f(R)$ is a commutative ring.

2.9.1 Kernal of a homomorphism:

Let $f: R \rightarrow R^1$ be a homomorphism from a ring R into a ring R^1. Then $\{x \in R/ \; f(x) = 0^1\}$ is called Kernel of f and is denoted by $\ker f$. 0^1 is the zero element of R^1.

2.9.2 Theorem: If $f: R \rightarrow R^1$ is a homomorphism of a ring R into a ring R^1 then $\ker f$ is an ideal of R.

Proof: Let $f: R \rightarrow R^1$ be a homomorphism of a ring R into a ring R^1.

$\ker f \; = \{x \in R/ \; f(x) = 0^1\}$ where 0^1 is the zero element of R^1.

Clearly $0 \in R$ and $f(0) = 0^1$

$\Rightarrow 0 \in \ker f$

$\therefore$ $\ker f$ is a nonempty subset of R.

Let $a, b \in \ker f$ and $r \in R$

$a, b \in \ker f \Rightarrow f(a) = 0^1$ and $f(b) = 0^1$

$\Rightarrow f(a) - f(b) = 0 - 0^1 = 0^1$

$\Rightarrow f(a-b) = 0^1$

$\therefore a - b \in \ker f$ ...... (1)

$f(ar) = f(a)f(r) = 0^1 f(r) = 0^1$

$f(ra) = f(r)f(a) = f(r)0^1 = 0^1$

$\therefore ar \in \ker f$ and $ra \in \ker f$ ..... (2)

From (1) and (2) $\ker f$ is an ideal of R.

2.9.3 Theorem: A homomorphism f from a ring R into a ring R^1 is a monomorphism iff $\ker f = \{0\}$.

Proof: Let $f : R \to R^1$ be a homomorphism from a ring R into a ring R^1.

Suppose f is a homomorphism.

To prove that $\ker f = \{0\}$

Let $a \in \ker f \Rightarrow f(a) = 0^1$

$\Rightarrow f(a) = f(0)$

$\Rightarrow a = 0$ since f is one-one

$\therefore$ $\ker f = \{0\}$.

Suppose $\ker f = \{0\}$.

To prove that f is one - one.

Suppose $a,b \in R$ and $f(a) = f(b)$

$\Rightarrow f(a) - f(b) = 0^1$

$\Rightarrow f(a-b) = 0^1$

$\Rightarrow a-b \in \ker f$

$\Rightarrow a-b = 0$ since $\ker f = \{0\}$.

$\Rightarrow a = b$

$\therefore f$ is one - one.

Hence f is a monomorphism.

2.9.4 Theorem: If I is an ideal of a ring R then there is an epimorphism f from R onto R/I such that $\ker f = \mathrm{I}$

Proof: Let R be a ring and I am ideal of R.

Define $f : R \to R/I$ as $f(a) = a + I, \forall a \in R$

f is well defined : Suppose $a,b \in R$ and $a = b$

$\Rightarrow a + I = b + I$

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

$\therefore f$ is well defined.

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

$f(ab) = ab + I = (a+I)(b+I) = f(a)(b), \forall a,b \in R$

$\therefore f$ is a homomorphism.

f is onto : Let $x + I \in R/I$

$\Rightarrow x \in R$

Now $f(x) = x + I$

$\therefore x + I \in R/I \Rightarrow \exists\ x \in R$ such that $f(x) = x + I$

Hence f is onto.

$\therefore\ f : R \to R/I$ is an epimorphism.

$\therefore \ker f = \{a \in R / f(a) = 0 + I\} = \{a \in R / f(a) = I\}$

$= \{a \in R / a + I = I\}$

$= \{a \in R / a \in I\}$

$= I$

2.9.5 Note: The above epimorphism is called canonical mapping or natural mapping.

2.10.1 Fundamental Theorem of Homomorphism:

If $f : R \to R^1$ is an epimorphism from a ring R into a ring R^1 then $R/\ker f \cong R^1$

(OR)

Every homomorphic image of a ring is isomorphic to some quotient ring.

Proof: Suppose $f : R \to R^1$ is an epimorphism from a ring R into a ring R^1.

Let $\ker f = I$

We know that $\ker f$ is an ideal of R.

$\therefore R/\ker f$ is a quotient ring.

Define $\phi : R/I \to R^1$ by

$\phi(a + I) = f(a) \forall a + I \in R/I$

ϕ is well defined and one - one :

Let $a + I, b + I \in R/I$

$a + I = b + I$

$\Leftrightarrow a - b \in I$

$\Leftrightarrow f(a - b) = 0^1$ since $I = \ker f$.

$\Leftrightarrow f(a) - f(b) = 0^1$

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

$\Leftrightarrow \phi(a+I)=\phi(b+I)$

$\therefore \phi$ is well defined and one - one.

ϕ is onto :

Let $y \in R^1$

since $f: R \to R^1$ is onto there is $x \in R$ such that

$f(x)=y$

$x \in R \Rightarrow x+I \in R/I$

$\therefore y \in R^1 \Rightarrow$ there is $x+I \in R/I$ and $\phi(x+I)=f(x)=y$

$\therefore \phi$ is onto.

ϕ is a homomorphism:

Let $a+I, b+I \in R/I$

$\phi\left[(a+I)+(b+I)\right]=\phi\left[(a+b)+I\right]=f(a+b)$

$= f(a)+(b)$

$=\phi\left[(a+I)+\phi(b+I)\right]$

$\phi\left[(a+I)(b+I)\right]=\phi\left[ab+I\right]=f(ab)$

$= f(a)f(b)$

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

$\therefore \phi$ is a homomorphism which is one-one and onto i.e. ϕ is an isomorphism from R/I to R^1.

$\therefore R/I \cong R^1$.

2.10.2 Definition: A ring R is said to be embedded in a ring R^1 if there exists a monomorphism from R into R^1.

2.10.3 Example: A subring S of a ring R can be embedded in ring R.

For $I: S \to R$ defined by $I(a)=a, \forall a \in s$ is a monomorphism from S into R.

2.10.4 Theorem: Every integral domain can be embedded in a field.

Proof: Let D be an integral domain.

Let $R = \{(a,b)/a,b \in D, b \neq 0\}$

Define $\sim$ on R as $(a,b) \sim (c,d) \Leftrightarrow ad = bc$

We claim that $\sim$ is an equivalence relation on R.

Let $(a,b) \in R \Rightarrow a,b \in D$ and $b \neq 0$

$\Rightarrow ab = ba$ Since D is commutative

$\Rightarrow (a,b) \sim (a,b)$

$\therefore \sim$ is reflexive.

Let $(a,b),(c,d) \in R$

Suppose $(a,b) \sim (c,d) \Rightarrow ad = bc$

$\Rightarrow da = cb$ Since D is commutative

$\Rightarrow cb = da$

$\Rightarrow (c,d) \sim (a,b)$

$\therefore \sim$ is symmetric.

Let $(a,b),(c,d),(e,f) \in R$

Suppose $(a,b) \sim (c,d)$ and $(c,d) \sim (e,f)$

$\Rightarrow ad = b$ and $cf = de$

$\Rightarrow (ad)f = (bc)f$

$\Rightarrow (ad)f = b(cf)$

$\Rightarrow (ad)f = b(de)$ Since $cf = de$

$\Rightarrow (da)f = (bd)e$ since D is commutative

$\Rightarrow d(af) = d(be)$

$\Rightarrow af = be$ by left cancellation law.

$\Rightarrow (a,b) \sim (e, f)$

$\therefore \sim$ is transitive.

Hence $\sim$ is an equivalence relation on R.

Let $\frac{a}{b}$ be the equivalence class containing (a,b) with respect to the equivalence relation $\sim$.

Let $F = \left\{ \frac{a}{b} \middle/ a, b \in D, a \neq 0 \right\}$ the set of all equivalence classes under $\sim$.

We define addition + and multiplication . as $\frac{a}{b} + \frac{c}{d} = \frac{ad+bc}{bd}$ and $\frac{a}{b} . \frac{c}{d} = \frac{ac}{bd}, \forall \frac{a}{b}, \frac{c}{d} \in F$

To show that + and . are well defined.

Suppose $\frac{a}{b} = \frac{a_1}{b_1}$ and $\frac{c}{d} = \frac{c_1}{d_1}$

$\Rightarrow (a,b) \sim (a_1, b_1)$ and $(c,d) \sim (c_1, d_1)$

$[x] = [y]$ iff $x \sim y$

$\Rightarrow ab_1 = ba_1$ and $cd_1 = dc$

Now $(ad + bc)b_1 d_1 = adb_1 d_1 + bcb_1 d_1$

$$= ab_1 dd_1 + bb_1 cd_1$$

$$= ba_1 dd_1 + bb_1 dc_1$$

$$= a_1 d_1 bd + b_1 c_1 dd$$

$$= (a_1 d_1 + b_1 c_1) bd$$

$\therefore (ad + bc, bd) \sim (a_1 d_1 + b_1 c_1, b_1 d_1)$

$$\Rightarrow \frac{ad+bc}{bd} = \frac{a_1 d_1 + b_1 c_1}{b_1 d_1}$$

$$\Rightarrow \frac{a}{b} + \frac{c}{d} = \frac{a_1}{b_1} + \frac{c_1}{d_1}$$

$\therefore$ + is well defined.

Also $acb_1d_1 = ab_1cd_1$

$= ba_1dc_1$ Since $ab_1 = ba_1$ and $cd_1 = dc_1$

$= bda_1c_1$

$\therefore (ac, bd) \sim (a_1c_1, b_1d_1)$

$$\Rightarrow \frac{ac}{bd} = \frac{a_1c_1}{b_1d_1}$$

$$\Rightarrow \frac{a}{b} \cdot \frac{c}{d} = \frac{a_1}{b_1} \cdot \frac{c_1}{d_1}$$

$\therefore$. is well defined.

$\therefore$ + and . are binary operations on F.

We prove that $(F, +, .)$ is a field.

First we prove that if $x \neq 0$ the $\frac{a}{b} = \frac{ax}{bx}$ and

if $x \neq 0, y \neq 0$ the $\frac{o}{x} = \frac{o}{y}$ and $\frac{x}{x} = \frac{y}{y}$ .... (1)

We have $a(bx) = (ab)x = (ba)x = b(ax)$

$\Rightarrow (a, b) \sim (ax, bx)$

$$\Rightarrow \frac{a}{b} = \frac{ax}{bx}$$

Also $oy = o = xo \Rightarrow (o, x) \sim (o, y)$

$$\Rightarrow \frac{o}{x} = \frac{o}{y}$$

Also $xy = xy \Rightarrow (x, x) \sim (y, y)$

$$\Rightarrow \frac{x}{x} = \frac{y}{y}$$

Addition is associative:

Let $\frac{a}{b}, \frac{c}{d}, \frac{e}{f} \in F$

$$\left(\frac{a}{b}+\frac{c}{d}\right)+\frac{e}{f}=\frac{ad+bc}{bd}+\frac{e}{f}$$

$$=\frac{(ad+bc)f+bde}{bdf}$$

$$=\frac{adf+bcf+bde}{bdf}$$

$$=\frac{adf+b(cf+de)}{b(df)}$$

$$=\frac{a}{b}+\frac{cf+de}{df}$$

$$=\frac{a}{b}+\left(\frac{c}{d}+\frac{e}{f}\right)$$

$\therefore$ addition is associative.

<u>Existance of additive identity:</u>

If $x \neq 0$ then $\frac{o}{x} \in F$

Now $\frac{a}{b}+\frac{o}{x}=\frac{ax+bo}{bx}=\frac{ax}{bx}=\frac{a}{b}$ from (1)

$\frac{o}{x}+\frac{a}{b}=\frac{bo+ax}{xb}=\frac{ax}{bx}=\frac{a}{b} \quad \forall \frac{a}{b} \in F$

$\therefore \frac{o}{x}$ is the additive identity of F.

<u>Existance of Inverse:</u>

Let $\frac{a}{b} \in F \Rightarrow a, b \in D$ and $b \neq 0$

$\Rightarrow -a, b \in D$ and $b \neq 0$

$\Rightarrow \frac{-a}{b} \in F$

Now $\frac{a}{b} + \frac{(-a)}{b} = \frac{ab + b(-a)}{bb} = \frac{ab - ba}{bb} = \frac{ab - ab}{bb} = \frac{o}{bb} = \frac{o}{x}$ from (1)

$$\frac{-a}{b} + \frac{a}{b} = \frac{(-a)b + ab}{bb} = \frac{-ab + ab}{bb} = \frac{o}{bb} = \frac{o}{x}$$

$\therefore \frac{-a}{b}$ is the additive inverse of $\frac{a}{b}$.

<u>Addition is commutative:</u>

Let $\frac{a}{b}, \frac{c}{d} \in F$

$$\frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd} = \frac{bc + ad}{db} = \frac{cb + dc}{db} = \frac{c}{d} + \frac{a}{b}$$

$\therefore +$ is commutative.

$\therefore (F, +)$ is an abelian group.

<u>Multiplication is associative:</u>

Let $\frac{a}{b}, \frac{c}{d}, \frac{e}{f} \in F$

$$\left[\left(\frac{a}{b}\right)\left(\frac{c}{d}\right)\right]\left(\frac{e}{f}\right) = \left(\frac{ac}{bd}\right)\left(\frac{e}{f}\right) = \frac{(ac)e}{(bd)f} = \frac{a(ce)}{b(df)} = \left(\frac{a}{b}\right)\left[\left(\frac{c}{d}\right)\left(\frac{e}{f}\right)\right]$$

$\therefore$ Multiplication is assocative in F.

<u>Existance of Multiplicative Identity:</u>

Let $0 \neq x \in D$ then $\frac{x}{x} \in F$

For $\frac{a}{b} \in F$, $\left(\frac{a}{b}\right)\left(\frac{x}{x}\right) = \frac{ax}{bx} = \frac{a}{b}$ and

$\left(\frac{x}{x}\right)\left(\frac{a}{b}\right) = \frac{xa}{xb} = \frac{ax}{bx} = \frac{a}{b}$ from (1)

$\therefore \frac{x}{x}$ when $x \neq 0$ is the multiplicative identity of F.

Existance of Multiplicative Identity:

Let $\frac{a}{b} \neq \frac{o}{b} \in F$

$\Rightarrow a \neq o$

$\Rightarrow \frac{b}{a} \in F$

Now $\left(\frac{a}{b}\right)\left(\frac{b}{a}\right) = \frac{ab}{ba} = \frac{ab}{ab} = \frac{x}{x}$ from (1)

$\left(\frac{b}{a}\right)\left(\frac{a}{b}\right) = \frac{ba}{ab} = \frac{ab}{ab} = \frac{x}{x}$

$\therefore \frac{b}{a}$ is the multiplicative inverse of $\frac{a}{b}$ in F.

Multiplication is commutative:

Let $\frac{a}{b}, \frac{c}{d} \in F$

$\left(\frac{a}{b}\right)\left(\frac{c}{d}\right) = \frac{ac}{bd} = \frac{ca}{db} = \left(\frac{c}{d}\right)\left(\frac{a}{b}\right)$

$\therefore$. is commutative.

$\therefore \left(F - \left\{\frac{o}{x}\right\}, .\right)$ is an abelion group.

Distributive Law:

Let $\frac{a}{b}, \frac{c}{d}, \frac{e}{f} \in F$

$$\left(\frac{a}{b}\right)\left[\frac{c}{d}+\frac{e}{f}\right]=\frac{a}{b}\left(\frac{cf+de}{df}\right)=\frac{a(cf+de)}{bdf}=\frac{acf+ade}{bdf}$$

$$\left(\frac{a}{b}\right)\left(\frac{c}{d}\right)+\left(\frac{a}{b}\right)\left(\frac{e}{f}\right)=\frac{ac}{bd}+\frac{ae}{bf}=\frac{acbf+bdae}{bdbf}=\frac{(acf+ade)b}{bdfb}=\frac{acf+ade}{bdf}$$

$$\therefore\left(\frac{a}{b}\right)\left[\frac{c}{d}+\frac{e}{f}\right]=\left(\frac{a}{b}\right)\left(\frac{c}{d}\right)+\left(\frac{a}{b}\right)\left(\frac{e}{f}\right)$$

Similarly $\left[\frac{c}{d}+\frac{e}{f}\right]\left(\frac{a}{b}\right)=\left(\frac{c}{d}\right)\left(\frac{a}{b}\right)+\left(\frac{e}{f}\right)\left(\frac{a}{b}\right)$

$\therefore (F,+,.)$ is a field.

Define $f:D\to F$ by $f(a)=\frac{ax}{x}$ where $o\neq x\in D$

Clearly f is a mapping.

To show that f is a monomorphism.

<u>f is a homomorphism:</u>

Let $a,b\in D$

Now $f(a+b)=\frac{(a+b)x}{x}=\frac{(a+b)x^2}{x^2}$

$$=\frac{ax^2+bx^2}{x^2}$$

$$=\frac{(ax)x+(bx)x}{xx}$$

$$=\frac{ax}{x}+\frac{bx}{x}$$

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

$$f(ab)=\frac{abx}{x}=\frac{abx^2}{x^2}=\frac{axbx}{xx}=\frac{ax}{x}.\frac{bx}{x}$$

$= f(a).f(b)$

$\therefore$ f is a homomorphism.

f is one-one:

Let $a, b \in D$ and $f(a) = f(b)$

$$\Rightarrow \frac{ax}{x} = \frac{bx}{x}$$

$$\Rightarrow (ax, x) \sim (bx, x)$$

$$\Rightarrow axx = xbx$$

$$\Rightarrow ax^2 = bx^2$$

$$\Rightarrow (a-b)x^2 = o$$

$$\Rightarrow a - b = o$$

$$\Rightarrow a = b$$

$\therefore f$ is one - one and hence a monomorphism.

$\therefore$ Every integral domain can be embedded in a field.

2.11.11 Summary:

In this lesson we defined ideal, principal ideal, prime ideal, maximal ideal, quotient ring homomorphism of rings, kernal of a homomorphism. We learnt algebra of ideals. Theorems on ideals, quotient rings homomorphism of rings, kernel of a homomorphism, fundamental theorem and embedding of rings.

2.12 Technical Terms:

i) Ideal

ii) Ideal generated by a subset

iii) Principal ideal

iv) Prime ideal

v) Maximal ideal

vi) Quotient ring

vii) Homomorphism of rings

viii) Kernal of a homomorphism

iv) Embedding of rings

2.13.1 Model Questions:

2.4.3 The set $I = \left\{ \begin{pmatrix} 0 & a \\ 0 & b \end{pmatrix} \Big/ a, b \in z \right\}$ is a left ideal of the ring of 2 x 2 matrices with integers as entries but not a right ideal.

Solution: Clearly I is a non-empty subset of the ring M of 2 x 2 matrices with integers as entries.

Let $A = \begin{pmatrix} 0 & a \\ 0 & b \end{pmatrix}$ and $B = \begin{pmatrix} 0 & c \\ 0 & d \end{pmatrix} \in I$ and $C = \begin{pmatrix} x & y \\ r & s \end{pmatrix} \in M$

$\Rightarrow a, b, c, d, x, y, r, s \in z$

$A - B = \begin{pmatrix} 0 & a-c \\ 0 & b-d \end{pmatrix} \in I$ Since $a-c, b-d \in z$

$CA = \begin{pmatrix} x & y \\ r & s \end{pmatrix} \begin{pmatrix} 0 & a \\ 0 & b \end{pmatrix} = \begin{pmatrix} 0 & ax+by \\ 0 & ar+bs \end{pmatrix} \in I$ Since $ax+by, ar+bs \in z$

$\therefore$ I is a left ideal.

But $AC = \begin{pmatrix} 0 & a \\ 0 & b \end{pmatrix} \begin{pmatrix} x & y \\ r & s \end{pmatrix} = \begin{pmatrix} ar & as \\ br & bs \end{pmatrix} \notin I$

Hence I is not a right ideal.

2.13.2 If R is a commutative ring and $a \in R$ then $Ra = \{ra / r \in R\}$ is an ideal of R.

Solution: Let $Ra = \{ra / r \in R\}$

Clearly $o \in R \Rightarrow oa \in Ra$

$\therefore ox = o \in Ra$

$\therefore Ra$ is a non-empty subset of R.

Let $x, y \in Ra$ and $r \in R$

$\Rightarrow x = r_1 a$ and $y = r_2 a$ where $r_1, r_2 \in R$

Now $x - y = r_1 a - r_2 a = (r_1 - r_2)a \in Ra$

$rx = r(r_1 a) = r r_1 a \in Ra$

$xr = (r_1 a)r = r(r_1 a) = rr_1 \in Ra$ Since R is commutative.

$\therefore x, y \in Ra$ and $r \in R \Rightarrow x - y, rx, xr \in Ra$

$\therefore Ra$ is an ideal of R.

2.13.3 If R is a ring and $a \in R$ then Ra is a left ideal and $aR = \{ar/r \in R\}$ is an ideal of the ring of integers z.

2.13.4 If m is a fixed integer. Then $I = \{mx/x \in Z\}$ is an ideal of the ring of integers Z.

Solution: Let z be the ring of integers and $m \in z$.

We know that z is a commutative ring and $m \in z$.

$\therefore I = \{mx/x \in z\}$ is an ideal of z by 2.13.2.

2.13.5 Union of two ideals need not be an ideal (2.4.10)

Solution: Let $I_1 = \{2x/x \in z\}$ and $I_2 = \{3x/x \in z\}$

Then by 2.13.4 I_1 and I_2 are ideals of the ring of integers Z.

$I_1 \cup I_2 = \{... -4, -3, -2, 0, 2, 3, 4, 6, 8, 9...\}$

$3, 2 \in I_1 \cup I_2$ but $3 - 2 = 1 \notin I_1 \cup I_2$

$\therefore I_1 \cup I_2$ is not an ideal.

2.13.6 A commutative ring R with unity element $1 \neq 0$ is a field if R has no proper ideals.

Solution: Let R be a commutative ring with unity element 1.

Suppose R has no proper ideals.

To prove that R is a field.

Let $0 \neq a \in R$

$\Rightarrow Ra = \{ra/r \in R\}$ is an ideal of R by 2.13.2.

Since R has no proper ideals either $Ra = \{0\}$ or $Ra = R$

But $0 \neq a \in Ra$ since $a = 1.a$ and $1 \in R$

$\therefore Ra \neq \{0\}$

$\Rightarrow Ra = R$

But $1 \in R \Rightarrow 1 \in Ra$

$\Rightarrow 1 = ba$ for some $b \in R$

$\therefore$ b is the multiplicative inverse of a.

$\therefore$ Every non zero element of the commutative ring R with unity element has multiplicative inverse.

$\Rightarrow R$ is a field.

2.13.7 If I is an ideal of a ring R and $rI = \{x \in R / xa = 0$ for all $a \in I\}$

then rI is an ideal of R.

Solution: Let R be a ring and I an ideal of R.

Let $rI = \{x \in R / xa = 0$ for all $a \in I\}$

We know that $o \in R$ and $oa = a$ for all $a \in I$

$\therefore o \in rI$

$\Rightarrow rI$ is a non-empty subset of R.

Let $x, y \in rI$ and $r \in R$.

$x, y \in rI \Rightarrow xa = 0$ and $ya = 0$ for all $a \in I$

$\Rightarrow xa - ya = 0$

$\Rightarrow (x - y)a = 0$ for all $a \in I$

$\therefore x - y \in rI$.

$x \in rI \Rightarrow xa = 0 \quad \forall a \in I$

$\Rightarrow r(xa) = 0 \quad \forall a \in I$

$\Rightarrow (rx)a = 0 \quad \forall a \in I$

$\Rightarrow rx \in r(I)$

Further I is an ideal $\Rightarrow ra \in I$ for all $a \in I$

$x \in rI \Rightarrow xa = 0$ for all $a \in I$

$\Rightarrow x(ra)=0$ since $ra \in I$ for all $a \in I$

$\Rightarrow (xr)a=0$ for all $a \in I$

$\Rightarrow xr \in r(I)$

$\therefore r(I)$ is an ideal of R.

2.13.8 If n is a fixed positive integer then the mapping $f: z \rightarrow nz$ defined by $f(x) = nx$ is not a homomorphism .

Solution: Z is the set of integers.

$f: z \rightarrow nz$ defined by $f(x)=nx$ where n is a fixed postive integer.

For $x, y \in z$

$$f(x+y)=n(x+y)$$

$$=nx+ny$$

$$= f(x)+f(y)$$

$$f(xy)=nxy$$

$$f(x).(y) = nx.ny = n^2xy$$

$$\therefore f(xy) \neq f(x)f(y)$$

$\therefore f$ is not a homomorphism.

2.13.9 If R is a ring with unity element 1 and f is an epimorphism from R onto a ring R^1 then $f(1)$ is the unity element of R^1.

Solution: Let R be a ring with unity element 1.

Suppose $f: R \rightarrow R^1$ is an epimorphism.

$1 \in R \Rightarrow f(1) \in R^1$

To show that $f(1)$ is the unity element of R^1.

Let $b \in R^1$

$f: R \rightarrow R^1$ is onto $\Rightarrow$ there is $a \in R$ such that $f(a)=b$

$a \in R \Rightarrow a.l = a = l.a$

Now $f(1)b = f(1).f(a)$

$= f(1.a)$

$= f(a)$

$= b$

Also $bf(1) = f(a).f(1)$

$= f(a1)$

$= f(a)$

$= b$

$\therefore f(1)$ is the unity element of R^1.

2.13.10 The mapping $f: C \to C$ defined by $f(x+iy) = x - iy$ is an isomorphism on the set of complex numbers C.

Solution: We know that the set of complex numbers C is a ring.

$f: C \to C$ defined by $f(x+iy) = x - iy$

Let $a+ib, c+id \in C$

$f[(a+ib)+(c+id)] = f[(a+c)+i(b+d)]$

$= (a+c) - i(b+d)$

$= a + c - ib - id$

$= a - ib + c - ib$

$= f(a+ib) + f(c+id)$

$f[(a+ib)(c+id)] = f[(ac-bd)+i(ad+bc)]$

$= (ac-bd) - i(ad+bc)$

$= (a-ib)(c-id)$

$= f(a+ib)\,f(c+id)$

$\therefore f$ is a homomorphism.

f is one - one:

Suppose $f(a+ib)=f(c+id)$

$\Rightarrow a-ib=c-id$

$\Rightarrow a=c$ and $b=d$

$\Rightarrow a+ib=c+id$

$\therefore f$ is one - one.

f is onto:

Let $x+iy\in c$

$\Rightarrow x-iy\in c$ and $f(x-iy)=x+iy$

$\therefore f$ is onto.

Hence f is an isomorphism.

2.13.11 If F is a field and $f:F\rightarrow F$ is a homomorphism then f is an isomorphism or zero homomorphism.

Solution: Suppose F is a field and $f:F\rightarrow F$ is a homomorphism.

Then $\ker f$ is an ideal of F.

But we know that $\{0\}$ and F are the only ideals of F by 2.4.8.

$\therefore \ker f=\{0\}$ or $\ker f=F$

Suppose $\ker f=\{0\}$

We know that a homomorphism f is a monomorphism

Iff $\ker f=\{0\}$ by 2.9.3

$\therefore f:F\rightarrow F$ is a monomorphism.

Hence f is an isomorphism.

Suppose $\ker f=F$

By definition of $\ker f$ we have $f(x)=0^1, \forall x\in F$.

Where 0^1 is the zero element of F.

$\Rightarrow$ f is zero homomorphism.

2.14 Exercises:

1. Show that the set $I = \left\{ \begin{pmatrix} a & b \\ 0 & 0 \end{pmatrix} \Big/ a, b \in z \right\}$ is a right ideal but not a left ideal of the ring of 2 x 2 matrices M over integers.

2. Show that if R is a commutative ring with unity element and $a \in R$ then $Ra = \{ra/r \in R\}$ is a principal ideal generated by a.

3. Find all the principal ideals of the ring $(z_6, +_6, x_6)$.

4. If R is a commutative ring with unity and $a, b \in R$ then show that $\{ax + by/x, y \in R\}$ is the ideal generated by a, b.

5. If I is an ideal of R and $[R : I] = \{x \in R/rx \in I$ for every $r \in R\}$ then $[R : I]$ is an ideal of R and $I \subseteq [R : I]$.

6. Let $R^1 = \left\{ \begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix} \Big/ a \in R \right\}$ where R is the ring of real numbers. Prove that $f : R^1 \to R$ defined by $f \begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix} = a$ for all $\begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix} \in R^1$ is an isomorphism.

7. Show that every homomorphic image of a commutative ring is commutative.

8. Give an example to show that a homomorphic image of an integral domain may not be an integral domain.

9. Let R be a ring with unity. For each inversible element $a \in R$, the mapping $f : R \to R$ defined by $f(x) = axa^{-1}, \forall x \in R$ is an automorphism.

10. Let R be the ring of all real valued continuous functions defined on $[0,1]$ and $M = \left\{ f(x) \in R \Big/ f\left(\frac{1}{3}\right) = 0 \right\}$. Show that M is maxmimal ideal of R.

- Smt. K. Ruth

LESSON - 3

RINGS OF POLYNOMIALS

3.1 Objective of the Lesson:

To learn the definition of a Polynomial over a ring, polynomial ring, degree of a polynomial and evaluation homomorphism.

3.2 Structure

3.3 Introduction:

In this lesson we will introduce the notion of a polynomial ring over a ring, over an integral domain and over a field. We also introduce evaluation homomorphism.

3.4.1 Definition of a Polynomial over a ring:

Let R be a ring and x an indeterminate. If $a_0, a_1, a_2 ... \in R$ and $a_i = 0$ for all except a finite number of values of i then a formal sum $f(x) = a_0 + a_1x + a_2x^2 + ...$ is called a polynomial over R. if $f(x) = a_0 + a_1x + ... + ...a_nx^n + ...$ has $a_i = 0$ for all i > n, then we donate $f(x)$ by $a_0 + a_1x + ... + a_nx^n$.

3.4.2 Example: 1. $f(x) = 3 + x + 4x^2 - 5x^4$ is a polynomial over z.

2. $f(x)=\frac{1}{2}+2x^2-\frac{3}{7}x^3$ is a polynomial over Q.

3. $f(x)=1+x^2+4x^3$ is a polynomial over Z_5.

3.4.3 Note : 1. The set of all polynomials over a ring R with intdeterminate x is denoted by $R[x]$.

2. We omit altogether from the formal sum any term of the form $0x^i$.

3.4.4 Note: 1. Let R be a ring. A polynomial over R can also be defined as a sequence $(a_0, a_1, a_2...a_n,....)$ of elements of R, where all but a finite number of a_i's are zero.

$\therefore$ A polynomial $f(x)=a_0+a_1x+a_2x^2+...$ can also be denoted by $(a_0, a_1...a_n...)$.

2. If $f(x)=a_0+a_1x+a_2x^2+...+a_nx^n+...$ is a polynomial over a ring R then $a_0, a_1, a_2...$ are called coefficients of $f(x)$. $a_0, a_1x, a_\alpha x^2...$ are called constant term, x term, x^2 term .... x^n term... of $f(x)$.

3.4.4. Zero Polynomial: If 0 is the zero element of a ring R then $f(x)=0+0x+0x^2+0x^n+....=(0,0,0,...0...)$ is called the zero polynomial. It is denoted by 0 or $0(x)$.

3.4.5 Constant Polynomial: Let R be a ring. Then any element a of the ring is a constant polynomial $f(x)=a+0x+0x^2+....$ It is usually written as $f(x)=a$.

3.5.1 Algebra of Polynomials:

Equity of Polynomials: Two polynomials $(a_0, a_1...a_2...)$ and $g=(b_0, b_1.....b_n....)$ over a ring R are said to be equal if $a_i=b_i$ for all $i \geq 0$.

If f and g are equal polynomials we write $f=g$.

3.5.2 Addition of Polynomials:

Let $f=(a_0, a_1, a_2...a_m...)$ and $g=(b_0, b_1, b_2...b_n...)$ be two polynomials over a ring R. the sum of f and g is denoted by $f+g=(c_0, c_1, c_\alpha...c_k...)$ where $c_i=a_i+b_i$ for $i=0,1,2,.....$

3.5.3 Example: If $f=(2,3,-4,0,0,...)$ and $g=(3,-2,0,-3,0,0...)$ are two polynomials over the ring of integers z then $f+g=(5,-1,-4,-3,0,0...)=5+x-4x^2-3x^3$

3.5.4 Multiplication of Polynomials:

Let $f=(a_0,a_1,a_2...a_m...)$ and $g=(b_0,b_1,b_2...b_n...)$ be two polynomials. Over a ring R. The product of f and g is denoted by fg and $fg=(c_0,c_1,c_2...c_p...)$ where $c_i=a_0b_i+a_1b_{i-1}+...+a_ib_0$ i.e. $c_i=\sum_{j=0}^{i}a_jb_{i-j}=\sum_{j+k=i}a_jb_k$.

3.5.5. Example : If $f=(1,3,4,0,0....)$ and $g=(2,2,0,5,0,0,....)$ over z_6 find fg.

$$f=1+3x+4x^2 \qquad g=2+2x+0x^2+5x^3$$

$$fg=1\times_6 2+(1\times_6 2+_6 3\times_6 2)x+(1\times_6 0+_6 3\times_6 2+_6 4\times_6 2)x^2$$

$$+(1\times_6 5+_6 3\times_6 0+_6 4\times_6 2)x^3$$

$$=2+2x+2x^2+x^3$$

3.6.1 Degree of a Polynomial:

Let $f=(a_0,a_1,...a_2...)$ be a non-zero polynomial over a ring R. The largest integer i for which $a_i\neq 0$ is called the degree of the polynomial f. It is denoted by $\deg f(x)$ or $\deg f$.

3.6.2 Note: 1. The degree of a non zero polynomial $f=(a_0,a_1,...a_2...)$ is n if $a_n\neq 0$ and $a_i=0 \ \forall i>n$.

2. The degree of a non zero constant polynomial is zero

3. The degree of the zero polynomial is not defined.

3.6.3 Definition: Leading Coefficient: If the degree of the polynomial $f(x)=a_0+a_1x+...+a_nx^n$ is n then $a_n\neq 0$ is called leading coefficient in $f(x)$.

3.6.4 Example: 1. The degree of the polynomial $f(x)=3-2x+x^4+x^6$ over the ring of integers is 6.

2. The degree of the polynomial $f(x)=\frac{7}{4}$ over the ring of rational numbers is zero.

3.6.5 Theorem: Let $f(x),g(x)$ be two non zero polynomials over a ring R the

i) $\deg\left(f(x)+g(x)\right)\leq\max\{\deg f(x),\deg g(x)\}$ if $f(x)+g(x)\neq 0(x)$.

ii) $\deg\left(f(x)g(x)\right) \leq \deg f(x) + \deg g(x)$ if $f(x)g(x) \neq 0(x)$ where $0(x)$ is the zero polynomial.

Proof: Let $f(x) = a_0 + a_1x + a_2x^2 + ... + a_mx^m$ and $g(x) = b_0 + b_1x + b_2x^2 + ... + b_nx^n$ be two polynomials over a ring R with $\deg f(x) = m$ and $\deg g(x) = n$.

$\therefore a_m \neq 0$ and $a_i = 0 \quad \forall i > m$ and

$b_n \neq 0$ and $b_i = 0 \quad \forall i > n$

Case i): Suppose $m > n$ then $\max\{m, n\} = m$

$$f(x) + g(x) = (a_0 + b_0) + (a_1 + b_1)x + (a_2 + b_2)x^2 + ... + (a_n + b_n)x^n + a_{n+1}x^{n+1} + ... + a_mx^m$$

$$\therefore \deg\left[f(x) + g(x)\right] = m = \max\{\deg f(x), \deg g(x)\}$$

Case ii) Suppose $m < n$ then $\max\{m, n\} = n$

$$f(x) + g(x) = (a_0 + b_0) + (a_1 + b_1)x + ... + (a_m + b_m)x^m + b_{n+1}x^{m+1} + ... + b_nx^n$$

$$\therefore \deg\left[f(x) + g(x)\right] = n = \max\{\deg f(x), \deg g(x)\}$$

Case iii) Suppose $m = n$

$$f(x) + g(x) = (a_0 + b_0) + (a_1 + b_1)x + ... + (a_m + b_m)x^m$$

$$\therefore \deg\left[f(x) + g(x)\right] \leq m = \max\{\deg f(x), \deg g(x)\}$$

ii) $f(x).g(x) = a_0b_0 + (a_1b_0 + a_0b_1)x....$

$$= d_0 + d_1x + d_2x^2 + ...$$

Where $d_k = \sum_{i+j=k} a_ib_j$ from the definition.

Suppose $k > m + n \Rightarrow i + j > m + n$

$\Rightarrow i > m$ or $j > n$.

But $i > m \Rightarrow a_i = 0$ and $j > n \Rightarrow b_j = 0$

$\Rightarrow a_i b_j = 0$ if $i > m$ or $j > n$

$\Rightarrow a_i b_j = 0$ if $i + j > m + n$

$\Rightarrow d_k = 0$ if $k > m + n$

$\therefore \deg\left(f(x)g(x)\right) \le m + n = \deg f(x) + \deg g(x)$

3.6.6 Corollary: If $f(x)$ and $g(x)$ are two nonzero polynomials over an integral domain R then $\deg\left[f(x).g(x)\right] = \deg f(x) + \deg g(x)$.

Proof: Let $f(x) = a_0 + a_1 x + \ldots + a_m x^m$ and $g(x) = b_0 + b_1 x + \ldots + b_n x^n$ be two polynomials over an integral domain R with $\deg f(x) = m$ and $\deg g(x) = n$.

$\therefore a_m \ne 0$ and $a_i = 0 \quad \forall i > m$

$b_n \ne 0$ and $b_i = 0 \quad \forall i > n$

$a_m \ne 0$, $b_n \ne 0 \Rightarrow a_m b_n \ne 0$ since R is an integral domain.

Now $f(x).g(x) = d_0 + d_1 x + d_2 x^2 + \ldots + d_{m+n} x^{m+n} + \ldots$

Where $d_{m+n} = a_0 b_{m+n} + a_1 b_{m+n-1} + \ldots + a_m b_n + a_{m+1} b_{n-1} + \ldots + a_{m+n} b_0$

$= a_m b_n$ since $a_i = 0$ for $i > m$ and

$b_i = 0$ for $i > n$

$\ne 0$

$\therefore \deg\left[f(x).g(x)\right] \ge m + n$

By 3.6.5 $\deg\left[f(x).g(x)\right] \le m + n$

Hence $\deg\left[f(x).g(x)\right] = \deg f(x) + \deg g(x)$.

3.6.7 Corollary: If $f(x)$ and $g(x)$ are non zero polynomials over an integral domain or field then $\deg f(x) \le \deg\left[f(x).g(x)\right]$.

Proof: Let $f(x)$ and $g(x)$ be two polynomials over an integral domain or field.

We always have $\deg f(x) \leq \deg f(x) + \deg g(x)$ since $\deg g(x) \geq 0$.

$\Rightarrow \deg f(x) \leq [\deg f(x).\deg g(x)]$ by 3.6.6.

3.7.1 Theorem: The set of all polynomials over a ring R is a ring with respect to addition and multiplication of polynomials.

Proof: Let $R[x]$ be the set of all polynomials over a ring R with indeterminate x.

Let $f(x) = a_0 + a_1x + ...,\quad g(x) = b_0 + b_1x + b_2x^2 + ...,\quad h(x) = c_0 + c_1x + c_2x^2 + ... \in R[x]$

$$f(x) + g(x) = (a_0 + b_0) + (a_1 + b_1)x + ... + (a_m + b_m)x^m .....$$

$$= c_0 + c_1x + c_2x^2 + ...$$

$f(x).g(x) = d_0 + d_1x + d_2x^2 + ...$ where $d_k = \sum_{i+j=k} a_i b_j$

$a_i, b_i \in R \Rightarrow a_i + b_i$ and $a_i b_i \in R \quad \forall i, j$

$\Rightarrow c_i$ and $d_k \in R \quad \forall i, k$

$\therefore f(x) + g(x)$ and $f(x)g(x) \in R[x]$.

$\therefore +$ and . are binary operations on $R[x]$.

Commutative w.r. to +

$$f(x) + g(x) = \sum_{i=1}^{\infty} a_i x^i + \sum_{i=1}^{\infty} b_i x^i$$

$$= \sum_{i=0}^{\infty} (a_i + b_i) x^i$$

$$= \sum_{i=0}^{\infty} (b_i + a_i) x^i \quad \text{since } a_i + b_i = b_i + a_i, \forall a_i, b_i \in R$$

$$= \sum_{i=0}^{\infty} (b_i x^i + a_i x^i)$$

$$= \sum_{i=0}^{\infty} b_i x^i + \sum_{i=0}^{\infty} a_i x^i$$

$$= g(x) + f(x)$$

$\therefore$ Addition is commutative.

Associative Law w.r.to +.

$$\left[f(x) + g(x) \right] + h(x) = \left[\sum_{i=0}^{\infty} a_i x^i + \sum_{i=0}^{\infty} b_i x^i \right] + \sum_{i=0}^{\infty} c_i x^i$$

$$= \sum_{i=0}^{\infty} (a_i + b_i) x^i + \sum_{i=0}^{\infty} c_i x^i$$

$$= \sum_{i=0}^{\infty} \left[(a_i + b_i) + c_i \right] x^i$$

$$= \sum_{i=0}^{\infty} \left[a_i + (b_i + c_i) \right] x^i$$ since $(a_i + b_i) + c_i = a_i + (b_i + c_i)$ in R.

$$= \sum_{i=0}^{\infty} \left[a_i x^i + (b_i + c_i) x^i \right]$$

$$= \sum_{i=0}^{\infty} a_i x^i + \sum_{i=0}^{\infty} (b_i + c_i) x^i$$

$$= \sum_{i=0}^{\infty} a_i x^i + \left[\sum_{i=0}^{\infty} b_i x^i + \sum_{i=0}^{\infty} c_i x^i \right]$$

$$= f(x) + \left[g(x) + h(x) \right]$$

$\therefore$ addition is associative.

<u>Existance of zero element:</u>

The zero polynomial $0(x) = 0 + 0x + 0x^2 + \ldots = \sum_{i=0}^{\infty} 0x^i \in R[x]$

$$f(x) + 0(x) = \sum_{i=0}^{\infty} a_i x^i + \sum_{i=0}^{\infty} 0x^i$$

$$= \sum_{i=0}^{\infty} (a_i + 0) x^i$$

$$= \sum_{i=0}^{\infty} a_i x^i \quad \text{Since } \; a_i + 0 = a_i \; \forall \, a_i \in R .$$

$\therefore 0(x)$ is the zero element of $R[x]$.

Existance of Additive inverse:

$$f(x) = \sum_{i=0}^{\infty} a_i x^i \in R[x] \Rightarrow a_i \in R \; \forall \, i$$

$$\Rightarrow -a_i \in R \text{ such that } a_i + (-a_i) = 0$$

$$\Rightarrow \sum_{i=0}^{\infty} (-a_i) x^i \in R[x]$$

$$\Rightarrow (-f)(x) = \sum_{i=0}^{\infty} (-a_i) x^i \in R[x]$$

Now $\Rightarrow f(x) + (-f)(x) = \sum_{i=0}^{\infty} a_i x^i + \sum_{i=0}^{\infty} (-a_i) x^i$

$$= \sum_{i=0}^{\infty} [a_i + (-a_i)] x^i$$

$$= \sum_{i=0}^{\infty} 0 x^i$$

$$= 0(x)$$

$\therefore$ Every element of $R[x]$ has additive inverse.

Associative Law w.r.to multiplication:

$$[f(x).g(x)] h(x) = \left(\sum_{i=0}^{\infty} a_i x^i \right) \left(\sum_{i=0}^{\infty} b_i x^i \right) \sum_{i=0}^{\infty} c_i x^i$$

$$= \left[\sum_{k=0}^{\infty}\left(\sum_{i+j=k} a_i b_i\right)x^k\right]\sum_{i=0}^{\infty} c_i x^i$$

$$= \sum_{n=0}^{\infty}\left[\sum_{i+j+s=n}(a_i b_i)c_s\right]x^n$$

$$= \sum_{n=0}^{\infty}\left[\sum_{k+s=n}\left(\sum_{i+j=k} a_i b_i\right)c_s\right]x^n$$

$$f(x)[g(x)h(x)] = \left[\sum_{i=0}^{\infty} a_i x^i\left(\sum_{i=0}^{\infty} b_i x^i\right)\left(\sum_{i=0}^{\infty} c_i x^i\right)\right]$$

$$= \sum_{i=0}^{\infty} a_i x^i\left[\sum_{s=0}^{\infty}\left(\sum_{j+k=s} b_j c_k\right)x^i\right]$$

$$= \sum_{n=0}^{\infty}\left[\sum_{i+s=n}\left(a_i \sum_{j+k=s} b_j c_k\right)\right]x^i$$

$$= \sum_{n=0}^{\infty}\left[\sum_{i+j+k=n} a_i(b_j c_k)\right]x^i$$

$$\therefore [f(x)g(x)]h(x) = f(x)[g(x)h(x)]$$

Multiplication is associative.

Distributive Law:

$$f(x)[g(x)+h(x)] = \sum_{n=0}^{\infty} a_i x^i\left[\sum_{j=0}^{\infty} b_j x^j + \sum_{j=0}^{\infty} c_j x^j\right]$$

$$= \sum_{n=0}^{\infty} a_i x^i\left[\sum_{j=0}^{\infty}(b_j + c_j)x^j\right]$$

$$= \sum_{n=0}^{\infty}\left[\sum_{i+j=n} a_i(b_j + c_j)\right]x^n$$

$$= \sum_{n=0}^{\infty} \left[\sum_{i+j=n} (a_i b_j + a_i c_j) \right] x^n$$

$$= \sum_{n=0}^{\infty} \left(\sum_{i+j=n} a_i b_j \right) x^n + \sum_{n=0}^{\infty} \left(\sum_{i+j=n} a_i c_j \right) x^n$$

$$= f(x).g(x) + f(x).h(x)$$

Similarly we can prove the other distributive law.

$$[g(x)+h(x)]F(x) = g(x).f(x)+h(x).f(x)$$

Hence $R[x]$ is a ring.

Definition: Let R be a ring. Then $R[x]$ is called the ring of polynomials in the inderminate x with coefficiants in R.

Theorem: The ring $R[x]$ of polynomials over an integral domain R is an integral domain.

Proof: Let R be an integral domain and $R[x]$ be the ring of polynomials over R with indeterminate x.

To prove $R[x]$ is an integral domain we have to prove $R[x]$ is commutative and without zero divisors.

Let $f(x) = a_0 + a_1 x + a_2 x^2 + ... = \sum_{i=0}^{\infty} a_i x^i$ and

$g(x) = b_0 + b_1 x + b_2 x^2 + ... = \sum_{i=0}^{\infty} b_i x^i$ be two polynomials in $R[x]$.

$$f(x)g(x) = \left(\sum_{i=0}^{\infty} a_i x^i \right) \left(\sum_{j=0}^{\infty} b_j x^j \right)$$

$$= \sum_{n=0}^{\infty} \left(\sum_{i+j=n}^{\infty} a_i b_j \right) x^n$$

$$= \sum_{n=0}^{\infty} \left(\sum_{i+j=n}^{\infty} b_j a_i \right) x^n \qquad \text{since R is commutative.}$$

$$=\left(\sum_{j=1}^{\infty} b_j x^j\right)\left(\sum_{i=0}^{\infty} a_i x^i\right)$$

$$= g(x)\,f(x).$$

$\therefore R[x]$ is commutative.

Suppose $f(x)\neq 0, g(x)\neq 0$ and deg $f(x)=m$, deg $g(x)=n$

$\Rightarrow a_m \neq 0$ and $\Rightarrow b_n \neq 0$

$\Rightarrow a_m b_n \neq 0$ Since R is an integral domain.

$\therefore$ Atleast one coefficient in $f(x)g(x)$ is non-zero.

$$\Rightarrow f(x)g(x)\neq 0$$

$\therefore R[x]$ has no zero divisors.

Hence $R[x]$ is an integral domain.

3.7.3 Note: If R is a ring with unity then the ring of polynomials $R[x]$ over R is also ring with unity.

Suppose R is a ring with unity 1.

Then f $(x)=1+0x+0x^2+\ldots\in R[x]$

i.e. $I(x)=\sum_{j=0}^{\infty} b_j x_j$ where $b_0=1$ and $b_j=0 \;\forall\; j\geq 1$

Let $f(x)=a_0+a_1x+a_2x^2+\ldots\in R[x]$

$$f(x)I(x)=\left(\sum_{i=0}^{n} a_i x^i\right)\left(\sum_{j=0}^{\infty} b_j x^j\right)$$

$$=\sum_{n=0}^{\infty}\left(\sum_{i+j=n} a_i b_j\right)x^n$$

$$=\sum_{n=0}^{\infty}(a_n.1)x^n$$

$$= \sum_{n=0}^{\infty} a_n x^n = f(x)$$

Similarly $I(x) f(x) = f(x)$.

$\therefore I(x)$ is the identity element of $R[x]$.

3.7.4 Corollary: If R is a field then $R[x]$ is an integral domain but not a field.

Proof: Suppose R is a field

$\Rightarrow R$ is an integral domain

$\Rightarrow R[x]$ is an integral domain by 3.7.2.

Let $0 \neq f(x) \in R[x]$

Suppose deg $f(x) > 0$

If g(x) is the multiplicative inverse of $f(x)$ in $R[x]$

Then $f(x)g(x) = I(x)$

$\Rightarrow \deg[f(x)g(x)] = 0$

$\Rightarrow$ deg $f(x)$ + deg $g(x)$ =0 by 3.6.6.

$\Rightarrow$ deg $f(x) = 0$

A contradiction.

$\therefore f(x)$ has no multiplicative inverse.

Hence $R[x]$ need not be a field.

3.8.1 The Evaluation:

Theorem: Let F be a subfield of a field E and $F[x]$ be the ring of polynomials over the field F. If $\alpha \in E$ then the mappying $\phi_\alpha : F[x] \rightarrow E$ defined by

$$\phi_\alpha(a_0 + a_1 x + a_2 x^2 + ... + a_n x^n) = a_0 + a_1\alpha + a_2\alpha^2 + ... + a_n\alpha^n$$

for all $a_0 + a_1 x + a_\alpha x^2 + ... + a_n x^n \in F[x]$ is a homomorphism.

Proof: Let F be a subfield of a field E and $F[x]$ be the ring of polynomials over the field F. For $\alpha \in E$

$\phi_\alpha : F[x] \to E$ is defined by

$\phi_\alpha(a_0 + a_1x + a_2x^2 + ... + a_nx^n) = a_0 + a_1\alpha + a_2\alpha^2 + ... + a_n\alpha^n$

for all $a_0 + a_1x + a_2x^2 + ... + a_nx^n \in F[x]$ is a homomorphism.

Cleary ϕ_α is well defined.

Let $f(x) = a_0 + a_1x + ... + a_mx^m$ and $g(x) = b_0 + b_1x + ... + b_nx^n$

$\Rightarrow f(x) + g(x) = c_0 + c_1x + ... + c_kx^k$ where $c_0 = a_0 + b_0$

$c_1 = a_1 + b_1 ... c_k = a_k + b_k$ and $k = \max\{m, n\}$

$\phi_\alpha[f(x) + g(x)] = \phi_\alpha[c_0 + c_1x + ... + c_kx^k]$

$= c_0 + c_1\alpha + c_2\alpha^2 + ... + c_k\alpha^k$

$= (a_0 + b_0) + (a_1 + b_1)\alpha + (a_2 + b_2)\alpha^2 + ... + (a_k + b_k)\alpha^k$

$= (a_0 + b_0) + (a_1\alpha + b_1\alpha) + (a_2\alpha^2 + b_2\alpha^2) + ... + (a_k\alpha^k + b_k\alpha^k)$

$= (a_0 + a_1\alpha + a_2\alpha^2 + ... + a_k\alpha^k) + (b_0 + b_1\alpha + b_2\alpha^2 + ... + b_k\alpha^k)$

Since a_i's, b_i's and α are elements of a field $= \phi_\alpha[f(x)] + \phi_\alpha[g(x)]$

$f(x)g(x) = d_0 + d_1x + d_2x^2 + ... + d_px^p$ where $d_n = \sum_{i+j=n} a_ib_j$

$\phi_\alpha[f(x)g(x)] = \phi_\alpha[d_0 + d_1x + d_2x^2 + ... + d_px^p]$

$= d_0 + d_1\alpha + d_2\alpha^2 + ... + d_p\alpha^p$

$= a_0b_0 + (a_0b_1 + a_1b_0)\alpha + (a_0b_2 + a_1b_1 + a_2b_0)\alpha^2 + ... + \sum_{i+i=p} a_ib_j\alpha^p$

$= (a_0 + a_1\alpha + a_2\alpha^2 + ... + a_m\alpha^m) \quad (b_0 + b_1\alpha + ... + b_n\alpha^n)$

$= \phi_{\alpha}\left[f(x)\right]\phi_{\alpha}\left[g(x)\right]$ .... (2)

From (1) and (2) ϕ_{α} is a homomorphism.

3.8.2 Note:

(1) The homomorphism ϕ_{α} defined in 3.8.1 is called the evaluation homorphism at α.

(2) If $\phi_{\alpha} : F\left[x\right] \to E$ is an evaluation homorphism at $\alpha \in E$ and

$f(x) = a_0 + a_1x + a_2x^2 + ... + a_mx^m \in F(x)$

then $\phi_{\alpha}\left[f(x)\right] = a_0 + a_1\alpha + a_2\alpha^2 + ... + a_m\alpha^m$ is denoted by $f(\alpha)$.

3.8.3 Definition: Zero of a Polynomial: Let F be a subfield of a field E and $\alpha \in E$. Let $\phi_{\alpha} : F\left[i\right] \to E$ be an evaluation homorphism. For $f(x) = a_0 + a_1x + ... + a_n\alpha^n\alpha \in F\left[x\right]$ if $f(\alpha) = \phi_{\alpha}\left[f(x)\right] = a_0 + a_1x + ... + a_n\alpha^n = 0$ then $\alpha \in E$ is called a zero of the polynomial $f(x)$ or $\alpha \in E$ is a solution of the equation $f(x) = 0$.

3.8.4 Note: Let $f(x)$ be a polynomial over a subfield F of a field E. Then $\alpha \in E$ is called a zero of the polynomial $f(x)$ if $f(\alpha)$=0.

3.8.5 Example: The zeros of $f(x) = x^4 + 4$ in $z_5\left[x\right]$ are $1, 2, 3, 4$

$z_5 = \{0, 1, 2, 3, 4\}$

$f(x) = x^4 + 4$

$f(x) = 0$ if $x = 1, 2, 3, 4$

$\therefore$ zeros of $f(x)$ are $1, 2, 3, 4$

3.8.6 Kernal of Evaluation Homomorphism : Let F be a subfield of a field E. For $\alpha \in E_n$ we have an evaluation homomorphism $\phi_{\alpha} : F\left[x\right] \to E$ defined as $\phi_{\alpha}\left[f(x)\right] = f(\alpha)$. Then the set $\left\{f(x) \in F[x] / \phi_{\alpha}\left[f(x)\right] = f(\alpha) = 0\right\}$ where 0 is the zero element of E is called kernal of ϕ_{α}.

It is denoted by $\ker\phi_{\alpha}$.

3.9 Summary:

In this lesson we defined polynomial over a ring. We proved set of all polynomials over a ring is a ring. Set of all polynomials over a field is an integral domain. We defined Evaluation homorphism, zero of a polynomial and Kernal of evaluation homomorphism.

3.10 Technical Terms:

i) Polynomial over a ring

ii) Degree of a polynomial

iii) Leading coefficient

iv) Evaluation homomorphism

v) Zero of a polynomial

vi) Kernal of Evaluation Homomorphism.

3.11 Model Examination Questions:

1. Find the sum and product of the polynomials $f(x)=2+3x+5x^2$ and $g(x)=1+2x+3x^2$ over Z_6.

Solution: $Z_6=\{0,1,2,3,4,5\}$

$f(x)=2+3x+5x^2$ and $g(x)=1+2x+3x^2$

$$f(x)+g(x)=(2+_6 1)+(3+_6 2)x+(5+_6 3)x^2$$

$$=3+5x+2x^2$$

$$f(x)g(x)=(2+3x+5x^2)(1+2x+3x^2)$$

$$=(2\times_6 1)+(2\times_6 2+_6 3\times_6 1)x+(2\times_6 +_6 3\times_6 2+_6 5\times_6 1)x^2+(3\times_6 3+_6 5\times_6 2)x^3+(5\times_6 3)x^4$$

$$=2+x+5x^2+x^3+3x^4$$

2. If $f(x)=5+3x+2x^2$ and $g(x)=1+3x+4x^3\in Z_6$ find deg $[f(x)+g(x)]$ and deg $[f(x).g(x)]$.

Solution: Given $f(x)=5+3x+2x^2$ and $g(x)=1+3x+4x^3$

deg $f(x)=2$ deg $g(x)=3$

We have deg $[f(x)+g(x)]=\max\{f(x),g(x)\}$

$$= \max\{2,3\}$$

$$= 3$$

$f(x).g(x) = (5+3x+2x^2)(1+3x+4x^3)$

$= 5 + [(5\times_6 3) +_6 (3\times_6 1)]x + [(3\times_6 3) +_6 (2\times_6 1)]x^2 + [5\times_6 4 +_6 (2\times_6 3)]x^3 + (3\times_6 4)x^4 + (2\times_6 4)x^5$

$= 5 + 5x^2 + 2x^3 + 2x^5$

deg $f(x).g(x) = 5$

3. If $\phi_5 : Z_7[x] \to Z_7$ is an evaluation homomorphism find $\phi_5(2+x^3), \phi_5(3+4x^2)$ and $\phi_5\left[(2+x^3)(3+4x^2)\right]$.

Solution: $\phi_5 : Z_7[x] \to Z_7$ is an evaluation homomorphism.

$$\phi_5(2+x^3) = 2+5^3 = 2+6 = 1$$

$$\phi_5(3+4x^2) = 3+4.5^2 = 5$$

$$\phi_5\left[(2+x^3)(3+4x^2)\right] = \phi_5(2+x^3)\phi_5(3+4x^2)$$

$$= 1.5 = 5$$

4. Find the zeros of $f(x) = 1+x+x^2$ in $Z_7[x]$.

Solution: We have $Z_7 = \{0,1,2,3,4,5,6\}$

$$f(0) = 1+0+0^2 = 1 \neq 0$$

$$f(1) = 1+1+1^2 = 3 \neq 0$$

$$f(2) = 1+2+2^2 = 0$$

$$f(3) = 1+3+3^2 = 6 \neq 0$$

$$f(4) = 1+4+4^2 = 0$$

$$f(5) = 1+5+5^2 = 3 \neq 0$$

$f(6) = 1 + 6 + 6^2 = 1 \neq 0$.

$\therefore$ Zeros of $f(x) = 1 + x + x^2$ in $Z_7[x]$ are 2 and 4.

3.12 Exercises:

1. If $f(x) = 5 - x^2$ and $g(x) = 3 + 2x + x^3$ are polynomials in $Z_5[x]$ find the sum and product of $f(x)$ and $g(x)$.

2. If $f(x) = 3 + 2x + x^3$ and $g(x) = 4 + x^4 \in Z_6[x]$ then prove that $\deg[f(x).g(x)] \neq \deg f(x) + \deg g(x)$

3. If $f(x) = 2 + 5x + 3x^2$, $g(x) = 1 + 4x + 2x^3$ find the sum product, $\deg[f(x) + g(x)]$ and $\deg[f(x).g(x)]$ in $Z_7[x]$.

4. Find the zeros of $f(x) = 1 + x^2$ in $Z_5[x]$.

5. Let F be a field. Prove that a polynomial $f(x) \in F[x]$ is a unit if and only if it is a nonzero constant polynomial.

6. Let F be a field and let $F[x]$ be the ring of polynomials over F. Let $f(x)$ and $g(x)$ be nonzero polynomials in $F[x]$. We say that $f(x)$ divides $g(x)$ if there exists a polynomial $h(x) \in F[x]$. Such that $g(x) = f(x).h(x)$.

Prove that $f(x)$ divides $g(x)$ and $g(x)$ divides $f(x)$ if and only if $g(x) = h(x) f(x)$, where $h(x)$ is a nonzero constant polynomial in $F[x]$.

- Smt. K. Ruth

LESSON - 4

FACTORIZATION OF POLYNOMIALS OVER A FIELD

4.1 Objectives of the Lesson:

To acquain the student with the division algorithm of polynomials, irreducible polynomials, the famous Eisenstein criteria for irreducibility of polynomials, ideals in $F[x]$ and the unique factorization of a nonconstant polynomial as a product of a finite number of irredducible polynomials.

4.2 Structure:

This lesson contains the following components:

4.3 Introduction:

Throughout the lesson, we assume that F is a field and $F[x]$ is the ring of polynomials over F. In this lesson, similar to the division algorithm of integers, we prove the division algorithm for $F[x]$. Some important corollaries are proved. The concept of irreducibility of polynomials is introduced and some criteria for determining irreducibility of quadratic and cubic polynomials is obtained. The famous eisenstein's irreducibility criterion is discussed. Suitable examples are given. We also prove that every ideal in $F[x]$ is a principal ideal. Finally, we prove that every nonconstant polynomial in $F[x]$ can be written uniquely as a product of a finite number of irreducible polynomials in $F[x]$.

Throughout the lesson we use the following notation.

Z : The ring of integers

Q : The field of rational numbers

R : The field of real numbers

C : The field of complex numbers

Z_n : The ring of integers modulo n.

$Z_n[x]$: The ring of polynomials over Z_n.

4.4 The Division Algorithm in $F[x]$:

In order for $F[x]$ to be an Enclidean ring we need to prove the division algorithm in $F[x]$. This is provided by the following.

4.4.1 Theorem: (The division algorithm) let $f(x)=a_0+a_1x+...+a_nx^n, g(x)=b_0+b_1x+...+b_mx^m \in F[x]$ be polynomials such that $a_n \neq 0$ and $b_m \neq 0$. Then there exist unique polynomials $q(x), r(x) \in F[x]$ such that $f(x)=q(x)g(x)+r(x)$ where $r(x)=0$ or $\deg r(x) < \deg g(x)$. ($q(x)$ is called the quotient and $r(x)$ the remainder).

Proof: If $\deg g(x) > \deg f(x)$, let $q(x)=0$ and $r(x)=f(x)$. Assume that $\deg g(x) \leq \deg f(x)$, i.e. $m \leq n$. Proof is by induction on n. If $n=0$ then $m=0, f(x)=a_0$ and $g(x)=b_0$. Let $q(x)=a_0b_0^{-1}$ and $r(x)=0$. Then $q(x)g(x)+r(x)=\left(a_0b_0^{-1}\right)b_0=a_0=f(x)$. Assume that the existence part of the theorem is true for polynomials of degree less than n where $n>0$. Now $\left(a_nb_m^{-1}x^{n-m}\right)g(x)=\left(a_nb_m^{-1}x^{n-m}\right)\left(b_0+b_1x+...+b_mx^m\right)=a_nb_m^{-1}b_0x^{n-m}+a_nb_m^{-1}x^{n-m+1}+...+a_nx^n$. Hence

$f(x)-\left(a_nb_m^{-1}x^{n-m}\right)g(x)=\left(a_0+a_1x+...+a_nx^n\right)-\left(a_nb_m^{-1}b_0x^{n-m}...+a_nx^n\right)$ is a polynomial of degree less than n. By induction hypothesis there are polynomials p(x) and $q(x)$ such that $f(x)-\left(a_nb_m^{-1}x^{n-m}\right)g(x)=P(x)g(x)+r(x)$ where $r(x)=0$ or $\deg r(x) < \deg g(x)$. Therefore if $q(x)=a_nb_m^{-1}x^{n-m}+P(x)$, then $f(x)=q(x)g(x)+r(x)$

For uniqueness, suppose $f(x)=q_1(x)g(x)+r_1(x)$ and $f(x)=q_2(x)g(x)+r_2(x)$, where $r_1(x)=0, r_2(x)=0$ of $\deg r_1(x) < \deg g(x)$ and $\deg r_2(x) < \deg g(x)$. Assume that $r_1(x) \neq) r_2(x)$.

Since $(q_1(x)-q_2(x))\,g(x)=r_2(x)-r_1(x)$, we get that $q_1(x)\neq q_2(x)$. Thus $\deg(q_1(x)-q_2(x))+\deg g(x)=\deg(q_1(x)-q_2(x))g(x)=\deg(r_2(x)-r_1(x)$. But $\deg(q_1(x)-q_2(x))+\deg g(x)\geq \deg g(x)$ and $\deg(r_2(x)-r_1(x)\leq \max\{\deg r_2(x),\deg r_1(x)\}<\deg g(x)$. This is a contradiction. Therefore $r_1(x)=r_2(x)$ and so $q_1(x)=q_2(x)$.

We compute the polynomials $q(x)$ and $r(x)$ of theorem 4.4.1 by long division.

4.4.2. Example: For polynomials $f(x)=x^6+3x^5+4x^2-3x+2$ and $g(x)=x^2+2x-3$ in $Q[x]$, find $q(x)$ and $r(x)$ as described by the division algorithm so that $f(x)=g(x)q(x)+r(x)$ with $r(x)=0$ or $\deg r(x)<\deg g(x)$

$$
\begin{array}{l|l}
 & x^4+x^3+x^2+x+5 \\
\hline
x^2+2x-3 & x^6+3x^5+\qquad 4x^2-3x+2 \\
 & \underline{x^6+2x^5-3x^4} \\
 & \quad x^5+3x^4 \\
 & \quad \underline{x^5+2x^4-3x^3} \\
 & \qquad x^4+3x^3+4x^2 \\
 & \qquad \underline{x^4+2x^3-3x^2} \\
 & \qquad\quad x^3+7x^2-3x \\
 & \qquad\quad \underline{x^3+2x^2-3x} \\
 & \qquad\qquad 5x^2\quad +2 \\
 & \qquad\qquad \underline{5x^2+10x-15} \\
 & \qquad\qquad\quad -10x+17
\end{array}
$$

Thus $q(x)=x^4+x^3+x^2+x+5$ and

$r(x)=-10x+7$

4.4.3 Corollary: (Remainder theorem) Let $f(x)=a_0+a_1x+...+a_nx^n\in F[x]$. For any $a\in F$, there exists a unique polynomial $q(x) \in F[x]$ such that $f(x) = q(x)(x-a) + f(a)$.

Proof: If $f(x) = 0$, let $q(x) = 0$. Then $f(x) = q(x)(x-a) + f(a)$. suppose that $f(x) \neq 0$. By theorem 4.4.1 there exist unique polynomials $q(x), r(x) \in F[x]$ such that $f(x) = q(x)(x-a) + r(x)$ where $r(x) = 0$ or $\deg r(x) < 1$. Thus we have that $r(x) = c$ for some $c \in F$. So, $f(x) = q(x)(x-a) + c$. Then $f(a) = q(a)(a-a) + c = c$. Thus $f(x) = q(x)(x-a) + f(a)$.

4.4.4 Corollary (Factor Theorem): An element $a \in F$ is a zero of $f(x) \in F[x]$ if and only if $x-a$ is a factor of $f(x)$.

Proof: If a is a zero of $f(x)$, then $f(a) = 0$. By corollary 4.4.3, $f(x) = q(x)(x-a) + f(a)$, for some $q(x) \in F[x]$. since $f(a) = 0$, we get that $(x-a)$ is a factor of $f(x)$. Conversely if $(x-a)$ is a factor of $f(x)$, then $f(x) = q(x)(x-a)$ for some $q(x) \in F[x]$. Then $f(a) = q(a)(a-a) = 0$. Thus a is a zero of $f(x)$.

4.4.5 Example: Let $x^3 + 4x^2 + 4x + 1 \in Z_5[x]$. We divide this polynomial by $x-1$ and get

$$
\begin{array}{r|l}
 & x^2 + 4 \\ \hline
x-1 & x^3 + 4x^2 + 4x + 1 \\
 & \underline{x^3 - x^2} \\
 & \quad 0 + 4x + 1 \\
 & \quad\quad \underline{4x - 4} \\
 & \quad\quad\quad\quad 0
\end{array}
$$

Therefore $x^3 + 4x^2 + 4x + 1 = (x-1)(x^2 + 4)$ and so, $(x-1)$ is a factor of $x^3 + 4x^2 + 4x + 1$. By corollary 4.4.4, we get that 1 is a zero of $x^3 + 4x^2 + 4x + 1$.

4.4.6 corollary. A nonzero polynomial $f(x) \in F[x]$ of degree n has at most n zeros in F.

Proof: If $f(x)$ has no root in F then the corollary is true. So, suppose that $f(x)$ has at least one root in F. Proof is by induction on n. If $n = 1$, then $f(x) = ax + b$ and $-b/a$ is the only root of $f(x)$ in F. In this case, the corollary is true. Assume that the corollary is true for all polynomials of degree $n-1$. Let $a \in F$ be a root of $f(x)$. Then $f(x) = (x-a)q(x)$, where $q(x) \in F[x]$. Therefore

$\deg q(x) = n-1$. By induction hypothesis, $q(x)$ has at most $n-1$ roots in F. If $b \in F$ and b is a zero of $f(x)$ other than a, then $0 = f(b) = q(b)(b-a)$. since $b \neq a$, we get that $q(b)$=0. So b is a zero of $q(x)$. Thus every zero of $f(x)$ other than a is a zero of $q(x)$. Hence $f(x)$ has atmost n roots in F.

4.4.7 Example: Let $f(x) = x^3 + x^2 + 6x + 6 \in Z_7[x]$. Note that 1 is a zero of $f(x)$. By corollary 4.4.4, ($x-1$) is a factor of $f(x)$. Let us find the quotient by long division.

$$\begin{array}{r|l} & x^2+2x+1 \\ \hline x-1 & x^3+x^2+6x+6 \\ & \underline{x^3-x^2} \\ & \quad 2x^2+6x \\ & \quad \underline{2x^2-2x} \\ & \qquad\quad x+6 \\ & \qquad\quad \underline{x-1} \\ & \qquad\qquad 0 \end{array}$$

Thus $f(x) = (x-1)(x^2+2x+1)$. Obviously $x^2+2x+1 = (x+1)^2$. Therefore $x^3+x^2+6x+6 = (x-1)(x+1)^2$ in $Z_7[x]$

4.4.8 Example: For polynomials $f(x) = x^4 + 5x^3 - 3x^2$ and $g(x) = 5x^2 - x + 2$ in $Z_{11}[x]$, find the quotient and remainder when $f(x)$ is divided by $g(x)$.

Solution:

$$\begin{array}{r|l} & 9x^2+5x-1 \\ \hline 5x^2-x+2 & x^4+5x^3-3x^2 \\ & \underline{x^4-9x^3+7x^2} \\ & \quad 3x^3-10x^2 \\ & \quad \underline{3x^3-5x^2+10x} \\ & \qquad -5x^2-10x \\ & \qquad \underline{-5x^2+x-2} \\ & \qquad\qquad\qquad 2 \end{array}$$

Thus $f(x) = (9x^2 + 5x - 1)(5x^2 - x + 2) + 2$

$\therefore$ quotient $q(x) = 9x^2 + 5x - 1)$ and remainder $r(x) = 2$.

4.5 Irreducible Polynomials:

We start with the following important definition.

4.5.1 Definition: A nonconstant polynomial $f(x) \in F[x]$ is called irreducible over F (or an irreducible polynomial in $F[x]$, if $f(x)$ cannot be expressed as a product $g(x)h(x)$ of two polynomials $g(x)$ and $h(x)$ in $F[x]$ such that $\deg g(x) < \deg f(x)$ and $\deg h(x) < \deg f(x)$.

If $f(x) \in F[x]$ is a nonconstant polynomial that is not irreducible over F, then $f(x)$ is said to be reducible over F.

4.5.2. Examples: 1. Every first degree polynomial in $F[x]$ is irreducible over F. In particular, $2x + 2 \in Q[x]$ is irreducible over Q.

2. Since $x^2 + 1 = (x - i)(x + i)$ in $C[x], x^2 + 1$ is reducible over C.

Irreducible polynomials play an important role in the study of field theory. The problem of determining whether a given $f(x) \in F[x]$ is irreducible over F is difficult. We now give some criteria for determining irreducibility of quadratic and cubic polynomials.

4.5.3 Theorem: Let $f(x) \in F[x]$ be a polynomial of degree 2 or 3. Then $f(x)$ is reducible over F if and only if it has a zero in F.

Proof: It $f(x)$ is reducible over F, then $f(x) = f_1(x) f_2(x)$ where $f_1(x)$ and $f_2(x)$ are polynomials in $F[x]$ such that $\deg f_1(x) < \deg f(x)$ and $\deg f_2(x) < \deg f(x)$. But $\deg f(x) = \deg f_1(x) + \deg f_2(x)$. If $\deg f_1(x) > 1$ and $\deg f_2(x) > 1$ then $\deg f(x) \geq 4$, a contradiction. Therefore either $f_1(x)$ or $f_2(x)$ is of degree 1. If, say, $f_1(x)$ is of degree 1, then $f_1(x) = ax + b,$ where $a, b \in F$. Then $f_1(-ba^{-1}) = 0$ and hence $f(-ba^{-1}) = 0$, which proves that $-ba^{-1}$ is a zero of $f(x)$ in F.

Conversely, if $f(a) = 0$ for $a \in F$, then $x - a$ is a factor of $f(x)$, so $f(x) = (x - a)q(x)$, where $q(x) \in F[x]$. Since $\deg(x - a) < \deg f(x)$ and $\deg q(x) < \deg f(x)$, we get that $f(x)$ is reducible over F.

4.5.4 Examples: 1. Since the polynomial $x^2-2\in Q[x]$ has no zeros in Q, by theorem 4.5.3, x^2-2 is irreducible over Q.

2. Note that $\sqrt{2}$ is a zero of x^2-2 in R, by theorem 4.5.3, we get that x^2-2 is ireducible over R.

From examples (1) and (2), we observe that irreducibility depends on fields.

3. Since $x^3+3x+2\in Z_5[x]$ has no zeros in Z_5, by theorem 4.5.3, x^3+3x+2 is irreducible over Z_5.

We now state a theorem which is useful in proving some interesting theorems and whose proof is beyond the scope of this book.

4.5.5. Theorem: If $f(x)\in Z[x]$ then $f(x)$ factors into a product $g(x)h(x)$ of two polynomials $g(x)$ and $h(x)$ in $Q[x]$ such that $\deg g(x)=r<\deg f(x)$ and $\deg h(x)=s<\deg f(x)$ if and only if $f(x)$ factors into a product $u(x)\ v(x)$ of two polynomials $u(x)$ and $v(x)$ in $Z[x]$ such that $\deg k(x)=r<\deg f(x)$ and $\deg v(x)=s<\deg f(x)$.

4.5.6 Example: Let $f(x)=x^3+x+1\in Z_2[x]$. It can be checked that none of the elements of Z_2 is a zero of $f(x)$. So, $f(x)$ has no zero in Z_2 and by theorem 4.5.3, $f(x)$ is irreducible over Z_2.

4.5.7 Corollary: Let $f(x)=a_0+a_1x+...+a_{n-1}x^{n-1}+x^n\in Z[x]$ with $a_0\neq 0$. If $f(x)$ has a zero $a\in Q$, then $a\in Z$ and a divides a_0.

Proof: Since $a\neq 0$, we can write a as $a=\dfrac{\alpha}{\beta}$, where $\alpha,\beta\in Z$ and their gcd (α,β) =1. Then

$$a_0+a_1\left(\frac{\alpha}{\beta}\right)+...+a_{n-1}\left(\frac{\alpha^{n-1}}{\beta^{n-1}}\right)+\frac{\alpha^n}{\beta^n}=0$$

Multiply the above equation by β^{n-1} to obtain $a_0\beta^{n-1}+a_1\alpha\beta^{n-1}+...+a_{n-1}\alpha^{n-1}=-\dfrac{\alpha^n}{\beta}$. Because $\alpha,\beta\in Z$ it follows that $-\dfrac{\alpha^n}{\beta}\in Z$ since $(\alpha^n,\beta)=1$ and β divides α^n, we have that $\beta=\pm 1$. There $a=\pm\alpha\in Z$. The last equation shows that α/a_0 and hence a/a_0.

4.5.8 Example: Let us show that $f(x) = x^3 - x - 1 \in Q[x]$ is irreducible over Q. If $f(x)$ is reducible over Q, then $f(x)$ has a zero a in Q. By corollary 4.5.7 $a \in Z$ and $a/-1 \therefore a = \pm 1$. But $f(1) = 1-1-1 = -1 \neq 0$ and $f(-1) = -1+1-1 = -1 \neq 0$. So a is not a zero of $f(x)$, a contradiction. Hence $f(x)$ is irreducible over Q.

The question of deciding whether a given polynomial is irreducible or not can be a difficult and laborious one. Few criteria exist which declare that a given polynomial is or is not irreducible. One of these few is the following.

4.5.9 Theorem: (Eisenstein Criterion) Let $f(x) = a_0 + a_1 x + .. + a_n x^n \in Z[x], n \geq 1$. If there is a prime number p such that $p^2 \not| a_0, p \not| a_0, p \not| a_{n-1}$ and $p \times a_n$, then $f(x)$ is irreducible over Q.

Proof: Assume that $f(x)$ is reducible over Q then $f(x)$ factors into a product of two polynomials in $Q[x]$ of lower degrees r and s. By theorem 4.5.5, $f(x)$ has such a factorization with polynomials of the same degrees r and s in $Z[x]$. Accordingly $f(x) = (b_0 + b_1 x + ... + b_r x^2)(c_0 + c_1 x + ... + c_s x^s)$ with $b_i, c_i \in Z, b_r \neq 0, c_s \neq 0, r < n$ and $s < n$. Then $a_0 = b_0 c_0$ and $a_n = b_r c_s$. Clearly $r + s = n$. Since p/a_0 and $p^2 \not| a_0$, either p/b_0 and $p \not| c_0$ or p/c_0 and $p \not| b_0$. Consider the case p/c_0 and p/b_0.

Because p/a_n, it follows that p/b_r and p/cs. Let c_m be the first coefficient in $c_0 + c_1 x + ... + c_s x^s$ such that $p \not| c_m$. Observe that $a_m = b_0 c_m + ... + b_m c_0$ from this we get that $p \not| a_m$, since $p \not| b_0 c_m$ and $p/(b_1 c_{m-1} + ... + b_m c_0)$. By hypothesis, $m = n$. Thus $n = m \leq s < n$, which is impossible. Similarly if p/b_0 and $p \not| c_0$, we arrive at a contradiction. Therefore our assumption is wrong and $f(x)$ is irreducible over Q.

4.5.10 Examples: 1. Consider the polynomial $x^3 - 5x + 10 \in Z[x]$. 5 is a prime number such that $5 \not| 1, 5/-5, 5/10$ and $5^2 \not| 10$. By theorem 4.5.9, we get that $x^3 - 5x + 10$ is irreducible over Q.

2. Similarly, by using the prime number 3, we get, by theorem 4.5.9, that $25x^5 - 3x^4 - 3x^2 - 12$ is irreducible over Q.

3. The polynomial $\phi_p(x)=1+x+.....+x^{p-1}$ is irreducible over Q. Where p is any prime number.

Solution: Note that $\phi_p(x)=\dfrac{x^p-1}{x-1}$. Let

$$g(x)=\phi_p(x+1)=\frac{(x+1)^p-1}{(x+1)-1}=\frac{1}{x}\left(x^p+\binom{p}{1}x^{p-1}+....+\binom{p}{p-1}x\right)$$

$$=x^{p-1}+\binom{p}{1}x^{p-2}+....+\binom{p}{p-1}$$

Note that $p \Big/ \binom{p}{r}$, for $r=1,..,p-1, p\not|1$ and $p^2 \not| \binom{p}{p-1}$. By theorem 4.5.9, $g(x)$ is irrudible over Q. But if $\phi_p(x)=f(x)h(x)$ were a nontrivial factorization of $\phi_p(x)$ in $Z[x]$, then $\phi_p(x+1)=f(x+1)h(x+1)$ would give a nontrivial factorization of $g(x)$ in $Z[x]$. By theorem 4.5.5., $g(x)$ is a product of two nonconstant polynomials in $Q[x]$. This a contradiction, since $g(x)$ is irreducible over Q]. Thus $\phi_p(x)$ must also be irreducible over Q.

4. Prove that x^3+3x^2-8 is irreducible over Q.

Solution: Let $f(x)=x^3+3x^2-8$. If $f(x)$ is reducible over Q, by theorem 4.5.3, $f(x)$ has a zero a in Q. By corollary 4.5.7, $a\in Z$ and $a/8$. Therefore $a=\pm1,\pm,\pm4,\pm8$. But none of these are zeros of $f(x)$. This is a contradiction.

$\therefore f(x)$ is irreducible over Q.

4.6 Ideal Structure in F[x]:

We now introduce the notation $<a>=\{ra/r\in R\}$ to represent the ideal of all multiples of a.

4.6.1 Definition: Let R be a commutative ring with unity. An ideal I of R is called a principal ideal if $I=<a>$ for some $a\in R$.

4.6.2 Examples: 1. Every ideal of Z is a principal ideal.

Solution: Let I be an idea of Z. If $I=\{0\}$, then clearly $I=<0>$. So, assume that $I\neq\{0\}$. Choose the least positive integer m in I. clearly $<m>=\{km/k\in Z\}\leq I$. Conversely, if $h\in I$, then

$h = qm + r$ where $q, r \in Z$ with $0 \le r < m$. Since $h - qm \in I$, we get that $r \in Z$. The minimality of m implies $r = 0$ and $h = qm$, i.e. $h \in < m >$. Thus

$I \subseteq < m >$. Therefore $I = < m >$ and hence I is a principal ideal.

2. We know that $F[x]$ is a commutative ring with unity. Clearly $x \in F[x]$. Then the principal ideal $< x >$ is the set $\left\{ xf(x) / f(x) \in F[x] \right\}$

$\therefore$ the ideal $< x >$ consists of all polynomials in $F[x]$ having zero constant term.

The next theorem is an application of the division algorithm for $F[x]$.

4.6.3 Theorem: Every ideal in $F[x]$ is a principal ideal.

Proof: Let M be an ideal of $F[x]$. If $M = \{0\}$, then $M = < 0 >$. Assume that $M \ne \{0\}$. Choose a nonzero polynomial $g(x)$ in M of minimal degree. If the degree of $g(x)$ is 0, then $g(x)$ is a unit in $F[x]$. Therefore $M = < g(x) > = F[x]$. If the degree of $g(x) \ge 1$, we claim that $M = < g(x) >$. If $f(x) \in M$, then $f(x) = q(x)g(x) + r(x)$, where $r(x) = 0$ or deg $r(x) < \deg g(x)$. Since $f(x), g(x) \in M$, we get that $r(x) = f(x) - q(x)g(x) \in M$. Since $g(x)$ is a nonzero element of minimal degree in M, we have that $r(x) = 0$. Thus $f(x) = q(x)g(x) \in < g(x) >$. Hence $M = < g(x) >$.

We now characterize the maximal ideals of $F[x]$.

4.6.4 Theorem: Let $p(x)$ be a nonzero polynomial in $F[x]$. The $p(x)$ is irreducible over F if and only if $< p(x) >$ is a maximal ideal of $F[x]$.

Proof: If $< p(x) >$ is a maximal ideal of $F[x]$, then $< p(x) > \ne F[x]$. Therefore $p(x)$ is nonconstant polynomial. Let $p(x) = u(x)v(x)$ where $u(x), v(x) \in F[x]$. Then $< p(x) > \subseteq < u(x) >$.

Hence $< p(x) = < u(x) >$ or $< u(x) > = F[x]$. If $< u(x) > = F[x]$, then $u(x)$ is a unit, i.e. $\deg u(x) = 0$. If $< u(x) > = < p(x) >$, then $u(x) = p(x)g(x)$ and hence $p(x) = u(x)v(x) = p(x)g(x)v(x)$. since $F[x]$ is an integral domain $1 = g(x)v(x)$. Hence $v(x)$ is a unit, i.e. $\deg v(x) = 0$. Therefore $p(x)$ is irreducible over F.

Conversely if $p(x)$ is irreducible over F, then $< p(x) > \neq F\{x\}$. If $< p(x) > \subseteq I \subseteq F\{x\}$, where I is an ideal of $F[x]$. By theorem 4.6.3, $I = < f(x) >$ for some $f(x) \in F[x]$. Since $< p(x) > \subseteq < f(x) >$, we get the $p(x) = f(x)h(x)$, for some $h(x) \in F[x]$. Since $p(x)$ is irreducible, either $f(x)$ is a constant polynomial (whence $< f(x) > = F[x]$) or $h(x)$ is a constant polynomial (whence $< p(x) > = < f(x) >$). Thus we have that either $< p(x) > = < f(x) >$ or $< f(x) > = F[x]$. Hence $< p(x) >$ is a maximal ideal of $F[x]$.

We now prove an useful theorem.

4.6.5 Theorem: Let $p(x)$ be an irreducible polynomial in $F[x]$. If $p(x)$ divides $g(x)h(x)$, for some $g(x)$, $h(x) \in F[x]$, then either $p(x)$ divides $g(x)$ or $p(x)$ divides $h(x)$.

Proof: If $p(x)$ divides $g(x)h(x)$, then $g(x)h(x) \in < p(x) >$. Since $p(x)$ is irreducible, by theorem 4.6.4, $< p(x) >$ is a maximal ideal of $F[x]$. Every maximal ideal is a prime ideal. Therefore $< p(x) >$ is a prime ideal. Since $g(x)h(x) \in < p(x) >$, it follows that either $g(x) \in < p(x) >$ or $h(x) \in < p(x) >$, i.e. either $p(x)$ divides $g(x)$ or $p(x)$ divides $h(x)$.

4.6.6 Corollary: If $p(x) \in F[x]$ is irreducible over F and $p(x)$ divides the product $r_1(x)r_2(x)......r_n(x)$, where $r_i \in F[x]$ for $i = 1,..,n$, then $p(x)$ divides $r_i(x)$ for at least one;

Proof: We prove the corollary by induction on n. By theorem 4.6.5, the result is true for $n = 2$. Assume that the result is true for $n-1$. Let $r(x) = r_1(x)......r_{n-1}(x)$ then $p(x)$ divides $r(x)r_n(x)$. Again by theorem 4.6.5, either $p(x) / r(x)$ or $p(x) / r_n(x)$. If $p(x) / r(x)$, then by induction hypothesis $p(x) / r_{\dot r}(x)$ for some $1 \leq \dot r \leq n-1$. thus in either case we get that $p(x) / r_{\dot r}(x)$ for some i.

4.6.7 Theorem: If F is a field, then every nonconstant polynomial $f(x) \in F[x]$ can be factored in $F[x]$ uniquely (up to order and units) as a product of a finite number of polynomials in $F[x]$.

Proof: Let $f(x) \in F[x]$ be a nonconstant polynomial. We first prove that $f(x)$ can be written as the product of a finite number of irreducible polynomials in $F[x]$. The proof is by induction on $\deg f(x)$. if deg $f(x) = 1$, then $f(x)$ is irreducible and the result is true in this case. We assume that the result is true for all polynomials $g(x)$ in $F[x]$ such that $\deg g(x) < \deg f(x)$. On the basis of this assumption we aim to prove the result for $f(x)$. If $f(x)$ is irreducible, then there is nothing to prove. If $f(x)$ is not irreducible, then $f(x) = g(x)h(x)$, where $\deg g(x) < \deg f(x)$ and $\deg h(x) < \deg f(x)$. Then by our induction hypothesis $g(x)$ and $h(x)$ can be written as a product of a finite number of irreducible polynomials in $F[x]$; $g(x) = r_1(x).....r_n(x)$ and $h(x) = t_1(x).....t_m(x)$ where $r_i(x)$ and $r_{\dot{r}}(x)$ are irreducible polynomials in $F[x]$. Consequently $f(x) = g(x)h(x) = r_1(x).....r_n(x)t_1(x).....t_m(x)$ and in this way $f(x)$ has been factored as a product of a finite number of irreducible polynomials.

It remains for us to prove uniqueness. Suppose that $f(x) = p_1(x)p_2(x).....p_r(x) = q_1(x).....q_s(x)$ are two factorizations of $f(x)$ into irreducible polynomials. By corollary 4.6.6., $p_1(x)$ divides $q_i(x)$ for some i; since $p_1(x)$ and $q_i(x)$ are both irreducible polynomials in $F[x]$ and $p_1(x) / q_i(x)$, we have that $q_i(x) = u_1 p_1(x)$, where u_1 is a unit in $F[x]$.

Thus $p_1(x)p_2(x).....p_r(x) = u_1 p_1(x) q_1(x)...q_{i-1}(x)q_{i+1}....q_s(x)$; cancel off $p_1(x)$ and we are left with $p_2(x).....p_r(x) = u_1 q_1(x)....q_{i-1}(x)q_{i+1}....q_s(x)$. Repeat the argument on this relation with $p_2(x)$. After r steps the left side becomes 1, the right side is a product of a certain number of $q(x)$'s (the excess of s over r). This would force $r \le s$ since the $p(x)$'s are not units. Similarly $s \le r$, so that $r = s$. In the process we have also showed that every $p_i(x) = u_i q_{\dot{r}}(x)$, where u_i is a unit; for some $\dot{r}$.

4.7 Summary:

In this lesson you have learnt the division algorithm of $F[x]$, irreducible polynomials, Eisenstein's irreducibility criterion, ideals in $F[x]$ and factorization of polynomials over $F[x]$.

4.8 Technical terms/Named theorems:

Irreducible polynomial

Principal ideal

Division algorithm of $F[x]$

Remainder theorem

Factor Theorem

Eisenstein's irreducibility criterion

4.9 Model Examination Questions:

1. State and prove the division algorithm in $F[x]$.

2. Define an irreducible polynomial. Give an example of an irreducible polynomial.

3. Prove that a nonzero polynomial $p(x)$ in $F[x]$ is irreducible if and only it $< p(x) >$ is a maximal ideal.

4. State and prove Eisensteins irreducibility criterian.

5. For any prime number p, prove that the polynomial $\phi_p(x) = 1 + x + ... + x^{p-1}$ is irreducible over Q.

6. Prove that every ideal in $F[x]$ is a principal ideal.

7. Prove that every nonconstant polynomial in $F[x]$ can be factored in $F[x]$ uniquely (upto order and units) as a product of a finite number of irreducible polynomials in $F[x]$.

8. Find all prime numbers p such that $x+2$ is a factor of $x^4 + x^3 + x^2 - x + 1 \in Z_p(x)$

9. Let $f(x) \in F[x]$ be a polynomial of degree 2 or 3. Prove that $f(x)$ is reducible over F if and only if it has a zero in F.

10. Let $f(x) = a_0 + a_1 x + ... + a_{n-1} x^{n-1} + x^n \in Z[x]$ with $a_0 \neq 0$. If $f(x)$ has a zero $a \in Q$ then prove that $a \in Z$ and a/a_0.

4.10 Exercises:

1. Determine which of the following are irreducible over Q.

   a) $2x^5 - 6x^3 + 9x^2 - 15$

   b) $x^4 - 3x^2 + 9$

   c) $3x^5 - 7x^4 + 7$

2. Prove that

   a) $x^2 + 1$ is irreducible over Z_7.

   b) $x^2 + x + 1$ is irreducible over Z_2

c) x^2+x+4 is irreducible over Z_{11}

3. Find all prime numbers p such that $x+2$ is a factor of $x^4+x^3+x^2-x+1\in Z_p[x]$

4. Show that x^2+8x-2 is irreducible over Q.

5. For polynomials $f(x)=x^6+3x^5+4x^2-3x+2$ and $g(x)=3x^2+2x-3$ in $Z_7[x]$, find the quotient $q(x)$ and remainder $r(x)$.

6. For polynomials $2x^7+x^6-3x^5+4x^3-x+5$ and x^2-2x+4 in $Q[x]$, find the quotient $q(x)$ and remainder $r(x)$.

7. If p is a prime number, prove that the polynomial x^n-p is irreducible over Q.

8. Determine whether the following polynomials in $Z[x]$ satisfies an Eisenstein criterion for irreducibility over Q.

a) x^2-12

b) $8x^3-6x^2-9x+24$

c) $4x^{10}-6x^3+24x-18$

d) $2x^{10}-25x^3+10x^2-30$

4.11 Model Practical Problems with Solution:

Problem: The polynomial $2x^3+3x^2-7x-5$ can be factored into linear factors in $Z_{11}[x]$. Find this factorization.

AIM: To find factorization of the polynomial $2x^3+3x^2-7x-5$ into linear factors in $Z_{11}[x]$.

Hypothesis : The polynomial $2x^3+3x^2-7x-5$ can be factored into linear factors in $Z_{11}[x]$.

Solution: By Corollary 4.4.4., a linear polynomial $(x-a)$ is a factor of a polynomial $f(x)$ in $F[x]$ if and only if a is a zero of $f(x)$ in F. So, by hypothesis, all the zeros of $2x^3+3x^2-7x-5$ are in Z_{11}. On verification, we get that 3 is a zero of the given polynomial, since $2.3^3+3.3^2-7.3-5=0$ in Z_{11}. Therefore $x-3$ divides $2x^3+3x^2-7x-5$

Let us find the factorization by long division.

$$
\begin{array}{r|l}
 & 2x^2+9x-2 \\
\hline
x-3 & 2x^3+3x^2-7x-5 \\
 & \underline{2x^3-6x^2} \\
 & \quad 9x^2-7x \\
 & \quad \underline{9x^2-5x} \\
 & \qquad -2x-5 \\
 & \qquad \underline{-2x+6} \\
 & \qquad\qquad 0
\end{array}
$$

$\therefore 2x^3+3x^2-7x-5=(x-3)(2x^2+9x-2)$

Since - 3 is a zero of $2x^2+9x-2$, we can divide this polynomial by $x+3$

$$
\begin{array}{r|l}
 & 2x+3 \\
\hline
x+3 & 2x^2+9x-2 \\
 & \underline{2x^2+6x} \\
 & \quad 3x-2 \\
 & \quad \underline{3x+9} \\
 & \qquad 0
\end{array}
$$

Conclusion: $2x^3+3x^2-7x-5=(x-3)(x+3)(2x+3)$ is the required factorization.

4.12 Problems for Practicals:

1. Find the remainder and quotient, when the polynomial $x^8+5x^7+3x^6-2x^5+x^4-8x^3+6x^2-2x+1$ is divided by $x^4-x^3+3x^2-1$ over $Q[x]$.

2. Find the remainder and quotient when the polynomial $x^7-2x^6+4x^5+x^3+2x^2-x+3$ is divided by $3x^4+2x^3+4x^2-1$ over $Q[x]$.

3. If $f(x)=x^4+5x^3+3x^2+2x+1$ and $q(x)=x^2+x-1$ are polynomials in $Z_5[x]$ then find the quotient and remainder when $f(x)$ is divided by $g(x)$.

4. Show that the polynomial x^2+1 is irreducible over the field of real numbers and reducible over the field of complex numbers.

5. Show that the polynomial $x^2+x-4 \in Z_{11}[x]$ is irreducible over Z_{11}.

6. Resolve x^4+4 into linear factors in Z_5.

7. Show that $f(x)=x^2+8x-2$ is irreducible over the field of rational numbers. Is it irreducible over the field of real numbers? Give reasons for your answer.

8. If $f(x)=1+x+x^2+x^3+x^4$, prove that $f(x)$ is irreducible over Q.

9. The polynomial x^3+2x^2+2x+1 can be factored into linear factors in $Z_7[x]$. Find this factorization.

10. (a) Let F be a field and $f(x), g(x) \in F[x]$. Show that $f(x)$ divides $g(x)$ if and only if $g(x) \in < f(x) >$

(b) Show that the polynomial x^4+2 is irreducible over Q.

- Prof. Y. Venkateswara Reddy

LESSON - 5

VECTOR SPACES

5.1 Objective of the Lesson:

To learn the definition of vector space, subspace, some examples of vector spaces, algebra of subspaces and related theorems.

5.2 Structure

5.3 Introduction:

In this lesson we introduce vector space, subspace, linear sum of subspaces, linear span of a set, linear combination of vectors, linearly dependent and independent vectors.

5.4 Vector Spaces:

Let F be a field. A vector space over F is an additive abelian group V together with a function $F \times V \to V((a,\alpha) \to a\alpha)$ such that

i) $a(\alpha+\beta) = a\alpha + b\beta$

ii) $(a+B)\alpha = a\alpha + b\alpha$

iii) $a(B\alpha) = (ab)\alpha$

iv) $1\alpha = \alpha$ for all $\alpha, \beta \in v; a, b \in F$ and 1 is the unity element of F.

5.4.4 Note: 1) If V is a vector space over F we write V(F) is a vector space. If the field is understood we simply say V is a vector space.

2. Elements of V are called vectors and elements of F are called scalars.

3. The internal composition + in V is called addition and the external composition . is called scalar multiplication.

5.4.5 Example: 1. Let F be a field and K be a subfield of F then F is a vector space over K.

Solution: Suppose F is a field and K a subfield of F.

F is a field $\Rightarrow (F,+)$ is an abelian group.

Let $a \in K$ and $\alpha \in F \Rightarrow a \in F$ and $\alpha \in F$ Since $K \subseteq F$

$\Rightarrow a.\alpha \in F$ Since F is a field.

$\therefore .$ is an external composition in F over K.

Let $a,b \in K$ and $\alpha, \beta \in F \Rightarrow a,b \in F$ and $\alpha, \beta \in F$

$a(\alpha+\beta) = a\alpha + b\beta$ by distributive law

$(a+b)\alpha = a\alpha + b\alpha$ by distributive law

$a(b\alpha) = (ab)\alpha$ by associative law

$1\alpha = \alpha$ Since 1 is the unity element of F.

$\therefore F$ is a vector space over the subfield K.

5.4.6 Note: Every field F is a subfield of itself

$\therefore$ Every field is a vector space over itself.

5.4.7 Example: Let F be a field $V_n = \{(a_1, a_2 ... a_n)/a_i \in F for 1 \le i \le n\}$ is the set of n-tuples. V_n is a vector space over F with respect to addition '+' and scalar multiplication '•' defined by

$(a_1, a_2 ... a_n) + (b_1, b_2 ... b_n) = (a_1 + b_1, ... a_n + b_n)$

$a(a_1, a_2 ... a_n) = (aa_1, aa_2, ... aa_n)$ for $(a_1, a_2 ... a_n), (b_1, b_2 ... b_n) \in V_n$ and $a \in F$.

Solution: Let F be a field and $V_n = \{(a_1, a_2 ... a_n)/a_i \in F, 1 \le i \le n\}$

Let $\alpha = (a_1, a_\alpha ... a_n), \beta = (b_1, b_\alpha ... b_n), r = (c_1, c_2, ... c_n) \in V_n \Rightarrow a_i$'s, b_i's and c_i's $\in F$

$\alpha + \beta = (a_1 + b_1, a_2 + b_2 ... a_n + b_n) \in V_n$ since $a_i + b_i \in F$ for $1 \le i \le n$

$\therefore V_n$ is closed w.r.to '+'.

$$(\alpha+\beta)+\gamma=\left[(a_1,a_2...a_n)+(b_1,b_2...b_n)\right]+(c_1,c_2...c_n)$$

$$=(a_1+b_1,a_2+b_2...a_n+b_n)+(c_1,c_2...c_n)$$

$$=\left[(a_1+b_1)+c_1(a_2+b_2)+c_2,...(a_n+b_n)+c_n\right]$$

$$=\left[a_1+(b_1+c_1),a_2+(b_2+c_2),...a_n+(b_n+c_n)\right]$$

$$=(a_1,a_2...a_n)+(b_1+c_1,b_2+c_2,...b_n+c_n)$$

$$=\alpha+(\beta+\gamma)$$

$\therefore$ Addition is associative.

$$\alpha+\beta=(a_1,a_2...a_n)+(b_1,b_2...b_n)$$

$$=(a_1,+b_1,a_2+b_2...a_n+b_n)$$

$$=(b_1+a_1,b_2+a_2...b_n+a_n) \quad \text{since} \quad a_is,b_is\in F$$

$$=(b_1b_2...b_n)+(a_1,a_2...a_n)$$

$$=\beta+\alpha$$

$\therefore$ Addition is commutative.

We have $0\in F$

$$\therefore \overline{0}=(0,0...0)\in V_n$$

Now $\alpha+\overline{0}=(a_1,a_2...a_n)+(0,0...0)$

$$=(a_1+0,a_2+0...a_n+0)$$

$$=(a_1,a_2...a_n)$$

$$=\alpha$$

$\therefore \overline{0}$ is the additive identity of V_n.

$$\alpha=(a_1,a_2...a_n)\in V_n \Rightarrow a_1,a_2...a_n\in F$$

$$\Rightarrow -a_1,-a_2...-a_n\in F$$

$$\Rightarrow (-a_1,-a_2...-a_n)\in V_n$$

Now $(a_1, a_2 ... a_n) + (-a_1, -a_2 ... - a_n)$

$$= (a_1 + -(a_1), a_2 + (-a_2), ... a_n + (-a_n)$$

$$= (0, 0 ... 0)$$

$$= \overline{0}$$

$\therefore (-a_1, -a_2 ... a_n) = -\alpha$ is the inverse of α.

Hence $(V_n, +)$ is an abelian group.

$a\alpha = a(a_1, a_2 ... a_n) = (aa_1, aa_2 ... aa_n) \in V_n$, Since $aa_1, aa_2 ... aa_n \in F$

$\therefore V_n$ is closed under scalar multiplication.

Let $a, b \in F$ and $\alpha = (a_1, a_2 ... a_n), \beta = (b_1, b_2 ... b_n) \in V_n$

$$a(\alpha + \beta) = a(a_1 + b_1, a_2 + b_2 ... a_n + b_n)$$

$$= \left(a(a_1 + b_1), a(a_2 + b_2) ... a(a_n + b_n)\right)$$

$$= \left(aa_1 + ab_1, aa_2 + ab_2 ... aa_n + ab_n\right)$$

$$= \left(aa_1, aa_2 ... aa_n) + (ab_1, ab_2 ... ab_n\right)$$

$$= a(a_1, a_2 ... a_n) + a(b_1, b_2 ... b_n)$$

$$= a\alpha + a\beta$$

$$(a + b)\alpha = (a + b)(a_1, a_2 ... a_n)$$

$$= \left((a + b)a_1, (a + b)a_2 ... (a + b)a_n\right))$$

$$= \left(aa_1 + ba_1, aa_2 + ba_2, ... aa_n + ba_n\right)$$

$$= \left(aa_1, aa_2 ... aa_n\right) + \left(ba_1, ba_2 ... ba_n\right)$$

$$= a\left(a_1, a_2 ... a_n\right) + b\left(a_1, a_2 ... a_n\right)$$

$$= a\alpha + b\alpha$$

$a(b\alpha) = a\left(ba_1, ba_2 ... ba_n\right)$

$= \left(a(ba_1), a(ba_2), ... a(ba_n)\right)$

$= \left((ab)a_1, (ab)a_2, ... (ab)a_n\right)$

$= ab\left(a_1, a_2 ... a_n\right)$

$= ab\alpha$

$1\alpha = 1\left(a_1, a_2 ... a_n\right)$

$= \left(1a_1, 1a_2 ... 1a_n\right)$

$= \left(a_1, a_2 ... a_n\right)$

$= \alpha$

$\therefore V_n(F)$ is a vector space.

5.4.8 Elementary Properties of a vector space:

Theorem : Let $V(F)$ be a vector space then

i) $a\overline{0} = \overline{0}$

ii) $0\alpha = \overline{0}$

iii) $a(-\alpha) = -a\alpha$

iv) $(-a)\alpha = -a\alpha$

v) $(-a)(-\alpha) = a\alpha$

vi) $a(\alpha - \beta) = a\alpha - a\beta$

vii) $(a - b)\alpha = a\alpha - b\alpha$

viii) $\alpha = \overline{0} \Rightarrow a = 0$ or $\alpha = \overline{0}$ $\forall \alpha, \beta \in V$ and $a \in F$

Let $V(F)$ be a vector space and $\alpha.\beta \in \vee, a \in F$

i) $a\overline{0} = a(\overline{0} + \overline{0}) = a\overline{0} + a\overline{0}$

$\Rightarrow a\overline{0} + \overline{0} = a\overline{0} + a\overline{0}$

$\Rightarrow \overline{0} = a\overline{0}$ by left cancellation law.

ii) $0\alpha = (0+0)\alpha$ since 0 is the zero element of F.

$= 0\alpha + 0\alpha$

$\Rightarrow 0\alpha + \overline{0} = 0\alpha + 0\alpha$ Since $\overline{0}$ is the identity in V.

$\Rightarrow \overline{0} = 0\alpha$ by left cancellation law.

iii) $a(-\alpha) + a\alpha = a(-\alpha + \alpha)$

$= a\overline{0}$

$= \overline{0}$ by i.

$\therefore a(-\alpha) = -a\alpha$

iv) $(-a)\alpha + a\alpha = (-a + a)\alpha$

$= 0\alpha$

$= \overline{0}$ by ii.

$\therefore (-a)\alpha = -a\alpha$

v) $(-a)(-\alpha) = -\left[(-a)\alpha\right]$ by iii.

$= -\left[-(a\alpha)\right]$ by iv.

$= a\alpha$

vi) $a(\alpha - \beta) = a\left[\alpha + (-\beta)\right]$

$= a\alpha + a(-\beta)$

$= a\alpha + (-a\beta)$

$= a\alpha - a\beta$

vii) $(a-b)\alpha = \left[a + (-b)\right]\alpha = a\alpha + (-b)\alpha$

$= a\alpha + (-b\alpha)$

$= a\alpha - b\alpha$

Suppose $a\alpha = \overline{0}$.

If $a \neq 0$ then there exists $a^{-1} \in F$ such that $aa^{-1} = a^{-1}a = 1$.

$\therefore a\alpha = \overline{0}$ and $a \neq 0$

$\Rightarrow a^{-1}(a\alpha) = a^{-1}\overline{0}$

$\Rightarrow (a^{-1}a)\alpha = \overline{0}$

$\Rightarrow 1\alpha = \overline{0}$

$\Rightarrow \alpha = \overline{0}$

$\therefore a\alpha = \overline{0} \Rightarrow$ either $a = 0$ or $\alpha = \overline{0}$.

5.4.9 Note: Let $V(F)$ be a vector space. For $a, b \in F$ and $\alpha, \beta \in V$

i) $a\alpha = b\alpha$ and $\alpha \neq \overline{0}$ then $a = b$

ii) $a\alpha = a\beta$ and $a \neq 0$ then $\alpha = \beta$

5.5. Vector Subspaces:

5.5.1 Definition: Let $V(F)$ be a vector space and $\phi \neq W \subseteq V$.

W is said to be a subspace of V if W is an additive subgroup of V and $a\beta \in W$, for all $a \in F, \beta \in W$.

5.5.2 Example: (i) Let $V(F)$ be a vector space then $W = \{\overline{0}\}$

where $\overline{0}$ is the additive identity of V is a subspace of V.

2. The vector space V itself is a subspace of V.

These two subspaces are called trivial subspaces of V.

5.5.3 Theorem: A necessary and sufficient conditions for a non-empty subset W of a vector space $V(F)$ to be a subspace are

i) $\alpha, \beta \in \omega \Rightarrow \alpha - \beta \in \omega$

ii) $a \in F, \alpha \in \omega \Rightarrow a\alpha \in \omega$

Proof: Let $V(F)$ be a vector space and W a non-empty subset of V.

Suppose W is a subspace of V.

$\Rightarrow W$ itself is a vector space.

$\Rightarrow W$ is a group w.r.to +

$\therefore \alpha, \beta \in W \Rightarrow \alpha - \beta \in W$.

Clearly $a \in F$ and $\alpha \in W \Rightarrow a\alpha \in W$

$\therefore$ The conditions i and ii are necessary.

Now suppose i) $\alpha, \beta \in W \Rightarrow \alpha - \beta \in W$ and

ii) $a \in F$ and $\alpha \in W \Rightarrow a\alpha \in W$

W is a non-empty subset of the abelion group $(V,+)$ and $\alpha, \beta \in W \Rightarrow \alpha - \beta \in W$.

$\therefore$ From group theory we know that (W, +) is a subgroup of $(V,+)$.

From condition (ii) we get than W is a vector space by itself.

Hence W is a subspace of $V(F)$.

5.5.4 Theorem: A necessary and sufficient condition for a non-empty subset W of a vector space $V(F)$ to be a subspace of $V(F)$ is $a, b \in F$ and $\alpha, \beta \in W \Rightarrow a\alpha + b\beta \in W$.

Proof: Let $V(F)$ be a vector space and W is a non-empty subset of V.

Suppose W is a subspace of $V(F)$.

$\Rightarrow W$ itself is a vector space over F.

$\Rightarrow W$ is closed under addition '+' and scalar multiplication '. '.

$\therefore a, b \in F$ and $\alpha, \beta \in W \Rightarrow a\alpha \in W, b\beta \in W$

$$\Rightarrow a\alpha + b\beta \in W$$

$\therefore$ The condition is necessary.

Suppose the condition $a, b \in F$ and $\alpha, \beta \in W \Rightarrow a\alpha + b\beta \in W$.

To prove that W is a subspace of $V(F)$.

Let $\alpha, \beta \in W$

Since F is a field 1 and $-1 \in F$

$\therefore$ By the condition $1.\alpha + (-1)\beta \in W$

$$\Rightarrow \alpha - \beta \in W$$

$\therefore (W,+)$ is a sub group of $(V,+)$ from group theory

$\therefore (W,+)$ is a group.

All elements of W are elements of V.

$\therefore (W,+)$ is an abelion group.

Let $a \in F$ and $\alpha \in W$

$\therefore$ By the condition $a\alpha + 0.\alpha \in W$ since $0 \in F$

$$\Rightarrow a\alpha \in W$$

$\therefore W$ is closed under scalar multiplication.

Since elements of W are elements of V, we have

$$a(\alpha + \beta) = a\alpha + a\beta$$

$$(a+b)\alpha = a\alpha + b\alpha$$

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

$1.\alpha = \alpha$ for all $a,b \in F$ and $\alpha, \beta \in W$

$\therefore W$ is a vector space over F.

Hence W is a subspace of $V(F)$.

$\therefore$ The condition is sufficient.

5.5.4 Algebra of Subspaces:

Theorem: The intersection of any two subspaces of a vector space is a subspace.

Proof: Let $V(F)$ be a vector space and W_1, W_2 be two subspaces of V.

To prove that $W_1 \cap W_2$ is a subspace.

Take $W = W_1 \cap W_2$

clearly $\overline{0} \in W_1$ and $\overline{0} \in W_2 \Rightarrow \overline{0} \in W_1 \cap W_2 = W$

$\therefore W$ is a non-empty subset of V.

Let $a,b \in F$ and $\alpha, \beta \in W$

$\alpha, \beta \in W \Rightarrow \alpha, \beta \in W_1$ and $\alpha, \beta \in W_2$

$a,b \in F, \alpha, \beta \in W_1$ and W_1 is a subspace of V.

$\Rightarrow a\alpha + b\beta \in W_1$ .... (1)

$a,b \in F, \alpha, \beta \in W_2$ and W_2 is a subspace of V.

$\Rightarrow a\alpha + b\beta \in W_2$ .... (2)

From (1) and (2) $a\alpha + b\beta \in W_1 \cap W_2 = W$

$\therefore a,b \in F, \alpha, \beta \in W \Rightarrow a\alpha + b\beta \in W$

$\therefore W$ is a subspace of V. by 5.5.4.

5.5.6 Note: 1. Intersection of any family of subspaces of a vector space is a subspace.

2. Union of two subspaces of a vector space need not be a subspace.

Example: We know that set of real numbers R is a field.

$\therefore R^3 = V_3(R) = \{(a_1, a_2, a_3)/a_1, a_2, a_3 \in R\}$ is a vector space by 5.4.7.

Let $\therefore W = \{(0, x, 0)/x \in R\}$ and $W_2 = \{(0,0,y)/y \in R\}$

Clearly W_1 and W_2 are non-empty subsets of $V_3(R)$

Let $a,b \in R$ and $\alpha = (0, x_1, 0)\beta = (0, x_2, 0) \in W$;

Then $a\alpha + b\beta = a(0, x_1, 0) + b(0, x_2, 0)$

$= (0, ax_1, 0) + (0, bx_2, 0)$

$= (0, ax_1 + bx_2, 0)$

$\therefore a\alpha + b\beta \in W_1$ since $a, b, x_1, x_2 \in R \Rightarrow ax_1 + bx_2 \in R$

Similarly $a,b \in R$ and $\gamma = (0, 0, y_1)$ $\delta = (0, 0, y_2) \in W_2$

$a\gamma + b\delta = a(0, 0, y_1) + b(0, 0, y_2)$

$= (0, 0, ay_1) + (0, 0, by_2)$

$= (0, 0, ay_1 + by_2)$

$a\gamma + b\delta \in W_2$ since $a, b \in R, y_1, y_2 \in R \Rightarrow ay_1 + by_2 \in R$

Hence W_1 and W_2 are subspaces of $V_3(R)$

Now $\alpha = (0,2,0) \in W_1$ and $\beta = (0,0,3) \in W_2$

$$\alpha + \beta = (0,2,3) \notin W_1 \cup W_2$$

$\therefore W_1 \cup W_2$ need not be a subspace.

5.5.7 Theorem: Union of two subspaces of a vector space is a subspace iff one is contained in the other.

Proof: Let W_1 and W_2 be two subspaces of a vector space $V(F)$.

Suppose $W_1 \cup W_2$ is a subspace of the vector space $V(F)$.

To prove that either $W_1 \subseteq W_2$ or $W_2 \subseteq W_1$

If possible assume that $W_1 \not\subseteq W_2$ or $W_2 \not\subseteq W_1$

$W_1 \not\subseteq W_2 \Rightarrow$ there is an $\alpha \in W_1$ and $\alpha \notin W_2$ .... (1)

$W_2 \not\subseteq W_1 \Rightarrow$ there is a $\beta \in W_2$ and $\beta \notin W_1$ .... (2)

$\alpha \in W_1 \Rightarrow \alpha \in W_1 \cup W_2$

$\beta \in W_2 \Rightarrow \beta \in W_1 \cup W_2$

$\therefore \alpha + \beta \in W_1 \cup W_2$ since $W_1 \cup W_2$ is a subspace.

$\Rightarrow \alpha + \beta \in W_1$ or $\alpha + \beta \in W_2$

If $\alpha + \beta \in W_1$ then $\alpha + \beta - \alpha \in W_1$ since $\alpha \in W_1$

$$\Rightarrow \beta \in W_1$$

A contradiction to (2)

If $\alpha + \beta \in W_2$ then $\alpha + \beta - \beta \in W_2$ since $\beta \in W_2$

$$\Rightarrow \alpha \in W_2$$

A contradiction to (1)

$\therefore \alpha+\beta \notin W_1$ and $\alpha+\beta \notin W_2$

$\Rightarrow \alpha+\beta \notin W_1 \cup W_2$ A contradiction to $\alpha+\beta \in W_1 \cup W_2$.

Hence either $W_1 \subseteq W_2$ or $W_2 \subseteq W_1$

Now suppose either $W_1 \subseteq W_2$ or $W_2 \subseteq W_1$

If $W_1 \subseteq W_2$ then $W_1 \cup W_2 = W_2$

If $W_2 \subseteq W_1$ then $W_1 \cup W_2 = W_1$

$\therefore W_1 \cup W_2$ is a subspace.

5.6.1 Linear sum of two subspaces:

Definition: Let $V(F)$ be a vector space and W_1, W_2 be two subspaces of $V(F)$. The set $\{\alpha_1+\alpha_2 / \alpha, \in W_1, \alpha_2 \in W_2\}$ is called linear sum of W_1 and W_2 and is denoted by W_1+W_2.

5.6.2 Theorem: If W_1 and W_2 are two subspaces of a vector space $V(F)$ then W_1+W_2 is a subspace of V. Also $W_1 \cup W_2 \subseteq W_1+W_2$.

Proof: Let $V(F)$ be a vector space and W_1, W_2 be two subspaces of V.

$$W_1+W_2 = \{\alpha_1+\alpha_2 / \alpha_1 \in W_1, \alpha_2 \in W_2\}$$

Clearly $\overline{0} \in W_1$ and W_2

$\therefore \overline{0} = \overline{0}+\overline{0} \in W_1+W_2$

$\therefore W_1+W_2$ is a non-empty subset of V.

Suppose $a, b \in F$ and $\alpha, \beta \in W_1+W_2$

$\alpha \in W_1+W_2 \Rightarrow \alpha = \alpha_1+\alpha_2$ where $\alpha_1 \in W_1$ and $\alpha_2 \in W_2$

$\beta \in W_1+W_2 \Rightarrow \beta = \beta_1+\beta_2$ where $\beta_1 \in W_1$ and $\beta_2 \in W_2$

Now $a\alpha+b\beta = a(\alpha_1+\alpha_2)+b(\beta_1+\beta_2)$

$$= (a\alpha_1+a\alpha_2)+(b\beta_1+\beta_2)$$

$= (a\alpha_1 + b\beta_2) + (a\alpha_2 + b\beta_2)$

$a,b \in F$ and $\alpha_1, \beta_1 \in W_1 \Rightarrow a\alpha_1 + b\beta_1 \in W_1$ since W_1 is a subspace.

$a,b \in F$ and $\alpha_2, \beta_2 \in W_2 \Rightarrow a\alpha_2 + b\beta_2 \in W_2$ since W_2 is a subspace.

$\therefore a\alpha + b\beta = (a\alpha_1 + b\beta_1) + (a\alpha_2 + b\beta_2) \in W_1 + W_2$

Hence $W_1 + W_2$ is a subspace of V.

Let $\alpha_1 \in W_1 \Rightarrow \alpha_1 + \overline{0} \in W_1 + W_2$

$\Rightarrow \alpha_1 \in W_1 + W_2$

$\therefore W_1 \subseteq W_1 + W_2$ .... (1)

$\alpha_2 \in W_1 \Rightarrow \overline{0} + \alpha_2 \in W_1 + W_2$

$\Rightarrow \alpha_2 \in W_1 + W_2$

$\therefore W_2 \subseteq W_1 + W_2$ ....(2)

From (1) and (2) $W_1 \cup W_2 \subseteq W_1 + W_2$

5.6.3 Linear Combination : Let $V(F)$ be a vector space and $\alpha_1, \alpha_2 ... \alpha_n \in V$ if $a_1, a_2 ... a_n \in F$. then the vector $a_1\alpha_1 + a_2\alpha_2 + ... a_n\alpha_n$ is called linear combination of $\alpha_1, \alpha_2 |+\alpha_n| ...... \alpha_n$.

5.6.4 Linear Span of a set: Let V(F) be a vector space and S a non-empty subset of V. The set of all linear combinations of elements of all possible finite subsets of S is said to be linear span of S.

It is denoted by L (S).

5.6.5 Note: 1. If S is a non-empty subset of a vector space V then

$L(S) = \{a_1\alpha_1 + a_2\alpha_2 + ... + a_n\alpha_n / a_i \in F$ and $\alpha_i \in S$ for $1 \le i \le n\}$

2. S is a subset of L (S).

5.6.6. Theorem: The linear span L (S) of any subset S of a vector space V is a subspace of V.

Proof: Let V(F) be a vector space and S a non-empty subset of V.

$L(S)$ is the linear spance of S.

Clearly by the defintion of linear spance $L(S)$ is a non-empty subset of V.

Let $\alpha, \beta \in L(S)$ and $a, b \in F$

$\alpha \in L(S) \Rightarrow \alpha = a_1\alpha_1 + a_2\alpha_2 + ... + a_m\alpha_m$ where a_i's $\in F$ and α_i's $\in S$

$\beta \in L(S) \Rightarrow \beta = b_1\beta_1 + b_2\beta_2 + ... + b_n\beta_n$ where b_i's $\in F$ and β_i's $\in S$.

Now $a\alpha + b\beta = (a_1\alpha_1 + a_2\alpha_2 + ... + a_m\alpha_m) + b(b_1\beta_1 + b_2\beta_2 + ... + b_n\beta_n)$

$$= (aa_1)\alpha_1 + (aa_2)\alpha_2 + ... + (aa_m)\alpha_m + (bb_1)\beta_1 + (bb_2)\beta_2 + ... + (bb_n)\beta_n$$

i.e. $a\alpha + b\beta$ is a linear combination of elements of S.

$\therefore a\alpha + b\beta \in L(S)$

Hence L (S) is a subspace of V.

5.6.7 Theorem: If S is a non-empty sub-set of a vector space V(F) then linear span of S is the intersection of all sub-spaces of V which contain S.

Proof: Let $V(F)$ be a vector space and S a non-empty sub-set of V.

Suppose W is a sub-space of V and $S \subseteq W$.

Let $\alpha \in L(S)$

$\Rightarrow \alpha = a_1\alpha_1 + a_2\alpha_2 + ... + a_n\alpha_n$ where a_i's $\in F$ and α_i's $\in S$.

$\alpha_1, \alpha_2, ...\alpha_n \in S \Rightarrow \alpha_1, \alpha_2, ...\alpha_n \in W$

$\Rightarrow a_1\alpha_1 + a_2\alpha_2 + ... + a_n\alpha_n \in W$ Since W is a subspace it is closed under addition and scalar multiplication.

$\therefore \alpha \in L(S) \Rightarrow \alpha \in W$

i.e. $L(S) \subseteq W$

$\therefore L(S)$ is a subset of every subspace which contains S

$\Rightarrow L(S) \subseteq$ Inter section of all subspaces of V. Which contains S .... (1)

We know that $L(S)$ is a subspace and $S \subseteq L(S)$ by theorem 5.6.6 and 5.6.5.

$\therefore$ Intersection of all subspaces of V which contains $S \subseteq L(S)$ .... (2)

From (1) and (2) $L(S)$ = Intersection of all subspaces of V.

Which contain S.

5.6.8 Note: If S is a non-empty sub-set of a vector space $V(F)$ then $L(S)$ is the smallest subspace of V containing S.

5.6.9 Theorem: If S is a non-empty subset of a vector space $V(F)$ then

i) S is a subspace of $V \Leftrightarrow L(S) = S$

ii) $L(L(S)) = L(S)$

Proof: Let $V(F)$ be a vector space and S be a non-empty subset of V.

i) Suppose S is a subspace of V.

Let $\alpha \in L(S)$

$\Rightarrow \alpha = a_1\alpha_1 + a_2\alpha_2 + ... + a_n\alpha_n$ where a_i's$\in F$ and α_i's$\in S$.

S is a subspace $\Rightarrow$ It is closed under addition and scalar multiplication.

$\therefore \alpha \in S$

$\therefore \alpha \in L(S) \Rightarrow \alpha \in S$

$\therefore L(S) \subseteq S$ .... (1)

Let $\beta \in S \Rightarrow 1.\beta \in L(S)$

$\Rightarrow \beta \in L(S)$

$\therefore S \subseteq L(S)$ .... (2)

Hence $L(S) = S$. From (1) and (2)

Now suppose $L(S) = S$

By theorem 5.6.6 $L(S)$ is a subspace of V.

$\therefore S$ is a subspace of V.

ii) We know that $L(S)$ is a subspace of V by 5.6.6.

$\therefore$ By (1) $L(L(S)) = L(S)$

5.6.10 Theorem: If S and T are two subsets of a vector space $V(F)$ then

(i) $S \subseteq T \Rightarrow L(S) \subseteq L(T)$

(ii) $L(S \cup T) = L(S) + L(T)$

Proof: Let $V(F)$ be a vector space and S, T be two subsets of V.

i) Suppose $S \subseteq T$

Let $\alpha \in L(S)$

$\Rightarrow \alpha = a_1\alpha_1 + a_2\alpha_2 + ... + a_n\alpha_n$ where a_i's$\in F$ and α_i's$\in S$.

i.e. α is linear combination of finite subset of S.

$\Rightarrow \alpha$ is linear combination of finite subset of T. Since $S \subseteq T$

$\Rightarrow \alpha \in L(T)$

$\therefore \alpha \in L(S) \Rightarrow \alpha \in L(T)$

$\therefore L(S) \subseteq L(T)$

ii) Let $\alpha \in L(S \cup T)$

$\Rightarrow \alpha = a_1\alpha_1 + a_2\alpha_2 + ... + a_n\alpha_n + b_1\beta_1 + b_2\beta_2 + ... + b_m\beta_m$

where $a_1, a_2, ..., a_n, b_1, b_2, ..., b_m \in F$ and α_i's$\in S$, β_j's$\in T$.

$\Rightarrow \alpha = L.C$ of elements of $S + L.C$ of elements of T.

$\Rightarrow \alpha =$ An element of L(S) + an element of L(T)

$\Rightarrow \alpha \in L(S) + L(T)$ .... (1)

Now suppose $\alpha \in L(S) + L(T)$

$\Rightarrow \alpha = \beta + \gamma$ where $\beta \in L(S)$ and $\gamma \in L(T)$

$\beta \in L(S) \Rightarrow \beta = L.C$ of finite elements of S.

$\gamma \in L(T) \Rightarrow \gamma = L.C$ of finite elements of T.

$\therefore \alpha = L.C$ of finite elements of $S \cup T$.

$\therefore \alpha \in L(S \cup T)$

$\therefore L(S)+L(T)\subseteq L(S\cup T)$ .... (2)

From (1) and (2) $L(S\cup T)=L(S)+L(T)$

5.6.11 Theorem: If W_1 and W_2 are two subspaces of a vectorspace $V(F)$ then $L(W_1\cup W_2)=W_1+W_2$.

Proof: Let $V(F)$ be a vector space and W_1,W_2 be two subspaces of V.

We know that W_1+W_2 is a subspace of V containing $W_1\cup W_2$ by theorem 5.6.2.

Also $L(W_1\cup W_2)$ is the smallest subspace of V containing $W_1\cup W_2$.

$\therefore L(W_1\cup W_2)\subseteq W_1+W_2$ .... (1)

Let $\alpha\in W_1+W_2$

$\Rightarrow \alpha=\beta+\gamma$ where $\beta\in W_1$ and $\gamma\in W_2$

i.e. α is L.C. of finite elements of $W_1\cup W_2$

$\therefore \alpha\in L(W_1\cup W_2)$

$\therefore W_1+W_2\subseteq L(W_1\cup W_2)$ .... (2)

From (1) and (2) $L(W_1\cup W_2)=W_1+W_2$.

5.7 Linearly dependent and independent vectors :

5.7.1 Definition: Linearly dependent vectors: A finite subset $\{\alpha_1,\alpha_2......\alpha_n\}$ of vectors of a vector space $V(F)$ is said to be linearly dependent if there exists scalars $a_1,a_2......a_n\in F$ not all zero such that $a_1\alpha_1+a_2\alpha_2+....+a_n\alpha_n=\overline{0}$.

5.7.2 Linearly Independent Vectors: A finite subset $\{\alpha_1,\alpha_2......\alpha_n\}$ of vectors of a vector space $V(F)$ is said to be linearly independent if it is not linearly dependent.

5.7.3 Note: A finite subset $\{\alpha_1,\alpha_2......\alpha_n\}$ of vectors of a vector space $V(F)$ is linearly independent if every relation of the form $a_1\alpha_1+a_2\alpha_2+....+a_n\alpha_n=\overline{0}$ where a_i's$\in F\Rightarrow a_1=a_2=.....=a_n=0$.

5.7.4 Note: A set of vectors which contains the zero vector is linearly dependent (L.D.)

Solution: Let $\{\alpha_1, \alpha_2 ......\alpha_n\}$ be a finite set of vectors of a vector space $V(F)$.

Suppose $\alpha_1 = \overline{0}$ and $\alpha_1, \alpha_1 ......\alpha_n$ are non zero

Then $1\alpha_1 + 0\alpha_2 + 0\alpha_3 + ... + 0\alpha_n = \overline{0}$

$\therefore$ We have scalars $a_1, a_2, ..., a_n$ not all zero such that $a_1\alpha_1 + a_2\alpha_2 + ... + a_n\alpha_n = \overline{0}$ {Since $1 \neq 0\}$

$\therefore \{\alpha_1, \alpha_2 ......\alpha_n\}$ is linearly dependent.

i.e. a set of vectors containing atleast one zero vector is L.D.

5.7.5 Note: A single non-zero vector forms a L.I. set.

Solution: Suppose $\{\alpha_1\}$ where $\alpha_1 \neq \overline{0}$ is a subset of a vector space $V(F)$.

If $a\alpha_1 = \overline{0}$ where $a \in F$ then $a = 0$ since $\alpha_1 \neq \overline{0}$

$\therefore a\alpha_1 = \overline{0} \Rightarrow a = 0$

$\therefore$ A single non-zero vector forms a L.I. Set.

5.7.6 Theorem: Every super set of a linearly dependent set is linearly dependent.

Proof: Let $V(F)$ be a vector space and $S = \{\alpha_1, \alpha_2 ......\alpha_n\}$ be a linearly dependent set.

$\therefore$ There exists scalars $a_1, a_2, ..., a_n \in F$ not all zero such that $a_1\alpha_1 + a_2\alpha_2 + ... + a_n\alpha_n = \overline{0}$ .....(1)

Let $S^1 = \{\alpha_1, \alpha_2 ......\alpha_n, \beta_1, \beta_1 ......\beta_m\}$ be a superset of S (i.e. $S \subseteq S^1$)

Now $a_1\alpha_1 + a_2\alpha_2 + ... + a_n\alpha_n + 0\beta_1 + 0\beta_2 + ... + 0\beta_m = \overline{0}$ .... (2) from (1)

In the above relation (2) all scalars are not zero.

$\therefore S^1$ is L.D.

5.7.7 Theorem: Every non-empty subset of a linearly independent set is linearly independent.

Proof: Let $V(F)$ be a vector space and $S = \{\alpha_1, \alpha_2 ......\alpha_n\}$ be a linearly independent set.

Let $S^1 = \{\alpha_1, \alpha_2 ......\alpha_k\} 1 \leq k \leq n$ be a subset of S.

Suppose $a_1\alpha_1 + a_2\alpha_2 + ... + a_k\alpha_k = \overline{0}$ where $a_1, a_2 .... a_k \in F$

$\Rightarrow a_1\alpha_1 + a_2\alpha_2 + ... + a_k\alpha_k + 0\alpha_{k+1} + 0\alpha_{k+2} + ... + 0\alpha_n = \overline{0}$

$\Rightarrow a_1 = a_2 = ... = a_k = 0$ since S is L.I.

$\therefore S^1 = \{\alpha_1, \alpha_2 ......\alpha_k\}$ is L.I.

5.7.8 Theorem: Let $V(F)$ be a vector space. A finite subset of non-zero vectors $S = \{\alpha_1, \alpha_2 ......\alpha_n\}$ of $V(F)$ is linearly dependent iff some vector $\alpha_k, 2 \le k \le n$ can be expressed as linear combination of the vectors which preced it.

Proof: Let $V(F)$ be a vector space and $S = \{\alpha_1, \alpha_2 ......\alpha_n\}$ be a finite subset of non-zero vectors of $V(F)$.

Suppose S is linearly dependent.

$\Rightarrow$ there exists scalars $a_1, a_2 ....a_k \in F$ not all zero such that $a_1\alpha_1 + a_2\alpha_2 + ... + a_n\alpha_n = \overline{0}$ .... (1)

Suppose k is the greatest suffix such that $a_k \ne 0$.

If $k = 1$ then $a_1\alpha_1 = \overline{0}$

$\Rightarrow \alpha_1 = \overline{0}$ since $a_k \ne 0$

A contradiction since S is a subset of nonzero vectors.

$\therefore 2 \le k \le n$

From (1) $a_1\alpha_1 + a_2\alpha_2 + ... + a_n\alpha_n = \overline{0}$

$\Rightarrow a_1\alpha_1 + a_2\alpha_2 + ... + a_k\alpha_k = \overline{0}$ since $a_i = 0$ for $i > k$

$\Rightarrow a_k\alpha_k = -a_1\alpha_1 - a_2\alpha_2 - .... - a_{k-1}\alpha_{k-1}$

$\Rightarrow \alpha_k = -(a_k^{-1}a_1)\alpha_1 - (a_k^{-1}a_2)\alpha_2 .... - (a_k^{-1}a_{k-1})\alpha_{k-1}$

$\therefore \alpha_k$ is L.C of its preceding vectors.

Now suppose some vector $\alpha_k, 2 \le k \le n$ is a L.C of its preceding vectors.

$\therefore \alpha_k = b_1\alpha_1 + b_2\alpha_2 + ... + b_{k-1}\alpha_{k-1}$ where $b_1, b_2, ..., b_{k-1} \in F$

$\Rightarrow b_1\alpha_1 + b_2\alpha_2 + ... + b_{k-1}\alpha_{k-1} + (-1)\alpha_k = \overline{0}$

$\Rightarrow b_1\alpha_1 + b_2\alpha_2 + ... + b_{k-1}\alpha_{k-1} + (-1)\alpha_k + 0\alpha_{k-1} + ... + 0\alpha_n = \overline{0}$

Since -1 is a non zero scalar we have $S = \{\alpha_1, \alpha_2 ......\alpha_n\}$ is L.D.

5.7.9 Theorem: Let $V(F)$ be a vector space and $S = \{\alpha_1, \alpha_2 ......\alpha_n\}$ be a subset of V. If $\alpha_i \in S$ is a linear combination of its preceding vectors then $L(S) = L(S^1)$ where $S^1 = \{\alpha_1, \alpha_2 .....\alpha_{i-1}, \alpha_{i-2}, ...\alpha_n\}$.

Proof: Let $V(F)$ be a vector space and $S = \{\alpha_1, \alpha_2 ......\alpha_n\}$ be a subset of V.

suppose $\alpha_i \in S$ is a linear combination of its preceding vectors.

i.e. $\alpha_i = a_1\alpha_1 + a_2\alpha_2 + a_{i-1}\alpha_{i-1}$ where $a_1, a_\alpha ....a_k \in F$

To prove that $L(S) = L(S^1)$

Clearly $S^1 \subseteq S \Rightarrow L(S^1) \subseteq L(S)$ .... (1)

Let $\alpha \in L(S)$

$\Rightarrow \alpha = b_1\alpha_i + b_2\alpha_2 + ... + b_{i-1}\alpha_{i-1} + b_i\alpha_i + b_{i-1}\alpha_{i-1} + ... + b_n\alpha_n$ where $b_1, b_2, ..., b_n \in F$

But $\alpha_i = a_1\alpha_1 + a_2\alpha_2 + ... + a_{i-1}\alpha_{i-1}$

$\therefore \alpha = b_1\alpha_1 + b_2\alpha_2 + ... + b_{i-1}\alpha_{i-1} + b_i(a\alpha_1 + a_2\alpha_2 + ... + a_{i-1}\alpha_{i-1}) + b_{i+1}\alpha_{i+1} + ... + b_n\alpha_n$

$= (b_1 + b_i a_1)\alpha_1 + (b_2 + b_i a_2)\alpha_2 + ... + (b_{i-1} + b_i a_{i-1})\alpha_{i-1} + b_{i+1}\alpha_{i+1} + a_2\alpha_2 + ... + a_{i-1}\alpha_{i-1}) + b_{i+1}\alpha_{i+1} + ... + b_n\alpha_n$

= L.C. of elements of S^1.

$\therefore \alpha \in L(S^1)$

$\alpha \in L(S) \Rightarrow \alpha \in L(S^1)$

$\therefore L(S) \subseteq L(S^1)$ .... (2)

Hence $L(S) = L(S^1)$ from (1) and (2).

5.8 Summary:

In this lesson we learnt definitions of vector space, subspace, linear sum, linear span, linearly dependent and linearly independent vectors. We proved theorems relating to Algebra of subspaces. Linear sum of subspaces, linear span of a set. We discussed the concepts linear combination, linear dependence and independance of vectors.

5.9. Technical Terms:

i) Internal composition

ii) External composition

iii) Vector space

iv) Vector subspace

v) Linear sum of subspaces

vi) Linear combination

vii) Linear span

viii) Linearly dependent vectors

ix) Linearly independent vectors.

5.10 Model Examination Questions:

1. The set of all m x n matrices with real entries is a vector space over the real field with respect to addition of matrices and scalar multiplication of a matrix.

Solution: Let M = Set of all m x n matrices with elements as real numbers.

i) We know that addition of two m x n matrices is again an m x n matrix.

$$\therefore A, B \in M \Rightarrow A+B \in M$$

M is closed w.r. to addition +.

ii) We know that addition of matrices is associative.

$$\therefore (A+B)+C = A+(B+C), \forall A, B, C \in M$$

Clearly the null matric $O_{m\times n} \in M$ and $A+O=O+A=A \forall A \in M$.

$\therefore$ The null matrix O is the additive identity.

iv) $A \in M \Rightarrow -A \in M$ and $A+(-A)=(-A)+A=0$

$\therefore -A$ is the additive inverse of A.

v) Addition of matrices is commutative.

$$\therefore A+B=B+A, \forall A, B \in M$$

$\therefore (M,+)$ is an abelian group.

vi) Let $a \in R$ and $A \in M$

$\Rightarrow aA \in M$, where $a(r_{ij}) = (ar_{ij})$ if $A = (r_{i\gamma})$.

$\therefore M$ is closed w.r.to scalar multiplication.

vii) $a(A+B) = aA + aB$ for all $a \in R$ and $A, B \in M$

viii) $(a+b)A = aA + bA$ for all $a, b \in R$ and $A \in M$

ix) $a(bA) = (ab)A$ for all $a, b \in R$ and $A \in M$

x) $1A = A$.

$\therefore M(R)$ is a vector space.

2. The set of all real valued continuous functions defined in the open internal (0,1) is a vector space over the field of real numbers with respect to addition and scalar multiplication defined by $(f+g)(x) = f(x) + g(x)$ and $(af)(x) = af(x)$ where $a \in R$ and $0 < x < 1$.

Solution: Let $S = \{ f / f : (0,1) \to \mathbb{R} \}$ is continuous}

i). We know that sum of two continuous functions is continuous

$\therefore f, g \in S \Rightarrow f + g \in S$

$\therefore S$ is closed w.r.to addition of functions.

ii) Let $f, g, h \in S$

$$[(f+g)+h](x) = (f+g)(x) + h(x) = [f(x) + g(x)] + h(x)$$

$$= f(x) + [g(x) + h(x)]$$

$$= f(x) + (g+h)(x)$$

$$= [f + (g+h)](x)$$

$\therefore$ Addition of functions is associative.

iii) $0 : (0,1) \to R$ defined by $0(x) = 0 \not\forall x \in (0,1)$ is a constanst function and is continuous.

$\therefore 0 \in S$

$$(f+0)(x) = f(x) + 0(x) = f(x) + 0 = f(x)$$

$\therefore 0$ is the additive identity.

iv) $f:(0,1) \to R \Rightarrow -f:(0,1) \to R$

f is continuous $\Rightarrow -f$ is continuous.

$$\Rightarrow -f \in S$$

$$f + (-f) = 0$$

$\therefore -f$ is the additive inverse of f.

v) For $f, g \in S$ we have $(f+g)(x) = f(x) + g(x)$

$$= g(x) + f(x)$$

$$= (g+f)(x)$$

$\therefore$ Addition is commutative.

Hence $(S,+)$ is an abelian group.

vi) $f \in S$ and $a \in R \Rightarrow af \in S$

$\therefore S$ is closed under scalar multiplication.

vii) $f, g \in S$ and $a \in R \Rightarrow$

$$[a(f+g)](x) = a(f+g)(x) = a[f(x) + g(x)]$$

$$= af(x) + ag(x)$$

$$= (af)(x) + (ag)(x)$$

$$= (af + ag)(x)$$

$$\therefore a(f+g) = af + ag$$

viii) $[(a+b)f](x) = (a+b)f(x) = af(x) + bf(x)$

$$= (af)(x) + (bf)(x)$$

$$= (af + bf)(x)$$

$\therefore (a+b)f = af + bf \ \forall a, b \in R$ and $f \in S$

For $1 \in R$, $[1f](x) = 1f(x) = f(x)$

$\therefore 1f = f$

$[(ab)f](x) = (ab)f(x) = a[bf(x)] = a[(bf)(x)]$

$= [a(bf)(x)]$

$\therefore (ab)f = a(bf)$

$\therefore S$ is a vector space.

3. Let $V_3(F)$ be a vector space. Then $W = \{(x, y, 0)/x, y \in F\}$ is a subspace of $V_3(F)$.

Solution: Let F be a field and $V_3(F) = \{(a,b,c)/a,b,c \in F\}$ the vector space of ordered triads.

$W = \{(x, y, 0)/x, y \in F\}$

Clearly W is a non-empty subset of $V_3(F)$.

Let $\alpha, \beta \in W$ and $a, b \in F$

$\alpha, \beta \in W \Rightarrow \alpha = (x_1, y_1, 0)$ and $\beta = (x_2, y_2, 0)$ where $x_1, y_1, x_2, y_2 \in F$

$a\alpha + b\beta = a(x_1, y_1, 0) + b(x_2, y_2, 0)$

$= (ax_1, ay_1, 0) + (bx_2, by_2, 0)$

$= (ax_1 + bx_2, ay_1 + by_2, 0 + 0)$

$= (ax_1 + bx_2, ay_1 + by_2, 0)$

$\therefore a\alpha + b\beta \in W$ since $ax_1 + bx_2, ay_1 + by_2 \in F$

$\therefore W$ is a subspace of $V_3(F)$.

4. Let $W = \{(x, 3x, 3x+2)/x \in \mathbb{R}\}$ show that W is not a subspace of $V_3(\mathbb{R})$.

Solution : Clearly $\alpha = (x, 3x, 3x+2), \beta = (y, 3y, 3y+2) \in W$

But $\alpha + \beta = (x+y, 3(x+y), 3(x+y)+4) \notin W$

$\therefore W$ is not a subspace of $V_3(\mathbb{R})$.

5. Express the vector $\alpha=(1,-2,5)$ as a linear combination of the vectors $e_1=(1,1,1), e_\alpha=(1,2,3), e_3=(2,-1,1)$

Solution: Suppose $\alpha = ae_1+be_2+ce_3$ where $a,b,c\in\mathbb{R}$

Then $(1,-2,5)=a(1,1,1)+b(1,2,3)+c(2,-1,1)$

$$=(a,a,a)+(b,2b,3b)+(2c,-c,c)$$

$$=(a+b+2c, a+2b-c, a+3b+c)$$

$\therefore a+b-2c=1$ .... (1)

$a+2b-c=-2$ .... (2)

$a+3b+c=5$ .... (3)

Solving (1), (2) and (3)

(1) + 2 x (2) $a+b+2c=1$

$2a+4b-2c=-4$

$3a+5b=-3$ .... (4)

(2) + (3) $2a+5b=3$ .... (5)

(4) - (5) $a=-6$

From (5) $b=3$

From (1) $c=2$

$\therefore (1,-2,5)=-6(1,1,1)+3(1,2,3)+2(2,-1,1)$

6. Show that the vector $(1,2,3),(1,0,0),(0,1,0)$ and $(0,0,1)$ from a linearly dependent subset of $V_3(\mathbb{R})$.

Solution: Suppose $\therefore (1,-2,5)=a(1,2,3)+b(1,0,0)+c(0,1,0)+d(0,0,1)=\overline{0}$

$$\Rightarrow (a,2a,3a)+(b,0,0)+(0,c,0)+(0,0,d)=(0,0,0)$$

$$\Rightarrow (a+b, 2a+c, 3a+d)=(0,0,0)$$

$\Rightarrow a+b=0\rightarrow(1)$ $\quad 2a+c=0\rightarrow(2)$ $\quad 3a+d=0\rightarrow(3)$

$\Rightarrow b=-a$ $\quad c=-2a$ $\quad d=-3a$

If $a = K$ then $K(1,2,3) - K(1,0,0) - 2K(0,1,0) - 3K(0,0,1) = \overline{0}$

$\Rightarrow (1,2,3) - (1,0,0) - 2(0,1,0) - 3(0,0,1) = \overline{0}$

$\therefore$ There exist scalars not all zero such that the linear combination of the give n vectors is zero. Hence the given vectors are linearly dependent.

7. Show that the vectors $(1,0,-1)(1,2,1)(0,-3,2)$ are linearly independent in $V_3(\mathbb{R})$.

Solution: Suppose $a(1,0,-1) + b(1,2,1) + c(0,-3,2) = \overline{0}$

$\Rightarrow (a,0,-a) + (b,2b,b) + (0,-3c,2c) = (0,0,0)$

$\Rightarrow (a+b, 2b-3c, -a+b+2c) = (0,0,0)$

$\Rightarrow a+b=0 \rightarrow (1) \quad 2b-3c=0 \rightarrow (2) \quad -a+b+2c=0$

(1)+(3) gives $2b + 2c = 0 \rightarrow (3)$

(2) - (3) given $-5c = 0 \Rightarrow c = 0$

From (3) $b = 0$

From (1) $a = 0$

$\therefore a(1,0,-1) + b(1,2,1) + c(0,-3,2) = \overline{0}$

$a = b = c = 0$

$\therefore$ The given vectors are linearly independent.

8. If α, β, γ are linearly independent vectors in $V(\mathbb{R})$ show that $\alpha+\beta, \beta+\gamma, \gamma+\alpha$ are also linearly independent.

Solution: Suppose $a(\alpha+\beta) + b(\beta+\gamma) + c(\gamma+\alpha) = \overline{0}$

$\Rightarrow a\alpha + a\beta + b\beta + b\gamma + c\gamma + c\alpha = \overline{0}$

$\Rightarrow (a+c)\alpha + (a+b)\beta + (b+c)\gamma = \overline{0}$

α, β, γ are linearly independent.

$\Rightarrow a+c = a+b = b+c = 0$

Solving $a = b = c = 0$

$\therefore \alpha+\beta, \beta+\gamma, \gamma+\alpha$ are also linearly independent.

5.11 Exercises:

(1) Prove that the set of all polynomials in an indeterminate x over a field F is a vector space.

(2) Show that $\{(a,b,c)/a,b,c \in \mathbb{R}$ and $a+b+2c=0\}$ is a subspace of $V_3(\mathbb{R})$.

(3) Show that $\{(a,b,c)/a,b,c \in Q\}$ is not a subspace of $V_3(\mathbb{R})$.

(4) Write the vector $A=\begin{bmatrix} 3 & -1 \\ 1 & -2 \end{bmatrix}$ in the vector space of all 2 x 2 matrices as the linear combination of the vectors.

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

(5) Show that the subspace spanned by $S=\{\alpha,\beta\}$ and $T=\{\alpha,\beta,\gamma\}$ are the same in the vector space $V_3(\mathbb{R})$ if $\alpha=(1,2,1)$, $\beta=(3,1,5)$ and $\gamma=(3,-4,7)$

(6) Show that $(1,2,1)(3,1,5)$ in $\mathbb{R}^3$ are linearly independent.

(7) Show that $(1,3,2)(1,-7,-8)(2,1,-1)$ in $\mathbb{R}^3$ are linearly dependent.

(8) Prove that the set $\{1,x,x(1-x)\}$ is linearly independent set of vectors in the space of all polynomials over the real field.

- Smt. K. Ruth

LESSON - 6

BASES AND DIMENSION

6.1 Objective of the Lesson:

To learn the definition of vector space. subspace, some examples of vector spaces, algebra of subspaces and related theorems.

6.2 Structure

6.3 Introduction:

In this lesson we define basis and dimension of a vector space, Quotient space. We prove some important theorems relating to dimension.

6.4.1 Definition: Basis of a vector space : A subset S of a vector space $V(F)$ is said to be a basis of V, if

i) S is linearly independent

ii) S spans V i.e. $L(S) = V$

6.4.2 Example: The set $S = \{e_1, e_2 ... e_n\}$ where $e_1 = (1,0,0,...0)$, $e_2 = (0,1,0,...0), e_3 = (0,01,0...0) e_n = (0,0...0,1)$ is a basis of the vector space $V_n(F)$.

Solution: $V_n(F) = \{(a_1, a_2 ... a_n) / a_1, a_2 ... a_n \in F\}$ where F is a field.

First we show that S is L.I.

Suppose $b_1 e_1 + b_2 e_2 + ... + b_n e_n = \overline{0}$

$\Rightarrow b_1(1,0,...0) + b_2(0,1,0...0) + b_n(0,0...0,1) = \overline{0}$

$\Rightarrow (b_1, 0, ...0) + (0, b_2, 0, ...0) + ... + (0, 0...0, b_n) = \overline{0}$

$\Rightarrow (b_1, b_2...b_n) = \overline{0} = (0, 0...0, b_n) = \overline{0}$

$\Rightarrow b_1 = b_2 = ...b_n = 0$

$\therefore S$ is L.I.

Let $\alpha \in V_n(F)$

$\Rightarrow \alpha = (a_1, a_2...a_n)$

Now $\alpha = a_1 e_1 + a_2 e_2 + ... + a_n e_n$

$\therefore S$ spans V i.e. $L(S) = V_n(F)$

Hence S is a basis of $V_n(F)$

6.4.3 Note: The basis $S = \{e_1, e_2...e_n\}$ is called standard basis of $V_n(F)$.

6.4.4 Definition: Finite Dimensional Vector Space: A vector space V(F) is said to be finite dimensional vector space if there is a finite subset S of V which spans V. i.e. $L(S) = V$.

6.4.5 Existance of basis theorem:

Every finite dimensional vector space has a basis.

Proof: Let $V(F)$ be a finite dimensional vector space.

$\Rightarrow$ There is a finite subset S of V such that $L(S) = V$.

Let $S = \{\alpha_1, \alpha_2...\alpha_n\}$

We may assume that $\overline{0} \notin S$.

If S is linearly independent then S itself is a basis of V.

If S is linearly independent then there exists a vector α_k, $2 \le K \le n$ which is a linear combination of its preceding vectors. by 5.7.8.

Take $S_1 = \{\alpha_1, \alpha_2...\alpha_{k-1}, \alpha_{k+1}...\alpha_n\}$

Now $S_1 \subseteq S$ and $L(S_1) = L(S)$ by 5.7.9.

$\Rightarrow L(S_1) = V$ since $L(S) = V$

If S_1 is linearly independent then S_1 is a bais of V.

Suppose S_2 is linearly dependent.

Proceeding as above we get a subset S_3 of S with $n-2$ elements and $L(S_3)=V$.

Continuing the above process, after finite no. of steps we get a subset which is linearly independent and spans V.

$\therefore$ We get a basis of V.

Hence every finite dimentional vector space has a basis.

6.4.6 Invariance Theorem: Let $V(F)$ be a finite dimensional vector space, then any two basis will have the same number of elements.

Proof: Let $V(F)$ be a finite dimensional vector space.

Then $V(F)$ has a basis by 6.4.5.

Let $S_1=\{\alpha_1,\alpha_2...\alpha_m\}$ and $S_2=\{\beta_1,\beta_2...\beta_n\}$ be two bases of $V(F)$.

$\Rightarrow S_1$ is $L.I, L(S_1)=V$ and S_2 is $L.I, L(S_2)=V$

$\beta_1 \in V \Rightarrow \beta_1$ is linear combination of S_1.

$\therefore S_3=\{\beta_1,\alpha_1,\alpha_2...\alpha_m\}$ is linearly dependent.

Also $L(S_3)=V$

$\therefore$ There exists a vector $\alpha_i \in S_3$ which is a linear combination of the preceding vectors and $\alpha_i \neq \beta_1$ by 5.7.8.

Let $S_4=\{\beta_1,\alpha_1,\alpha_2...\alpha_{i-1},\alpha_{i+1},\alpha_{i+2}...\alpha_m\}$

Now $L(S_4)=V$ by 5.7.9.

$\therefore \beta_2$ is a linear combination of elements of S_4.

$\Rightarrow S_5=\{\beta_1,\beta_2\alpha_1,\alpha_2...\alpha_{i-1},\alpha_{i+1},...\alpha_n\}$ is L.D.

$\Rightarrow$ There exists a vector $\alpha_j \in S_5$ which is a linear combination of the preceding vectors and $\alpha_j \neq \beta_1, \alpha_j \neq \beta_2$.

Now $S_6 = \{\beta_1, \beta_2, \alpha_1, \alpha_2 ... \alpha_{i+1}, ... \alpha_{j-1}, \alpha_{j+1} ... \alpha_n\}$ generates V i.e. $V(S_6) = V$

If we continue this process at each step one α is excluded and a β is included in the set S_1 obviously the set S_1 of α 's can not be exhausted before the set S_2 of βs otherwise $V(F)$ will be a linear span of a proper subset of S_2 and thus S_2 will become linearly dependent.

$\therefore$ We must have $n \not\leq m$

Now interchanging the roles of S_1 and S_2 we shall get that $m = n$.

$\therefore$ Any two bases of a finite dimentional vector space have the same number of elements.

6.5.1 Definition: The number of elements in any basis of a finite dimensional vector space $V(F)$ is called the dimension of the vector space $V(F)$ and will be denoted by dim V.

6.5.2 Example: The dimension of the vector space $V_n(F)$ is n.

6.5.3 Theorem: Every linearly independent subset of a finite dimensional vector space $V(F)$ is either a basis of V or can be extended to form a basis of V.

Proof: Let $V(F)$ be a finite dimensional vector space and $S = \{\alpha_1, \alpha_2 ... \alpha_m\}$ be a linearly independent subset of V.

Suppose dim $V = n$

$\Rightarrow V$ has a basis $\beta = \{\beta_1, \beta_2 ... \beta_m\}$

consider $S_1 = \{\alpha_1, \alpha_2 ... \alpha_m, \beta_1, \beta_2 ... \beta_n\}$

Since B is basis of V, each α can be expressed as linear combination of β 's.

$\therefore S_1$ is linearly dependent.

$\Rightarrow$ One of the vectors of S_1 is a linear combination of its preceding vectors.

This vector can not be any of the α 's, since α 's are linearly independent.

$\therefore$ That vector is one of β 's let it be β_i.

Let $S_2 = \{\alpha_1, \alpha_2 ... \alpha_m, \beta_1, \beta_2 ... \beta_{i-1}, \beta_{i+1}, ... \beta_n\}$

Clearly $L(S_2) = V$.

If S_2 is linearly independent then S_2 will be a basis of V which is the extended set of S.

If S_2 is linearly dependent we repeat the above process a finite no. of times and we get a linearly independent set containing S and spanning V.

This set is a basis of V which is the extention of S.

$\therefore$ Every linearly independent subset of finite dimensional vector space is either a basis or can be extended to form a basis of V.

6.5.4 Corollary: Each subset of $(n+1)$ or more vectors of an n-dimensional vector space $V(F)$ is linearly dependent.

Proof: Let S be a subset of an n-dimensional vector space $V(F)$.

Suppose S has $(n+1)$ or more vectors.

If S is linearly independent then either S is a basis of V or can be extendend to form a basis of V by the theorem 6.5.3.

Thus a basis of V contain $(n+1)$ or more vectors.

Which is a conradiction to the fact that dim $V = n$.

$\therefore S$ is linearly dependent.

6.5.5 Corollary: Any set of n linearly independent vectors of a n-dimensional vector space $V(F)$ forms a basis of V.

Proof: Let $V(F)$ be a n-dimensional vector space and S be a linearly independent subset of V with n vectors.

$\therefore S$ is a basis of V or can be extended to form a basis of V.

Since dim $V = n$, S must be a basis of V.

6.5.6 Corollary: Every set of n vectors of a n-dimensional vector space $V(F)$, which generates V is a basis of V.

Proof: Let $V(F)$ be a n-dimensional vector space and S be a set of n vectors which generates V. If S is linearly dependent then we get a proper subset of S which is a basis of V.

$\therefore$ We get a basis of V with less than n vectors. It is a contradiction to the fact that dim $V = n$.

$\therefore S$ is linearly independent.

Hence S is a basis of V.

6.5.7 Theorem: If $S = \{\alpha_1, \alpha_2 ... \alpha_n\}$ is a basis of a finite dimensional vector space $V(F)$ of dimension n then every element $\alpha \in V$ can be uniquely expressed as $\alpha = a_1\alpha_1 + a_2\alpha_2 + ... + a_n\alpha_n$ where $a_1, a_2 ... a_n \in F$.

Proof: Let $V(F)$ be a finite dimensional vector space with dim $V = n$.

$S = \{\alpha_1, \alpha_2 ... \alpha_n\}$ is a basis of $V \Rightarrow L(S) = V$

$\alpha \in V \Rightarrow \alpha = a_1\alpha_1 + a_2\alpha_2 + ... + a_n\alpha_n$ where $a_1, a_2 ... a_n \in F$

Suppose $\alpha = b_1\alpha_1 + b_2\alpha_2 + ... + b_n\alpha_n$ where $b_1, b_2 ... b_n \in F$

$\therefore a_1\alpha_1 + a_2\alpha_2 + ... + a_n\alpha_n = b_1\alpha_1 + b_2\alpha_2 + ... + b_n\alpha_n$

$\Rightarrow (a_1 - b_1)\alpha_1 + (a_2 - b_2)\alpha_2 + ... + (a_n - b_n)\alpha_n = \overline{0}$

$\Rightarrow a_1 - b_1 = a_2 - b_2 = ... = a_n - b_n = 0$ since S is L.I.

$\Rightarrow a_1 = b_1, a_2 = b_2 .... a_n = b_n$

Hence every vector $\alpha \in V$ can be uniquely expressed as $\alpha = a_1\alpha_1 + a_2\alpha_2 + ... + a_n\alpha_n$ where $a_1, a_2 ... a_n \in F$.

6.5.8 Note: If $B = \{\alpha_1, \alpha_2 ... \alpha_n\}$ is a basis of a finite dimensional vector space $V(F)$ then every vector $\alpha \in V$ is uniquely expressed as $\alpha = a_1\alpha_1 + a_2\alpha_2 + ... + a_n\alpha_n$ where $a_1, a_2 ... a_n \in F$. The scalars $a_1, a_2 ... a_n$ are called coordinates of α, relative to the basis B.

6.5.9 Theorem: If W is a subspace of a finite dimensional vector space $V(F)$ then W is also finite dimensional and $\dim W \le \dim V$. Also $V = W$ iff $\dim V = \dim W$.

Proof: Let $V(F)$ be a finite dimensional vector space with dim $V = n$ and W be a subspace of V.

$\dim V = n \Rightarrow$ Every subset of $(n+1)$ or more vectors of W is L.D.

W is a subspace of $V \Rightarrow$ every subset of $(n+1)$ or more vectors of W is L.D.

$\therefore$ Any set of L.I. vectors in W must contain at the most n vectors.

Suppose $S = \{\alpha_1, \alpha_2 ... \alpha_m\}$ be the largest L.I. subset of W.

Where $m \le n$

Now we prove that S is a basis of W.

Let $\alpha \in W$

Now $S_1 = \{\alpha_1, \alpha_2 ... \alpha_m, \alpha\}$ is L.D. {Since S is the largest L.I. subset of W}

$\Rightarrow$ There exists scalars $a_1, a_2 ... a_m \in F$. not all zero $\ni a_1\alpha_1 + a_2\alpha_2 + ... + a_m\alpha_m + a\alpha = \overline{0}$

If $a = 0$ then $a_1\alpha_1 + a_2\alpha_2 + ... + a_m\alpha_m = \overline{0}$

$\Rightarrow a_1 = a_2 = ... = a_m = 0$ since S is L.I.

$\Rightarrow S_1$ is L.I.

A contradiction.

$\therefore a \neq 0 \Rightarrow \exists \ a^{-1} \in F$ such that $aa^{-1} = 1$

$\therefore a_1\alpha_1 + a_2\alpha_2 + ... + a_m\alpha_m + a\alpha = \overline{0} \Rightarrow$

$\alpha = -a^{-1}a_1\alpha_1 - a^{-1}a_2\alpha_2 .... - a^{-1}a_m\alpha_m$

$\therefore \alpha$ is L.C. of elements of $S \Rightarrow \alpha \in L(S)$

$\therefore W = L(S)$

Hence S is a basis of $W \Rightarrow \dim W = m \leq n$

$\therefore W$ is also a finite dimensional vector space with $\dim W \leq \dim V$.

Now suppose $V = W$

Then every basis of V is a basis of W.

$\therefore \dim V = \dim W$

Conversely suppose $\dim V = \dim W = n$

Let S be a basis of $W \Rightarrow S$ contains n vectors and $L(S) = W$

$S \subseteq W \Rightarrow S \subseteq V$

Now S is a linearly independent subset of a finite dimensional vector space $V(F)$ with $\dim V = n$

$\therefore S$ is a basis of V by 6.5.5.

$\Rightarrow L(S) = V$

$\therefore V = W$

6.5.10 Theorem: If W_1 and W_2 are two subspace of a finite dimensional vector space $V(F)$ then $\dim(W_1 + W_2) \leq \dim W_1 + \dim W_2 - \dim(W_1 \cap W_2)$

Proof: Let W_1 and W_2 be two subspaces of a finite dimensional vector space $V(F)$.

$\therefore W_1 + W_2$ and $W_1 \cap W_2$ are also subspaces of $V(F)$ and hence they are finite dimensional by 6.5.9.

Let dim $(W_1 \cap W_2) = k$ and $S = \{\gamma_1, \gamma_2 ... \gamma_k\}$ be a basis of $W_1 \cap W_2 \Rightarrow S$ is L.I. and $W_1 \cap W_2 = L(S) \to (1)$.

Now $S \subseteq W_1$ and $S \subseteq W_2$

Since S is a linearly independent subset of W_1.

S can be extended to form a basis of W_1 by 6.5.3.

Let $S_1 = \{\gamma_1, \gamma_2 ... \gamma_k, \alpha_1, \alpha_2 ... \alpha_m\}$ be a basis of W_1

$\Rightarrow \dim W_1 = k + m$ and $W_1 = L(S_1), S_1$ is L.I. .... (2)

S is also a linearly independent subset of W_2.

$\therefore S$ can be extended to form a basis of W_2 by 6.5.3.

Let $S_2 = \{\gamma_1, \gamma_2 ... \gamma_k, \beta_1, \beta_2 ... \beta_n\}$ be a basis of W_2

$\Rightarrow \dim W_2 = k + n$ and $W_2 = L(S_2)$, S_2 is L.I. ..... (3)

Now we claim that $B = \{\alpha_1, \alpha_2, ... \alpha_m, \beta_1, \beta_2 ... \beta_n, \gamma_1, \gamma_2 ... \gamma_k\}$ is a basis of $W_1 + W_2$.

To Prove that B is L.I.:

Suppose $a_1\alpha_1 + a_2\alpha_2 + ... + a_m\alpha_m + b_1\beta_1 + b_2\beta_2 + ... + b_n\beta_n + c_1\gamma_1 + c_2\gamma_2 + ... + c_k\gamma_k = \overline{0}$ for

$a_1, a_2, ..., a_m, b_1, b_2, ..., b_n, c_1, c_2, ...c_k \in F$.

$\Rightarrow a_1\alpha_1 + a_2\alpha_2 + ... + a_m\alpha_m = -(b_1\beta_1 + b_2\beta_2 + ... + b_n\beta_n + c_1\gamma_1 + c_2\gamma_2 + ... + c_k\gamma_k) \to (4)$

Now $\Rightarrow b_1\beta_1 + b_2\beta_2 + ... + b_n\beta_n + c_1\gamma_1 + c_2\gamma_2 + ... + c_k\gamma_k \in W_2$ and $a_1\alpha_1 + a_2\alpha_2 + ... + a_m\alpha_m \in W_1$

$\therefore\ a_1\alpha_1 + a_2\alpha_2 + ... + a_m\alpha_m \in W_1 \cap W_2 = L(S)$ by (1)

$\Rightarrow a_1\alpha_1 + a_2\alpha_2 + ... + a_m\alpha_m = d_1\gamma_1 + d_2\gamma_2 + ... + d_k\gamma_k$ where d_i's $\in F$

$\Rightarrow\ a_1\alpha_1 + a_2\alpha_2 + ... + a_m\alpha_m - d_1\gamma_1 - d_2\gamma_2 - .... - d_k\gamma_k = \overline{0}$

$\Rightarrow a_1 = a_2 = ... = a_m = d_1 = d_2 = ... = d_n = 0$ since S_1 is L.I.

$\Rightarrow\ a_1\alpha_1 + a_2\alpha_2 + ... + a_m\alpha_m = \overline{0}$

$\Rightarrow\ b_1\beta_1 + b_2\beta_2 + ... + b_n\beta_n + c_1\gamma_1 + c_2\gamma_2 + ... + c_k\gamma_k = \overline{0}$ from (4)

$\Rightarrow b_1 = b_2 = ... = b_n = c_1 = c_2 = ... = c_k = 0$ since S_2 is L.I.

$\therefore a_1 = a_2 = ... = a_m = b_1 = b_2 = ... = b_n = c_1 = c_2 = ... = c_k = 0$

$\therefore B$ is linearly independent.

<u>B Spans</u> $(W_1 + W_2)$

Let $\alpha \in W_1 + W_2$

$\Rightarrow \alpha = \beta + \gamma$ where $\beta \in W$ and $\gamma \in W_2$

$\beta \in W_1 \Rightarrow \beta = a_1\alpha_1 + a_2\alpha_2 + ... + a_m\alpha_m + c_1\gamma_1 + c_2\gamma_2 + ... + c_k\gamma_k$

where $a_1, a_2, ..., a_m, c_1, c_2, ..., c_k \in F$. Since $W_1 = L(S_1)$

$\gamma \in W_2 \Rightarrow \gamma = b_1\beta_1 + b_2\beta_2 + ... + b_n\beta_n + d_1\gamma_1 + d_2\gamma_2 + ... + d_k\gamma_k$

where $b_1, b_2, ..., b_m, d_1, d_2, ..., d_k \in F$ Since $W_2 = L(S_2)$.

Now $\alpha = a_1\alpha_1 + a_2\alpha_2 + ... + a_m\alpha_m + c_1\gamma_1 + c_2\gamma_2 + ... + c_k\gamma_k + b_1\beta_1 + b_2\beta_2 + ... + b_m\beta_m + d_1\gamma_1 + d_2\gamma_2 + ... + d_k\gamma_k$

= L.C of elements of B.

$\therefore \alpha \in L(B) \Rightarrow B$ spans $W_1 + W_2$

Hence B is a basis of $W_1 + W_2$.

$\therefore \dim(W_1 + W_2) = m + n + k$

$$= k + m + k + n - k$$

$$= \dim W_1 + \dim W_2 - \dim(W_1 \cap W_2)$$

6.6.1 Coset: Let W be a subspace of a vector space $V(F)$ and $\alpha \in V$. Then the set $\{\gamma + \alpha : \gamma \in W\}$ is called right coset of W in V generated by α. The set $\{\alpha + \gamma : \gamma \in W\}$ is called left coset of W in V generated by α.

6.6.2 Note: 1. $\{\gamma + \alpha : \gamma \in W\}$ is denoted by $W + \alpha$ and $\{\alpha + \gamma : \gamma \in W\}$ is denoted by $\alpha + W$.

2. The right coset $W + \alpha$ are subsets of V.

3. $W + \alpha = \alpha + W$ since $(V, +)$ is an abelian group.

$\therefore$ We say $W + \alpha$ is a coset of W generated by α.

6.6.3 Theorem: If W is a subspace of $V(F)$ then

i) $W + \overline{0} = W$

ii) $\alpha \in W \Leftrightarrow W + \alpha = W$

iii) $W + \alpha = W + \beta \Leftrightarrow \alpha - \beta \in W$

Proof: $\overline{0} \in V \therefore W + \overline{0} = \{\alpha + \overline{0} / \alpha \in W\}$

$$= \{\alpha / \alpha \in W\}$$

$$= W$$

ii) Suppose $\alpha \in W$

T.P.T. $W + \alpha \leq W$

Let $\beta \in W + \alpha \Rightarrow \beta = \gamma + \alpha, \gamma \in W$

Now $\gamma + W, \alpha \in W$ and W is a subspace

$\Rightarrow \gamma + \alpha \in W$

$\Rightarrow \beta \in W$

$\therefore W + \alpha \subseteq W$ .... (1)

T.P.T $W \subseteq W + \alpha$

Let $\beta \in W$

$\alpha \in W$ and W is a subspace $-\alpha \in W$

$\beta \in W, -\alpha \in W \Rightarrow \beta - \alpha \in W$, since W is a subspace.

$\therefore \beta = (\beta - \alpha) + \alpha \in W + \alpha$

$\therefore W \subseteq W + \alpha$ .... (2)

Hence $W + \alpha = W$ from (1) and (2)

iii) Suppose $W + \alpha = W + \beta$

Clearly $\overline{0} \in W$ since W is a subspace.

$\therefore \overline{0} + \alpha \in W + \alpha \Rightarrow \alpha \in W + \alpha$

$\Rightarrow \alpha \in W + \beta$

$\Rightarrow \alpha = \gamma + \beta$ when $\gamma \in W$

$\Rightarrow \alpha - \beta = \gamma \in W$

$\therefore W + \alpha = W + \beta \Rightarrow \alpha - \beta \in W$

Conversely suppose $\alpha - \beta \in W$

$\Rightarrow W + (\alpha - \beta) = W$ from ii.

$\Rightarrow W + \alpha - \beta + \beta = W + \beta$

$\Rightarrow W + \alpha = W + \beta$

6.6.4 Theorem: If W is a subspace of a vector space $V(F)$ then the set V/W of all cosets of W in V is a vector space over F w.r. to addition of cosets and scalar multiplication defined by $(W + \alpha) + (W + \beta) = W + (\alpha + \beta)$ and $a(W + \alpha) = W + a\alpha$, for all $\alpha, \beta \in V$ and $a \in F$.

Proof: Let W be a subspace of a vector space $V(F)$ and $V/W = \{W + \alpha / \alpha \in V\}$ be the set of all cosets of W in V.

Addition is well defined:

Suppose $W + \alpha = W + \gamma$ and $W + \beta = W + \delta$

$\Rightarrow \alpha - \gamma \in W$ and $\beta - \delta \in W$ by 6.6.3

$\Rightarrow \alpha - \gamma + \beta - \delta \in W$ since W is a subspace

$\Rightarrow (\alpha + \beta) - (\gamma + \delta) \in W$

$\Rightarrow W + (\alpha + \beta) = W + (\gamma + \delta)$ by 6.6.3.

Scalar multiplication is well defined:

Suppose $W + \alpha = W + \gamma$

$\Rightarrow \alpha - \gamma \in W$ by 6.6.3

$\Rightarrow a(\alpha - \gamma) \in W$ for $a \in F$

$\Rightarrow a\alpha - a\gamma \in W$

$\Rightarrow W + a\alpha = W + a\gamma$ by 6.6.3

$\Rightarrow a(W + \alpha) = a(W + \gamma)$

$\therefore$ Addition and scalar multiplication in V/W are well defined.

Addition is associative:

Let $W + \alpha, W + \beta, W + \gamma \in V/W$

$$\left[(W + \alpha)(W + \beta)\right] + (W + \gamma) = \left[W + (\alpha + \beta)\right] + (W + \gamma)$$

$$= W\left[(\alpha + \beta) + \gamma\right]$$

$$= W + \left[\alpha + (\beta + \gamma)\right]$$

$$= (W + \alpha) + \left[W + (\beta + \gamma)\right]$$

$$= (W + \alpha) + \left[(W + \beta) + (W + \gamma)\right]$$

Existance of Identity:

Clearly $W + \overline{0} \in V/W$ and $(W + \alpha) + (W + \overline{0}) = W + (\alpha + \overline{0}) = W + \alpha$

$(W + \overline{0}) + (W + \alpha) = W + (\overline{0} + \alpha) = W + \alpha, \forall W + \alpha \in V/W$

$\therefore W + \overline{0}$ is the additive identity of V/W.

Existance of Inverse:

Let $W+\alpha \in V/W \Rightarrow -\alpha \in V \Rightarrow W+(-\alpha) \in V/W$

Now $(W+\alpha)+[(W+(-\alpha)] = W+[\alpha+(-\alpha)] = W+\overline{0} = W$ and

$$[(W+(-\alpha)]+(W+\alpha) = W+(-\alpha+\alpha) = W+\overline{0} = W$$

$\therefore W+(-\alpha)$ is the additive inverse of $W+\alpha$.

Addition is commutative:

Let $W+\alpha, W+\beta \in V/W$

$$(W+\alpha)+(W+\beta) = W+(\alpha+\beta) = W+(\beta+\alpha)$$

$$= (W+\beta)+(W+\alpha)$$

$\therefore V/W$ is an abelian group w.r. to addition.

Let $W+\alpha, W+\beta \in V/W$ and $a, b \in F$

i) $a[(W+\alpha)+(W+\beta)] = a[W+(\alpha+\beta)] = W+(a\alpha+a\beta)$

$$= (W+a\alpha)+(W+a\beta) = a(W+\alpha)+a(W+\beta)$$

ii) $(a+b)(W+\alpha) = [W+(a+b)\alpha] = W+(a\alpha+b\alpha) = (W+a\alpha)+(W+b\alpha)$

$$= a(W+\alpha)+b(W+\alpha)$$

iii) $ab(W+\alpha) = W+(ab)\alpha = W+a(b\alpha) = a(W+b\alpha) = a[b(W+\alpha)]$

iv) $1(W+\alpha) = W+1\alpha = W+\alpha$

$\therefore V/W$ is a vector space.

6.6.5 Quotient Space: If W is a subspace of a vector space $V(F)$. Then V/W the set of all cosets of W in V is a vector space called quotient space.

6.6.6. Theorem: If W is a subspace of a finite dimensional vector space V(F) then $\dim\left(V/W\right) = \dim V$ - dim W.

Proof: Let W be a subspace of a finite dimensional vector space V(F).

Suppose $\dim W = m$ and $S = \{\alpha_1, \alpha_2 ... \alpha_m\}$ be a basis of W.

$\therefore S$ is linearly independent and $L(S) = W$ .... (1)

Now S is a linearly independent subset of $V\{\therefore W \le V\}$

$\therefore S$ can be extended to form a basis of V.

Let $S_1 = \{\alpha_1, \alpha_2 ... \alpha_m, \beta_1, \beta_2, \beta_k\}$ be a basis of V.

$\therefore \dim V = m + k, S_1$ is $L.I$ and $L(S_1) = V$ .... (2)

Now we claim that $B = \{W + \beta_1, W + \beta_2, ... W + \beta_k\}$ is the basis of V/W.

<u>B is Linearly Independent:</u>

Let $b_1(W + \beta_1) + b_2(W + \beta_2) + ... + b_k(W + \beta_k) = W$ where $b_1, b_2 ... b_k \in F$

$\Rightarrow (W + b_1\beta_1) + (W + b_2\beta_2) + ... + (W + b_k\beta_k) = W$

$\Rightarrow W + (b_1\beta_1 + b_2\beta_2 ... + b_k\beta_k) = W$

$\Rightarrow b_1\beta_1 + b_2\beta_2 + ... + b_k\beta_k \in W$

$\Rightarrow b_1\beta_1 + b_2\beta_2 + ... + b_k\beta_k = a_1\alpha_1 + a_2\alpha_2 + ... + a_m\alpha_m$ since L(S)=W by (1)

$\Rightarrow b_1\beta_1 + b_2\beta_2 ... + b_k\beta_k - a_1\alpha_1 - a_2\alpha_2 - ... - a_m\alpha_m = \overline{0}$

$\Rightarrow b_1 = b_2 ... = b_k = a_1 = a_2 = ... a_m = 0$ since S_1 is L.I.

$\therefore B$ is linearly independent.

<u>B Spans V/W</u>:

Let $W + \alpha \in V/W$

$\Rightarrow \alpha \in V$

$\Rightarrow \alpha = c_1\alpha_1 + c_2\alpha_2 + ... + c_m\alpha_m + d_1\beta_1 + d_2\beta_2 + ... + d_k\beta_k$ where $c_1, c_2 ... c_m$ and $d_1, d_2 ... d_k \in F$

$\therefore W + \alpha = W + (c_1\alpha_1 + c_2\alpha_2 + ... + c_m\alpha_m + d_1\beta_1 + d_2\beta_2 + ... + d_k\beta_k)$

$= W + (W + d_1\beta_1 + d_2\beta_2 + ... + d_k\beta_k)$ when $W = c_1\alpha_1 + ... + c_m\alpha_m$

$= (W + \gamma) + (d_1\beta_1 + d_2\beta_2 + ... + d_k\beta_k)$

$= W + d_1\beta_1 + d_2\beta_2 + ... + d_k\beta_k$ since $W + \gamma = W$

$= d_1(W + \beta_1) + d_2(W + \beta_2) + ... + d_k(W + \beta_k)$

$\therefore B$ spans V/W.

Hence B is a basis of V/W.

$\therefore \dim\left(V/W\right) = K = m + k - m = \dim V - \dim W$.

6.7 Summary:

In this lesson we learnt the definitions of basis of a vector space, dimension of a vector space and quotient space. We proved theorems relating to basis and dimension.

6.8 Technical Terms:

i) Basis

ii) Finite dimensional vector space

iii) Dimension of a vector space

iv) Quotient space

6.9 Model Examination Questions:

1. Show that vectors $(1,2,1)(2,1,0)(1,-1,2)$ form a basis for R^3.

Solution: Suppose $a(1,2,1) + b(2,1,0) + c(1,-1,2) = (0,0,0)$ where $a,b,c \in R$

$\Rightarrow (a + 2b + c, 2a + b - c, a + 2c) = (0,0,0)$

$\Rightarrow a + 2b + c = 0; 2a + b - c = 0; a + 2c = 0$

$a + 2c = 0 \Rightarrow a = 2c$

$\therefore a + 2b + c = 0 \Rightarrow -2c = +2b + c = 0 \Rightarrow 2b - c = 0 \Rightarrow 2b = c$

$2a + b - c = 0 \Rightarrow -4c = +b - c = 0 \Rightarrow b - 5c = 0$

$\Rightarrow b - 10c = 0 \Rightarrow -7b = 0 \Rightarrow b = 0$

$\therefore 2b = c \Rightarrow c = 0$ and $\therefore 2b = c \Rightarrow c = 0$

$\therefore \{(1,2,1)(2,1,0)(1,-1,2)\}$ is a basis of $\mathbb{R}^3$

We know that dim $\mathbb{R}^3 = 3$.

$\therefore \{(1,2,1)(2,1,0)(1,-1,2)\}$ forms a basis of $\mathbb{R}^3$ by 6.5.5.

6.9.2 Find the coordinates of the vector $(2,1,-6)$ of $\mathbb{R}^3$ relative to the basis $\{(1,1,2)(3,-1,0)(2,0,-1)\}$

Solution:

Suppose $(2,1,-6) = a(1,1,2) + b(3,-1,0) + c(2,0,-1)$ where $a,b,c \in R$

$\Rightarrow (2,1,-6) = (a+3b+2c, a-b, 2a-c)$

$a+3b+2c = 2$ .... (1)

$a-b=1$ .... (2)

$2a-c=-6$ .... (3)

From (2) $b = a-1$ from (3) $c = 2a+6$

$\therefore$ From (1) $a+3(a-1)+2(2a+6)=2$

$a+3a-3+4a+12=2$

$8a=-7$

$a=-7/8$

$\therefore b = \frac{-7}{8} - 1 = \frac{-15}{8}$ and $c = \frac{-14}{8} + 6 = \frac{34}{8} = \frac{17}{4}$

$\therefore$ The coordinates of $(2,1,-6)$ are $\left(\frac{-7}{8}, \frac{-15}{8}, \frac{17}{4}\right)$

6.9.3 If V is the vector space of ordered pairs of complex numbers over the real field R, then show that the set $S = \{(1,0)(i,0)(0,1)(0,i)\}$ is a basis of V.

Solution: Suppose $a(1,0) + b(i,0) + c(0,1) + d(0,i) = (0,0)$ for $a,b,c,d \in R$

$\Rightarrow (a+ib, c+id) = (0,0)$

$\Rightarrow a+ib=0$ and $c+id=0$

$\Rightarrow a=0, b=0, c=0, d=0$

$\therefore$ The set $S=\{(1,0)(i,0)(0,1)(0,i)\}$ is $L.I$

Let $(a+ib, c+id)\in V$

Then $(a+ib, c+id)=a(1,0)+b(i,0)+c(0,1)+d(0,i)$

$\therefore S$ spans V

Hence S is a basis of V.

6.9.4 Let V be the vector space of 2 x 2 matrices over a field F. Show that V has dimension 4 by exhibiting a basis for V which has four elements.

Solution: Let $S=\{A,B,C,D\}$ where $A=\begin{bmatrix}1 & 0\\0 & 0\end{bmatrix}$ $B=\begin{bmatrix}0 & 1\\0 & 0\end{bmatrix}$ $C=\begin{bmatrix}0 & 0\\1 & 0\end{bmatrix}$ $D=\begin{bmatrix}0 & 0\\0 & 1\end{bmatrix}$

Clearly $S\subseteq V$

Suppose $aA+bB+cC+dD=0_{2\times 2}$ where $a,b,c,d\in F$

$$\Rightarrow a\begin{bmatrix}1 & 0\\0 & 0\end{bmatrix}+b\begin{bmatrix}0 & 1\\0 & 0\end{bmatrix}+c\begin{bmatrix}0 & 0\\1 & 0\end{bmatrix}+d\begin{bmatrix}0 & 0\\0 & 1\end{bmatrix}=\begin{bmatrix}0 & 0\\0 & 0\end{bmatrix}$$

$$\Rightarrow \begin{bmatrix}a & b\\c & d\end{bmatrix}=\begin{bmatrix}0 & 0\\0 & 0\end{bmatrix}$$

$$\Rightarrow a=b=c=d=0$$

$\therefore S$ is linearly independent.

Further if $\alpha=\begin{bmatrix}a & b\\c & d\end{bmatrix}\in V$ then

$$\alpha=aA+bB+cC+dD$$

$\therefore S$ spans V.

Hence S is a basis of V.

$$\therefore \dim V=4$$

6.9.5 Do the vectors $(1,1,0)(0,1,2)$ and $(0,0,1)$ form a basis of $V_3(\mathbb{R})$.

Solution: Suppose $a(1,1,0)+b(0,1,2)+c(0,0,1)=(0,0,0)$ when $a,b,c \in R$

$$\Rightarrow (a, a+b, 2a+c)=(0,0,0)$$

$$\Rightarrow a=0; a+b=0; 2b+c=0$$

$$\Rightarrow a=0; b=-a=0; c=-2b=0$$

$\therefore$ The vectors $(1,1,0)(0,1,2)$ and $(0,0,1)$ are linearly independent.

Since $\dim V_3(\mathbb{R})=3$ the set $\{(1,1,0)(0,1,2)(0,0,1)\}$ is a basis of $V_3(\mathbb{R})$ by 6.5.5.

6.9.6 Show that the set $\{(1,0,0)(1,1,0)(1,1,1)(0,1,0)\}$ spans the vector space $\mathbb{R}^3$ but not a basis.

Solution: Let $(a,b,c) \in \mathbb{R}^3$

Suppose $(a,b,c)=x(1,0,0)+y(1,1,0)+z(1,1,1)+t(0,1,0)$

$\Rightarrow (a,b,c)=(x+y+z, y+z+t, z)$

$\Rightarrow x+y+z=a; y+z+t=b, z=c$

$\Rightarrow x+y=a-c; y+t=b-c$

If $y=0$ then $x=a-c$ and $t=b-c$

$\therefore (a,b,c)=(a-c)(1,0,0)+c(1,1,1)+(b-c)(0,1,0)$

$\therefore S=\{(1,0,0)(1,1,0)(1,1,1)(0,1,0)\}$ spans $\mathbb{R}^3$

Since $\dim \mathbb{R}^3=3$, S can not be a basis.

6.9.7 If $S=\{\alpha,\beta,\gamma\}$ is a basis of $C^3(c)$ show that $S^1=\{\alpha+\beta, \beta+\gamma, \gamma+\alpha\}$ is also a basis of $C^3(c)$.

Solution: Suppose $a(\alpha+\beta)+b(\beta+\gamma)+c(\gamma+\alpha)=\overline{0}$ where $a,b,c \in c$

$\Rightarrow a\alpha+a\beta+b\beta+b\gamma)+c\gamma+c\alpha=\overline{0}$

$\Rightarrow (a+c)\alpha+(a+b)\beta+(b+c)\gamma=\overline{0}$

$\Rightarrow a+c=0; a+b=0; b+c=0$

$\Rightarrow a=0; b=0; c=0$

$\Rightarrow S^1=\{\alpha+\beta, \beta+\gamma, \gamma+\alpha\}$ is linearly independent $\dim C^3(c)=3$.

$\therefore S^1$ is a basis of $C^3(c)$

6.9.8 Extend the set of linearly independent vectors $\{(1,0,1,0),(0,-1,1,0)\}$ to the basis of V_4.

Solution: $S=\{(1,0,1,0),(0,-1,1,0)\}$ is a linearly independent subset of V_4.

Clearly $B=\{(1,0,0,0),(0,1,0,0)(0,0,1,0),(0,0,0,1)\}$ is a basis of V_4.

Let $S_1=\{(1,0,1,0),(0,-1,1,0)(1,0,0,0),(0,1,0,0)(0,0,1,0)(0,0,0,1)\}$.

$\therefore S_1$ spans V_4 since $B \subseteq S_1$ and $L(B)=V_4$.

But $\dim V_4=4$

$\therefore S_1$ cannot be a basis.

$\Rightarrow$ S_1 is linearly dependent.

$\Rightarrow (0,0,0,1)$ is a L.C. of its preceding vectors.

$\therefore S_2=\{(1,0,1,0),(0,-1,1,0)(1,0,0,0),(0,1,0,0)(0,0,1,0)\}$ spans V_4.

But S_2 cannot be a basis as dim $V_4=4$

$\Rightarrow S_2$ is linearly dependent.

$\Rightarrow (0,0,1,0)$ is L.C. of its preceding vectors.

$\Rightarrow S_3=\{(1,0,1,0),(0,-1,1,0)(1,0,0,0),(0,1,0,0)\}$ spans V_4.

Since $\dim V_4=4$ and S_3 has 4 elements S_3 must be a basis of V_4.

6.9.9 Find a basis and dimension for the subspace of $\mathbb{R}^3$ spanned by the vectors $(2,7,3)(1,-1,0)(1,2,1)$ and $(0,3,1)$.

Solution : Let W be the subspace of $\mathbb{R}^3$ spanned by $S=\{(2,7,3)(1,-1,0)(1,2,1)(0,3,1)\}$

W is subspace of $\mathbb{R}^3 \Rightarrow \dim W \le \dim \mathbb{R}^3 = 3$

$\therefore S$ can not be linearly independent since S has 4 vectors.

$\therefore S$ is $L.D \Rightarrow (2,7,3)$ is L.C. of the remaining vectors.

Now $S_1=\{(1,-1,0)(1,2,1)(0,3,1)\}$ spans W.

Suppose $a(1,-1,0)+b(1,2,1)+c(0,3,1)=(0,0,0)$

$(a+b,-a+2b+3c,b+c)=(0,0,0)$

$\Rightarrow a+b=0$ .... (1) $-a+2b+3c=0$ .... (2) $b+c=0$............ (3)

(1) $\Rightarrow a=-b$

$\therefore$ From (2) $b+2b+3c=0 \Rightarrow 3b+3c=0 \Rightarrow b+c=0$

$\therefore c=-b$

$\therefore a=1, b=-1, c=1$

$\therefore S_1$ is linearly dependent.

$\Rightarrow (0,3,1)$ is L.C. of remaining vectors and $S_2=\{(1,-1,0)(1,2,1)\}$

Suppose $a(1,-1,0)+b(1,2,1)=(0,0,0)$

$(a+b,-a+2b,b)=(0,0,0)$

$\Rightarrow a+b=0 \quad -a+2b=0 \quad b=0$

$\Rightarrow a=0, b=0$

$\therefore S_2$ is L.I. and $L(S_2)=W$

Hence S_2 is a basis of W.

$\therefore \dim W = 2$

If W_1 and W_2 are subspaces of $V_4(R)$ generated by the sets $\{(1,1,-1,2)(2,1,3,0)(3,2,2,2)\}$ and $\{(1,-1,0,1)(-1,1,0,-1)\}$ respectively then find dim (W_1+W_2).

Solution:

Let $S_1 = \{(1,1,-1,2)(2,1,3,0)(3,2,2,2)\}$ and $S_2 = \{(1,-1,0,1)(-1,1,0,-1)\}$

Given $L(S_1) = W_1$ and $L(S_2) = W_2$

We know that $L(W_1 \cup W_2) = W_1 + W_2$

suppose a $(1,1,-1,2) + b(2,1,3,0) + c(3,2,2,2) + d(1,-1,0,1) + e(-1,1,0,-1) = (0,0,0,0)$

$\Rightarrow a + 2b + 3c + d - e = 0$ ..... (1) $a + b + 2c - d + e = 0$ .... (2)

$-a + 3b + 2c + = 0$ .... (3) $2a + 2c + d - e = 0$ .... (4)

(2) + (4) $3a + b + 4c = 0$ .... (5)

(1) - (4) $-a + 2b + c = 0$ .... (6)

(3) $-a + 3b + 2c = 0$

(6)-(3) $-b - c = 0 \Rightarrow b = -c$

From (5) $a + c = 0$

From (2) $d = e$

$\therefore a = 1; b = 0; c = -1; d = 1; e = 1$

$\Rightarrow S_1 \cup S_2$ is $L.D$

$\therefore$ One vector is the L.C. of remaining vectors.

$\Rightarrow S^1 = \{(1,1,-1,2)(2,1,3,0)(1,-1,0,1)(-1,1,0,-1)\}$ spans $W_1 + W_2$

Suppose $a(1,1,-1,2) + b(2,1,3,0) + c(1,-1,0,1) + d(-1,1,0,-1) = (0,0,0,0)$

$\Rightarrow a + 2b + c - d = 0$ .... (1) $a + b - c + d = 0$ .... (2)

$-a + 3b = 0$ ....(3) $2a + c - d = 0$ ....(4)

From (3) $a = 3b$ From (1) $5b + c - d = 0$

From (4) $6b + c - d = 0$

$\Rightarrow b = 0$

$\therefore a = 0 \therefore b = 0$ and $c = d$

$\Rightarrow S^1$ is $L.D$

$\Rightarrow$ One vector is S^1 is L.C of the remaining vectors.

$\therefore S^{11} = \{(1,1,-1,2)(1,-1,0,1)(-1,1,0,-1)\}$ spans $W_1 + W_2$

Suppose $a(1,1,-1,2) + b(1,-1,0,1) + c(-1,1,0,-1) = (0,0,0,0)$

$\Rightarrow a + b - c = 0$ .... (1) $a - b + c = 0$ .... (2)

$-a = 0$ .... (3) $2a + b - c = 0$ .... (4)

$a = 0$ we get from (1) b = c

$\therefore a = 0; b = 1; c = 1$

$\therefore S^{11}$ is L.I. and spans $W_1 + W_2$

Hence S^{11} is a basis of $W_1 + W_2$

$\therefore \dim(W_1 + W_2) = 3$

6.10 Exercises:

1. Show that $B = \{(1,0,-1)(1,2,1)(0,-3,2)\}$ from a basis for $\mathbb{R}^3$.

2. Find the coordinates of $(2,3,4,-1)$ w.r. to the basis $\{(1,1,0,0)(0,1,10)(0,0,1,1)(1,0,0,0)\}$ of $V_4(\mathbb{R})$

3. Extend the set $\{(1,0,1)(0,-1,1)\}$ to form a basis for $\mathbb{R}^3$.

4. Find a basis for the subspace spanned by the vectors $(1,2,0)(-1,0,1)(0,2,1)$ in $V_3(\mathbb{R})$.

5. If V is the vector space generated by the polynomials $x^3 + 2x^2 - 2x + 1, x^3 + 3x^2 - x + 4, 2x^3 + x^2 - 7x - 7$ then find a basis of V and its dimension.

- Smt. K. Ruth

LESSON - 7

LINEAR TRANSFORMATION

7.1 Objective of the Lesson:

In Chapter 5 and 6, we discussed vector spaces and some of the related topics. Now it is natural to consider functions from a vector space into a vector space. Such a function with a condition is called a linear transformation or a homomorphism.

In this chapter, we discuss the properties of these linear transformation and related problems.

7.2 Structure of the Lesson:

This lesson has the following components.

7.3 Introduction:

In I B.Sc./B.A. homomorphisms from a group into a group are discussed. In chapter 2 of this book, homomorphisms from a ring/ field into a ring/ field are discussed. In this chapter, we discuss the homomorphisms (linear transformations) from a vector space into a vector space.

7.4 Linear Transformation:

7.4.1 Definition: Let $U(F)$ and $V(F)$ be vector spaces. A function $T: U(F) \to V(F)$ is called a linear transformation from U into V if $T(a\alpha + b\beta) = aT(\alpha) + bT(\beta)$ for all $\alpha, \beta \in U$ and $a, b \in F$.

7.4.2 Note: 1) It is called i) a monomorphism of T is one-one

ii) an epimorphism if T in onto

iii) an isomorphism if T is one-one and onto.

2. A linear transformation $T: U \to U$ is called a linear operator on U and it is called an automorphism if T is one-one and onto.

3. A linear transformation $T: U(F) \to F$ is called a linear functional on U.

4. Taking a=1, b=1 we get $T(\alpha+\beta) = T(\alpha) + T(\beta) \forall \alpha, \beta \in U$. Taking b = 0, we get $T(\alpha\beta) = aT(\alpha) \forall a \in F, \forall \alpha \in U$.

5. Throughout this lesson $\mathbb{R}$ denotes the field of real numbers.

Solved Problems:

7.4.3 Define: $T: \mathbb{R}^2 \to \mathbb{R}^2$ by $T(x, y, z) = (2x + y, x), \forall (x, y, z) \in \mathbb{R}^3$ show that T is a linear transformation.

Solution: Let $a, b \in F, \alpha = (\alpha_1, \alpha_2, \alpha_3), \beta = (\beta_1, \beta_2, \beta_3) \in \mathbb{R}^3$

So $a\alpha + b\beta = (a\alpha_1 + b\beta_1, a\alpha_2 + b\beta_2, a\alpha_3 + b\beta_3)$

Now $T(a\alpha + b\beta) = \left(2(a\alpha_1 + b\beta_1) + (a\alpha_2 + b\beta_2), a\alpha_1 + b\beta_1\right)$

= $a(2\alpha_1 + \alpha_2, \alpha_1) + b(2\beta_1 + \beta_2, \beta_1) = aT(\alpha) + bT(\beta)$ for all $a, b \in F$ and $\alpha, \beta \in \mathbb{R}^3$.

Hence T is a linear transformation.

7.4.4 Define: $T: \mathbb{R}^3 \to \mathbb{R}^2$ by $T(a_1, a_2, a_3) = (a_1 - a_2, a_1 - a_3) \ \forall \ (a_1, a_2, a_3) \in \mathbb{R}^3$. Show that T is a linear transformation.

Solution : Let $a, b \in F$ and $\alpha = (a_1, a_2, a_3)\beta = (b_1, b_2, b_3) \in \mathbb{R}^3$

So $T(a\alpha + b\beta) = T(aa_1 + bb_1, aa_2 + bb_2, aa_3 + bb_3)$

$= (aa_1 + bb_1 - aa_2 - bb_2, aa_1 + bb_1 - aa_3 - bb_3)$

$= a(a_1 - a_2, a_1 - a_3) + b(b_1 - b_2, b_1 - b_3)$

$= aT(\alpha) + bT(\beta) \forall \alpha, \beta \in t\mathbb{R}^3, \forall a, b \in \mathbb{R}$

Hence T is a linear transformation.

7.4.5 Show that the mapping $T:\mathbb{R}^3 \to \mathbb{R}^2$ defined by $T(a_1, a_2, a_3) = (|a_1|, 0) \forall (a_1, a_2, a_3) \in \mathbb{R}^3$ is not a linear transformation.

Solution: Let $a, b \in F, \alpha = (\alpha_1, \alpha_2, \alpha_3), \beta = (\beta_1, \beta_2, \beta_3) \in \mathbb{R}^3$

Now $T(a\alpha + b\beta) = T(a\alpha_1 + \alpha\beta_1, a\alpha_2 + \alpha\beta_2 a\alpha_3 + b\beta_3)$

$= \left(|a\alpha_1 + b\beta_1|, 0\right)$ ...... (1)

But $aT(\alpha) + bT(\beta) = aT(\alpha_1, \alpha_2, \alpha_3) + bT(\beta_1, \beta_2, \beta_3)$

$= a\left(|\alpha_1|, 0\right) + b\left(|\beta_1| + 0\right)$ .... (2)

From 1 and 2, it is clear that $T(a\alpha + b\beta) \neq aT(\alpha) + bT(\beta)$

Hence T is not a linear transformation.

7.4.6 SAQ: Show that the mapping $T:\mathbb{R}^2 \to \mathbb{R}^2$ defined by $T(a_1, a_2) = (2a_1 + a_2, a_1)$, for all $(a_1, a_2) \in \mathbb{R}^2$ is a linear transformation.

Properties of linear transformations:

7.4.7 Theorem: Let $T:U \to V$ be a linear transformation from the vector space $U(F)$ to the vector space $V(F)$. Then

a) $T\left(\overline{0}\right) = \overline{0}$ (b) $T\left(-\alpha\right) = -T(\alpha) \forall \alpha \in U$

(c) $T\left(\alpha - \beta\right) = T(\alpha) - T(\beta) \forall \alpha, \beta \in U$

(d) $T(a_1\alpha_1 + a_2\alpha_2 + ... + a_n\alpha_n) = a_1T(\alpha_1) + a_2T(d_2) + ... + a_nT(\alpha_n)$ for all $\alpha_1, \alpha_2 ... \alpha_n \in U$ and $\forall a_1, a_2, ..., a_n \in F$.

Proof: a) We know that $T(\alpha) = T(\alpha + \overline{0}) = T(\alpha) + T\left(\overline{0}\right) \Rightarrow T(\alpha) + \overline{0} = T(\alpha) = T(\alpha) + T\left(\overline{0}\right)$

$\Rightarrow \overline{0} = T\left(\overline{0}\right)$, by LCL in the group $(V, +)$.

Hence $T\left(\overline{0}\right) = \overline{0}\left(= \overline{0}_V\right)$

(b) Let $\alpha \in \cup \Rightarrow -\alpha \in \cup$ and $\alpha + (-\alpha) = \overline{0} = \left(= \overline{0}_U\right)$

Now $T\left[\alpha+(-\alpha)\right]=T\left(\overline{0}\right)$

$\Rightarrow\ T\left[\alpha+(-\alpha)\right]=T(0)$

$\Rightarrow T(\alpha)+T(-\alpha)=T(\alpha)+\left(-\left[T(\alpha)\right]\right)$

$\Rightarrow T(-\alpha)=-T(\alpha)$, by LCL in the group $(V,+)$.

This is true for all $\alpha\in U$

c) We have $T(\alpha-\beta)=T\left[\alpha+(-\beta)\right]=T(\alpha)+T(-\beta)=T(\alpha)-T(\beta),\forall\alpha,\beta\in U$

d) We use induction on n.

If $n=1$, the $T(a_1\alpha_1)=a_1T(\alpha_1)\forall a\in F$ and $\forall\alpha_1\in U$

Assume the result for $n=k.$ So we have for any $a_1,a_z,......,a_k\in F$ and

$\alpha_1,\alpha_z,...,\alpha_k\in U,T(a_1\alpha_1+a_2\alpha_2+..+a_k\alpha_k)=a_1T(\alpha_1)+a_2(\alpha_2)+..+a_kT(\alpha_k)$

Now let $a_1,a_z,...,a_k,a_{k+1},\in F$ and $\alpha_1,\alpha_2,...,\alpha_k,a_{k+1}\in U$

$T(a_1\alpha_1+a_2\alpha_2+..+a_k\alpha_k+a_{k+1}\alpha_{k+1})=T\left[(a_1\alpha_1+a_2\alpha_2+..+a_k\alpha_k)+a_{k+1}\alpha_{k+1}\right]$

$=T(a_1\alpha_1+a_2\alpha_2+..+a_k\alpha_k)+T(a_{k+1},\alpha_{k+1})$

$=a_1T(\alpha_1)+a_2T(\alpha_2)+...a_kT(\alpha_k)+a_{k+1}T(\alpha_{n+1})$

So, the result is true for $n=k+1$. Now the result follows from Mathematical induction.

7.4.8 Theorem: Let $U(F),V(F)$ be vector spaces. Let $T:U\rightarrow V$ be defined by $T(\alpha)=\overline{0}\forall\alpha\in U$. Then T is a linear transformation.

Proof: Let $a,b\in F,\alpha,\beta\in U\Rightarrow a\alpha+b\beta\in U$

$\Rightarrow T(a\alpha+b\beta)=\overline{0}$

$=a\overline{0}+b\overline{0}=aT(\alpha)+bT(\beta),\forall a,b\in F,\forall\alpha,\beta\in U$

This shows that T is a linear transformation.

Note: The linear transformation $T:U\rightarrow V$ defined by $T(\alpha)=\overline{0}\forall\alpha\in U$ is called the zero transformation and is denoted by O.

7.4.9 Theorem: Let $V(F)$ be a vector spaces over a field F.

Let $T:V \to V$ be defined by $T(\alpha)=\alpha \forall \alpha \in V$. Then T is a linear transformation (operator) on V.

Proof: Let $a,b \in F, \alpha, \beta \in V \Rightarrow a\alpha + b\beta \in V$ and

$$T(a\alpha + b\beta) = a\alpha + b\beta$$

$$= aT(\alpha) + bT(\beta) \forall a,b \in F \text{ and } \alpha, \beta \in V$$

So, T is a linear transformation.

Note : The linear transformation $T:V \to V$ defined by $T(\alpha)=\alpha \forall \alpha \in V$ is called the identity operator an V and is denoted by I.

7.4.10 Theorem: Let $U(F), V(F)$ be vector spaces over a field F and let $T:U \to V$ be a linear transformation. Then the mapping $-T:U \to V$ defined by $(-T)(\alpha) = -T(\alpha) \forall \alpha \in U$ is a linear transformation.

Proof: Let $a,b \in F$ and $\alpha, \beta \in U \Rightarrow a\alpha + b\beta \in U$ and

$$(-T)(a\alpha + b\beta) = -\left[T(a\alpha + b\beta)\right]$$

$$= -\left[aT(\alpha) + bT(\beta)\right]$$

$$= -a.T(\alpha) - bT(\beta)$$

$$= a(-T)(\alpha) + -b(-T)(\beta) \forall a,b \in F, \forall \alpha, \beta \in U$$

So, -T is a linear transformation.

Note: - T is called the negative of the linear transformation T.

7.4.11 Theorem: Let T_1, T_2 be linear transformation from a vector space $U(F)$ into a vector space $V(F)$. Then the mapping $T_1 + T_2 : U \to V$ defined by $(T_1 + T_2)(\alpha) = T_1(\alpha) + T_2(\alpha) \forall \alpha \in U$ is a linear transformation.

Proof: Let $a,b \in F, \alpha, \beta \in U$

Now $(T_1 + T_2)(a\alpha + b\beta) = T_1(a\alpha + b\beta) + T_2(a\alpha + b\beta)$

$$= aT_1(\alpha) + bT_1(\beta) + aT_2(\alpha) + bT_2(\beta)$$

$$= a\left[T_1(\alpha) + T_2(\alpha)\right] + b\left[T_1(\beta) + T_2(\beta)\right]$$

$= a(T_1 + T_2)(\alpha) + b(T_1 + T_2)(\beta), \forall a, b \in F$ and $\forall \alpha, \beta \in U$

So, $T_1 + T_2$ is a linear transformation.

7.4.12 Theorem: Let $T : U \to V$ be a linear transformation and $a \in F$. Then the mapping $(aT) : U \to V$ defined by $(aT)(\alpha) = aT(\alpha) \forall \alpha \in U$ is a linear transformation.

Proof: $\alpha, \beta \in U$ and $x, y \in F \Rightarrow x\alpha + y\beta \in U$

Now $(aT)(x\alpha + y\beta) = a\left[T(x\alpha + y\beta)\right]$

$$= a\left[xT(\alpha) + yT(\beta)\right]$$

$$= a\left[xT(\alpha)\right] + a\left[yT(\beta)\right]$$

$$= (ax)T(\alpha) + ayT(\beta)$$

$$= (xa)T(\alpha) + (ya)T(\beta)$$

$$= x\left[aT(\alpha)\right] + y\left[aT(\beta)\right]$$

$$= x\left[(aT)(\alpha)\right] + y\left[(aT)(\beta)\right]$$

This is true for all $x, y \in F$ and for all $\alpha, \beta \in U$

So, aT is a linear transformation.

7.4.13 Theorem: Let $U = F^{n\times 1}, V = F^{m\times 1}$ and $A \in F^{m\times n}$. Define $T : \cup \to \vee$ by $T(X) = A(X), \forall X \in U$. Then T is a linear transformation.

Proof: Let $a, b \in F, X, Y \in U$.

So, $T(aX + bY) = A(aX + bY) = A(aX) + A(bY)$

$$= a(AX) + b(AY) = aT(X) + bT(Y)$$

for all $a, b \in F$ and $X, Y \in U$

Hence T is a linear transformation.

7.4.14 SAQ: Define: $T : \mathbb{R}^2 \to \mathbb{R}^2$ by $T(a_1, a_2) = (a_1, -a_2) \forall (a_1, a_2) \in \mathbb{R}^2$. Show that T is a linear transformation. (This T is called the reflection in X-axis).

7.4.15 Define: $T:\mathbb{R}^2 \to \mathbb{R}^2$ by $T(a_1, a_2) = (a_1, 0) \forall (a_1, a_2) \in \mathbb{R}^2$. Show that T is a linear transformation. (This T is called the projection on the X-axis).

7.4.16 Theorem: $T:U \to V$ is a linear transformation iff

$T(\alpha+\beta) = T(\alpha)+T(\beta)$ and $T(a\alpha) = aT(\alpha) \forall \alpha, \beta \in U$ and for all $a \in F$ (U, V are vector spaces over a field F).

Proof: 1. Suppose $T:U \to V$ is a linear transformation.

$\Rightarrow T(a\alpha+b\beta) = aT(\alpha)+bT(\beta), \forall a,b \in F$ and $\forall \alpha, \beta \in U$ .... (1)

Taking $a=1, b=1$, we get $T(1\alpha+1\beta) = 1T(\alpha)+1T(\beta), \forall \alpha, \beta \in U$

$$\Rightarrow T(\alpha+\beta) = T(\alpha)+T(\beta), \forall \alpha, \beta \in U$$

Taking $a=1, b=0$, in (1) we get $T(a\alpha+o\beta) = aT(\alpha)+oT(\beta)$

$$\Rightarrow T(a\alpha) = aT(\alpha) \forall a \in F \omega \forall \alpha \in \cup$$

Conversely suppose that $T(\alpha+\beta) = T(\alpha)+T(\beta)$ and

$$T(a\alpha) = aT(\alpha) \forall \alpha, \beta \in U \forall a \in F$$

Now let $a,b \in F$ and $\alpha, \beta \in U$

So, $T(a\alpha+b\beta) = T(a\alpha)+T(b\beta)$

$$= aT(\alpha)+bT(\beta)$$

This is true for all $\alpha, \beta \in U$ and $a,b \in F$.

Hence T is a linear transformation.

In view of theorem (7.4.16), a linear transformation may also be defined as:

7.4.17 Definition: Let $U(F), V(F)$ be vector spaces. A mapping $T:U \to V$ is a linear transformation if $T(\alpha+\beta)=T(\alpha)T+T(\beta)$ and $T(a\alpha) = aT(\alpha) \forall a \in F$ and $\alpha, \beta \in U$.

The two conditions can also be replaced by the single condition:

$T(a\alpha+\beta) = aT(\alpha)+T(\beta) \forall \alpha, \beta \in U$ and $\forall a \in F$.

This can be justified on similar lines.

7.4.18 Theorem: Let $U(F), V(F)$ be vector spaces and let $\{\alpha_1, \alpha_2, ...\alpha_n\}$ be a basis of U. Let $\{\delta_1, \delta_2, ...\delta_n\}$ be any set of n vectors in V. Then $\exists$ a unique linear transformation $T:U \to V$

such that $T(\alpha_i) = \delta_i, 1 \le i \le n$.

Proof: Let $S = \{\alpha_1, \alpha_2, ...\alpha_n\}$ Let $\alpha \in U$. Since S is a basis of U, $\exists a_1, a_2, ..., a_n \in F \ni$
$\alpha = a_1\alpha_1 + a_2\alpha_2 + ... + a_n\alpha_n$

Define $T: U \to V$ by $T(\alpha) = T(a_1\alpha_1 + a_2\alpha_2 + ... + a_n\alpha_n) = a_1\delta_1 + a_2\delta_2 + ... + a_n\delta_n$, for every $\alpha = a_1\alpha_1 + a_2\alpha_2 + ... + a_n\alpha_n \in U$.

Let $a, b \in F, \alpha = a_1\alpha_1 + a_2\alpha_2 + ... + a_n\alpha_n \in U$

$$\beta = b_1\alpha_1 + b_2\alpha_2 + ... + b_n\alpha_n \in U$$

Now $T(a\alpha + b\beta) = T\left[(a_1a_1 + bb_1)\alpha_1 + (a_1a_2 + bb_2)\alpha_2 + ... + (aa_n + bb_n)\alpha_n\right]$

$$= (aa_1 + bb_1)\delta_1 + (aa_2 + bb_2)\delta_2 + ... + (aa_n + bb_n)\delta_n$$

$$= a(a_1\delta_1 + a_2\delta_2 + ... + a_n\delta_n) + b(b_1\delta_1 + b_2\delta_2 + ... + b_n\delta_n)$$

$= aT(\alpha) + aT(\beta)$, for all $\alpha, \beta \in U$ and for all $a, b \in F$.

This shows that T is a linear transformation.

Now for each $i, 1 \le i \le n, T(\alpha_i) = T\left[o\alpha_1 + o\alpha_2 + ... + 1\alpha_i + o\alpha_{i+1} ... + o\alpha_n\right]$

$$= o\delta_1 + o\delta_2 + ... + 1\delta_i + o\delta_{i+1} + ... + o\delta_n$$

$$= \delta_i$$

To show that T is unique, let $T^1: U \to V$ be a linear transformation such that $T^1(\alpha_i) = \delta_i, 1 \le i \le n.$ Let $\alpha \in U$. So $\exists\ a_1, a_2, .., a_n \in F \ni$

$$\alpha = a_1\alpha_1 + a_2\alpha_2 + ... + a_n\alpha_n$$

Now $T^1(\alpha) = T^1(a_1\alpha_1 + a_2\alpha_2 + ... + a_n\alpha_n)$

$$= a_1T^1(\alpha_1) + a_2T^1(\alpha_2) + ... + \delta_nT^1(\alpha_n)$$

$$= a_1\delta_1 + a_2\delta_2 + ... + a_n\delta_n = T(\alpha), \forall \alpha \in U$$

So $T^1 = T$. Hence the uniqueness and hence the theorem.

Product of linear transformations.

7.4.19 Theorem: Let U, V, W be vector spaces over a field F. Let $T: U \to V, H: V \to W$ be linear transformations. Then the composite function $HT: U \to W$ defined by $HT(\alpha) = H(T(\alpha)) \forall \alpha \in U$ is a linear transformation. (This HT is called product of linear transformations)

Proof: If $\alpha \in U$, then $T(\alpha) \in V$ and $H(T(\alpha)) \in W$ so that $HT(\alpha) \in W$

Let $a, b \in F$ and $\alpha, \beta \in U$

$$HT(a\alpha + b\beta) = H[T(a\alpha + b\beta)] = H[aT(\alpha) + bT(\beta)]$$

$$= aH(T(\alpha)) + bH(T(\beta))$$

$$= a(HT)(\alpha) + b(HT)(\beta) \forall \alpha, \beta \in U \text{ and } a, b \in F$$

Hence HT is a linear transformation.

7.4.20 Theorem: Let U, V, W be vector spaces over a field F.

Let $T_1, T_2 : U \to V; H_1, H_2 : V \to W$ be linear transformations. Then

i) $H_1(T_1 + T_2) = H_1T_1 + H_1T_2$

ii) $(H_1 + H_2)T_1 = H_1T_1 + H_2T_1$

iii) $a(H_1T_1) = (aH_1)T_1 = H_1(aT_1)$

Proof: i) Let $\alpha \in U$

$$[(H_1 + H_2)T_1](\alpha) = (H_1 + H_2)[T_1(\alpha)] = H_1[T_1(\alpha)] + H_2[T_1(\alpha)]$$

$$= (H_1T_1)(\alpha) + (H_2T_1)(\alpha) = (H_1T_1 + H_2T_1)(\alpha) \forall \alpha \in U$$

$$\Rightarrow (H_1 + H_2)T_1 = H_1T_1 + H_2T_1$$

(ii) and (iii) can be proved on similar lines.

7.4.21 Theorem: Let T_1, T_2, T_3 be linear operators on a vector space V(F).

Then (a) $T_1O = OT_1 = O$; (b) $T_1I = IT_1 = T_1$

(c) $T_1(T_2 + T_3) = T_1T_2 + T_1T_3$ (d) $(T_1 + T_2)T_3 = T_1T_3 + T_2T_3$

(e) $T_1(T_2T_3)=(T_1T_2)T_3$

Proof: Let $\alpha\in V; T_1o(\alpha)=T_1[o(\alpha)]=T_1(\bar{o})=(\bar{o})=\bar{o}\;\forall\alpha\in V$

$\Rightarrow T_1O=O$. Similarly we can prove that $OT_1=O$. [Here O is the zero operator on V].

The remaining results can be proved similarly.

7.4.22 Theorem: Let $L(U,V)$ be the set of all linear transformations from a vector space $U(F)$ into the vector space $V(F)$. For $T_1,T_2\in L(U,V)$ and $a\in F$, define $(T_1+T_2)(\alpha)=T_1(\alpha)+T_2(\alpha)\;\forall\alpha\in U$ and $(aT_1)(\alpha)=aT_1(\alpha)\;\forall\alpha\in U$. Then $L(U,V)$ is a vector space over F.

Proof: By theorems 7.4.11 and 7.4.12, we have $T_1+T_2, aT_1\in L(U,V)$. This shows that $L(U,V)$ is closed under vector addition and scalar multiplication.

Let $T_1,T_2,T_3\in L(U,V)$ and $\alpha\in U$.

$\therefore[T_1+(T_2+T_3)](\alpha)=T_1(\alpha)+(T_2+T_3)(\alpha)=T_1(\alpha)+[T_2(\alpha)+T_3(\alpha)]$

$=[T_1(\alpha)+T_2(\alpha)]+T_3(\alpha)=(T_1+T_2)(\alpha)+T_3(\alpha)$

$=[(T_1+T_2)+T_3](\alpha)\;\forall\alpha\in U$

$\therefore T_1+(T_2+T_3)=(T_1+T_2)+T_3\;\forall T,T_2,T_3\in L(U,V)$

+ is associative in $L(U,V)$.

Similarly we can prove that + is commutative in $L(U,V)$.

The zero transformation $O:U\to V$ is identity with respect to +. For $T\in L(U,V)$ and $\alpha\in U$, we have

$(T+O)(\alpha)=T(\alpha)+O(\alpha)=T(\alpha)+\bar{O}=T(\alpha)\;\forall\alpha\in U$

$\Rightarrow T+O=O$. Similarly we can show that $O+T=O$.

$\therefore T+O=O=O+T,\forall T\in L(U,V)$

$\therefore O$ is the identity.

$T\in L(U,V)\Rightarrow -T\in L(U,V)$ by theorem 7.4.10.

For $\alpha \in U, [T+(-T)](\alpha) = T(\alpha)+(-T)(\alpha) = T(\alpha)-T(\alpha) = \overline{O}(U) \forall \alpha \in U$

$\therefore T+(-T)=0$. Similarly we can show that $(-T)+T = O$

$\therefore -T$ is the inverse of T w.r.t. +.

For $a \in F, T_1, T_2 \in L(U,V), \alpha \in U$.

$$\therefore [a(T_1+T_2)](\alpha) = a[(T_1+T_2)(\alpha)] = a[(T_1(\alpha)+T_2(\alpha)] = aT_1(\alpha)+aT_2(\alpha)$$

$$= (aT_1)(\alpha)+(aT_2)(\alpha) = a(T_1+aT_2)(\alpha) \forall \alpha \in U$$

$\therefore a(T_1+T_2) = aT_1+aT_2 \forall a \in F$ and $\forall T_1, T_2 \in L(U,V)$.

Similarly we can prove that $(a+b)T = aT+bT, (ab)T = a(bT)$ and $1T = T \forall a,b \in F$ and $\forall T \in L(U,V)$

Hence $L(U,V)$ is a vector space over F.

Note: $L(U,V)$ is also denoted by Hom $F(U,V)$ or Hom (U,V).

7.4.23 Theorem: Let dim $U = n$, dim $V = m$. Then dim $L(U,V) = mn$ (U, V are vector spaces over the field F).

Proof: Let $A = \{\alpha_1, \alpha_2, ..., \alpha_n\}$, $B = \{\beta_1, \beta_2, ..., \beta_m\}$ be ordered bases of U and V respectively. By Theorem 7.4.18, $\exists$ a unique linear transformation $T_{ij} \in L(U,V) \ni$.

$$T_{ij}(\alpha_k) = \begin{cases} \beta_j, & k=i \\ \overline{0} & \text{if } k \neq i \end{cases}$$

for all $i, j, 1 \le i \le n, 1 \le j \le m$ and for each fixed $k, 1 \le k \le n$.

For example: $T_{i1}(\alpha_1) = \overline{0}$, $T_{i1}(\alpha_2) = \overline{0}$, $T_{i1}(\alpha_3) = \overline{0}$ ...., $T_{i1}(\alpha_i) = \beta_1$, ...$T_{i1}(\alpha_n) = \overline{0}$

i.e. $T_{ij}(\alpha_i) = \beta_j$ and $T_{i1}(\alpha_k) = \overline{0}$ if $k \neq i$.

Let $S = \{T_{ij} / 1 \le i \le n, 1 \le j \le m\}$

a) S is L.I: Let $a_{ij} \in F$ and suppose $\sum_{i=1}^{n}\sum_{j=1}^{m} a_{ij}T_{ij} = O(\in L(U,V)$

For each $k, 1 \le k \le n$, we have $\left(\sum_{i=1}^{n}\sum_{j=1}^{m} a_{ij}T_{ij}\right)(\alpha_k) = O(\alpha_k) = \bar{O}$

$$\Rightarrow \sum_{i=1}^{n}\sum_{j=1}^{m} a_{ij}T_{ij}(\alpha_k) = \bar{0} \Rightarrow \sum_{j=1}^{m} a_{kj}\beta_j = \bar{0} \Rightarrow a_{k1}\beta_1 + a_{k2}\beta_2 + ... + a_{km}\beta_m = \bar{O}$$

$$\Rightarrow a_{k1} = 0, a_{k2} = 0, ...a_{km} = 0 \;\forall\; k, 1 \le k \le n \Rightarrow a_{ij} = 0 \;\forall\; i, j, 1 \le i \le n, 1 \le j \le m$$

$\therefore S$ is L.I.

b) L(S) = L(U,V):

Clearly $L(S) \subseteq L(U,V)$

Let $T \in L(U,V)$. Now $\alpha_1 \in U \Rightarrow T(\alpha_1) \in V$. So $T(\alpha_1)$ can be expressed on a l.c. of elements of B. So $\exists$ $b_{11}, b_{12}, b_{13}, ...b_{1m} \in F \ni$ $T(\alpha_1) = b_{11}\beta_1 + b_{12}\beta_2 + b_{13}\beta_3 + ... + b_{1m}\beta_m$.

In general for each $k, 1 \le k \le n$, $\exists b_{k1}, b_{k2}, ...b_{km} \in F \ni T(\alpha_k) = b_{k1}\beta_1 + b_{k2}\beta_2 + ... + b_{km}\beta_m$

Let $H = \sum_{i=1}^{n}\sum_{j=1}^{m} b_{ij}T_{ij} \Rightarrow H \in L(U,V)$. We have $T_{ij}(\alpha_k) = \bar{0}$ if $k \ne i$ and β_j if $k = i$.

For each $k, 1 \le k \le n$, $H(\alpha_k) = \sum_{i=1}^{n}\sum_{j=1}^{m} b_{ij}T_{ij}(\alpha_k) = \sum_{j=1}^{m} b_{kj}\beta_j = T(\alpha_k)$.

$$\therefore H(\alpha_k) = T(\alpha_k), \forall \alpha_k \in B_1$$

So H and T agree on a basis of U. If $\alpha \in U$, then $\exists a_1, ...a_n \in F \ni$

$$\alpha = a_1\alpha_1 + a_2\alpha_2 + ...a_n\alpha_n.$$

So, $H(\alpha) = H(a_1\alpha_1 + a_2\alpha_2 + ... + a_n\alpha_n) = a_1H(\alpha_1) + a_2H(\alpha_2) + ... + a_nH(\alpha_n)$

$= a_1T(\alpha_1) + a_2T(\alpha_2) + ... + a_nT(\alpha_n) = T(a_1\alpha_1 + a_2\alpha_2 + ... + a_n\alpha_n = T(\alpha)$

for all $\alpha \in U$. Hence $T = H \in L(S) \therefore L(U,V) \subseteq L(S) \Rightarrow L(S) = L(U,V)$.

$\therefore$ S is a basis of L(U,V) $\therefore$ dimL(U,V)= number of elements of S = mn. Hence the theorem.

7.5 Null space and Range:

7.5.1 Definition: Let U(F) and V(F) be vector spaces and $T: U \to V$ be a linear transformation.

The null space N(T) of T is defined as

$$N(T)=\{\alpha \in U/T(\alpha)=\overline{0}\}$$

The range R (T) of T is defined as

$$R(T)=\{T(\alpha)/\alpha \in U\}$$

Null space is also called kernel.

Range is also called image.

7.5.2 Theorem: Let $T:U \to V$ be a linear transformation. Then N(T) is a subspace of U and R(T) is a subspace of V.

Proof: (1) $N(T)=\{\alpha \in U/T(\alpha)=\overline{0}\}$

We know that $\overline{0} \in U$ and $T(\overline{0})=\overline{0}. \Rightarrow \overline{0} \in N(T)$

$\therefore \phi \neq N(T) \subseteq U$

Let $a,b \in F, \alpha, \beta \in N(T)$

$\Rightarrow T(\alpha)=\overline{0}, T(\beta)=\overline{0}$

$$T(a\alpha+b\beta)=aT(\alpha)+bT(\beta)=a.\overline{0}+b.\overline{0}=\overline{0}$$

$\Rightarrow a\alpha+b\beta \in N(T) \quad \forall \alpha, \beta \in N(T)$ and $\forall a,b \in F$.

$\therefore$ N(T) is a subspace of U.

2. $R(T)=\{T(\alpha)/\alpha \in U\}$

$\overline{0} \in U \Rightarrow T(\overline{0}) \in R(T) \Rightarrow \phi \neq R(T) \subseteq V$

Let $\beta_1, \beta_2 \in R(T)$ and $a,b \in F$.

$\Rightarrow \exists \alpha_1, \alpha_2 \in U \ni T(\alpha_1)=\beta_1, T(\alpha_2)=\beta_2$

$$\therefore a\beta_1+b\beta_2=aT(\alpha_1)+bT(\alpha_2)$$

$$=T(a\alpha_1+b\alpha_2)$$

$\in R(T)$, since $a\alpha_1+b\alpha_2 \in U$.

$\therefore a\beta_1 + b\beta_2 \in R(T), \forall \beta_1, \beta_2 \in R(T)$ and $\forall a, b \in F$

$\therefore R(T)$ is a subspace of V.

7.5.3 Example: Let V(F) be a vector space and $I = V \rightarrow V$ be the identify operator. Then $N(I) = \{\bar{0}\}$, $R(I) = V$.

7.5.4 Example: Let $U(F), V(F)$ be vector spaces. $O : U \rightarrow V$ be the zero linear transformation. Then $N(O) = U$ and $R(O) = \{\bar{O}\}$.

7.5.5 Theorem: Let $T : U(F) \rightarrow V(F)$ be a linear transformation. If U is finite dimensional, then $R(T)$ is finite dimensional.

Proof: Let $S = \{\alpha_1, \alpha_2, ..., \alpha_n\}$ be a basis of U.

and $S^1 = \{T(\alpha_1), T(\alpha_2), ..., T(\alpha_n)\}$

Let $\beta \in R(T) \Rightarrow \exists \alpha \in U \ni T(\alpha) = \beta$

$$\Rightarrow \exists a_1, a_2, ..., a_n \in F \ni \alpha = a_1\alpha_1 + a_2\alpha_2 + ... + a_n\alpha_n$$

$$\Rightarrow \beta = T(\alpha) = T(a_1\alpha_1 + a_2\alpha_2 + ... + a_n\alpha_n)$$

$$\Rightarrow \beta = a_1T(\alpha_1) + a_2T(\alpha_2) + ... + a_nT(\alpha_n)$$

$$\Rightarrow \beta \in L(S^1)$$

$\therefore R(T) \subseteq L(S^1)$

Also $S^1 \subseteq R(T)$ ($\because$ each element of S^1 is in R(T))

$\Rightarrow L(S^1) \subseteq R(T)$, since R(T) is a subspace of V containing S^1 and $L(S^1)$ is the smallest subspace containing S^1.

Hence $L(S^1) = R(T)$.

$\therefore R(T)$ is finite dimensional.

7.5.6 Note: 1. $\dim U = n \Rightarrow \dim R(T) \leq n$

2. If S is a basis of U, then R(T) = Linear span of T(S).

i.e. if $S = \{\alpha_1, \alpha_2, ..., \alpha_n\}$, then R(T)=linear span of $S^1 = \{T(\alpha_1), T(\alpha_2), ..., T(\alpha_n)\}$

7.5.7 Example : Let $T:\mathbb{R}^3 \to \mathbb{R}^2$ be defined by $T(a_1, a_2, a_3) = (a_1 - a_2, 2a_3) \forall (a_1, a_2, a_3) \in \mathbb{R}^3$

Then $N(T) = \{(a_1, a_2, a_3) / T(a_1, a_2, a_3) = \overline{0}\}$

$= \{(a_1, a_2, a_3) / (a_1 - a_2, 2a_3) = (0,0)\}$

$= \{(a_1, a_2, a_3) / a_1 = a_2, a_3 = 0\}$

= $\{a, a, 0) / a \in \mathbb{R}\}$

and $R(T) = \{T(a_1, a_2, a_3) / (a_1, a_2, a_3) \in \mathbb{R}^3\}$

$= \{(a_1, a_2, 2a_3) / a_1, a_2, a_3 \in \mathbb{R}\}$

By theorem 7.4.28, R(T) is spanned by $\{T(1,0,0), T(0,1,0), T(0,0,1)\}$, where $\{e_1, e_2, e_3\}$ is the standard basis of $\mathbb{R}^3$, R(T) = Span of $\{(1,0), (-1,0), (0,2)\}$

= Span of $\{(1,0), (0,1)\} = \mathbb{R}^2$

Rank and Nullity:

7.5.8 Definition: Let $T = U(F) \to V(F)$ be a linear transformation.

The rank of T, denoted by $\rho(T)$ is defined as :

$\rho(T) = \dim R(T)$

The nullity of T, denoted by $\upsilon(T)$ is defined as:

$\upsilon(T) = \dim N(T)$.

7.5.9 Theorem: Let $T:U \to V$ be a linear transformation and U be finite dimensional. Then $\rho(T) + \upsilon(T) = \dim U$.

Proof: Let $\dim U = n$, Since N(T) is a subspace of U and U is finite dimensional, it follows that N(T) is finite dimensional. Let $\dim N(T)$ be = k and let $A = \{\alpha_1, \alpha_2, ..., \alpha_k\}$ be a basis of N(T). Since A is a L.I. subset of U, A can be extended to form a basis of U. Let $B = \{\alpha_1, \alpha_2, ..., \alpha_k, \alpha_{k+1}, ..., \alpha_n\}$ be a basis of U. Let $S=\{T(\alpha_{k+1}),T(\alpha_{k+2}),...,T(\alpha_n)\}$. We now show that S is a basis of R(T).

a) S is LI: Let $a_{k+1}T(\alpha_{k+1})+a_{k+2}T(\alpha_{k+2})+.....+a_nT(\alpha_n)=\overline{0}$

$\Rightarrow T(a_{k+1}\alpha_{k+1}+a_{k+2}\alpha_{k+2}+.....+a_n\alpha_n)=\overline{0}$

$\Rightarrow a_{k+1}\alpha_{k+1}+a_{k+2}\alpha_{k+2}+.....+a_n\alpha_n \in N(T)$

$\Rightarrow \exists a_1,a_2,...,\alpha_k \in F \ni a_{k+1}\alpha_{k+1}+a_{k+2}\alpha_{k+2}.....+a_n\alpha_n = a_1\alpha_1+....+a_k\alpha_k$

$\Rightarrow -a_1\alpha_1-a_2\alpha_2......-a_k\alpha_k +-a_{k+1}\alpha_{k+1}+....+a_n\alpha_n=\overline{0}$

$\Rightarrow -a_1=0,-a_2=0,...,-a_k=0;a_{k+1}=0,a_{k+2}=0,....a_n=0$, since B is LI.

$\Rightarrow a_{k+1}=0,.......,a_n=0 \quad \therefore S$ is L.I.

b) $L(S)=R(T)$: Clearly $L(S)\subseteq R(T)$.

Let $\beta\in R(T)\Rightarrow \exists \alpha\in U \ni T(\alpha)=\beta$

Now $\alpha\in U\Rightarrow \exists a_1,a_2,....,a_k\alpha_{k+1},....,a_n\in F\ni \alpha=a_1\alpha_1+.....+a_k\alpha_k+...+a_{k+1}\alpha_{k+1}+....+a_n\alpha_n$

$\therefore \beta=T(\alpha)=T(a_1\alpha_1+a_2\alpha_2+.....+a_k\alpha_k+a_{k+1}\alpha_{k+1}+a_{k+2}\alpha_{k+2}+....+a_n\alpha_n)$

$=a_1T(\alpha_1)+a_2T(\alpha_2)+...+a_kT(\alpha_k)+a_{k+1}T(\alpha_{k+1})+a_{k+2}T(\alpha_{k+2})+....+a_nT(\alpha_n)$

$=a_1.0+a_2.0+...+a_k.0+a_{k+1}T(\alpha_{k+1})+a_{k+2}T(\alpha_{k+2})+....+a_nT(\alpha_n)$

$=a_{k+1}T(\alpha_{k+1})+a_{k+2}T(\alpha_{k+2})+....+a_nT(\alpha_n)\in L(S)$

$\therefore R(T)\subseteq L(S).\Rightarrow L(S)=R(T)\therefore S$ is a basis of R(T)

$\therefore \rho(T)=$ Number of elements of $S=n-k=n-\upsilon(T)$.

$\Rightarrow \rho(T)+\upsilon(T)=n=\dim U$; Rank T + nullity T = $\dim U$. Hence the theorem.

Solved Problems:

7.5.10 Problem: Let $T:\mathbb{R}^3\rightarrow\mathbb{R}^2$ be defined by $T(a_1,a_2,a_3)=(a_1-a_2,2a_3)$ for all $(a_1,a_2,a_3)\in\mathbb{R}^3$. Find the rank T, nullity T and verify $\rho(T)+\upsilon(T)=n=\dim U$.

Solution: We have $N(T)=\{(a_1,a_2,a_3)/T(a_1,a_2,a_3)=(0,0)\}$

$= \{(a_1, a_2, a_3)/(a_1 - a_2, 2a_3) = (0,0)\}$

$= \{(a_1, a_2, a_3)/(a_1 = a_2, a_3 = 0\}$

$= \{(a, a, 0)/a \in \mathbb{R}\}$

= Span of (1, 1, 0)

$\therefore \upsilon(T) = \dim N(T) = 1$

$R(T) = \{T(a_1, a_2, a_3)/(a_1, a_2, a_3) \in \mathbb{R}^3\}$

= $= \{(a_1 - a_2, 2a_3)/a_1, a_2, a_3 \in \mathbb{R}\}$

R(T) = Span of $\{(T(e_1), T(e_2), T(e_3)\}$, where $\{e_1, e_2, e_3\}$ is the standard basis of $\mathbb{R}^3$.

= Span of $\{(1,0), (-1,0), (0,2)\}$

= Span of $\{(1,0), (0,1)\}$

= $\mathbb{R}^2$

$\therefore \rho(T) = \dim R(T) = 2$

Now $P(T) + V(T) = 2 + 1 = 3 = \dim \mathbb{R}^3$

7.5.11 Problem: Find the null space, range, rank and nullity of the linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^3$ defined by $T(x, y) = (x + y, x - y, y)$ for all $(x, y) \in \mathbb{R}^2$.

Solution: $N(T) = \{(x, y)/T(x, y) = \overline{0}\} = \{(x, y)/x + y, x - y, y) = (0,0,0\}$

$= \{(x, y)/x + y = 0, x - y = 0, y = 0\} = \{(0,0)\}$

$\therefore$ Nullity T is = 0.

$R(T) = \{T(x, y)/(x, y) \in \mathbb{R}^2\} = \{(x + y, x - y, y/x, y \in \mathbb{R}\}$

Since rank T + nullity T = $\dim \mathbb{R}^2 = 2$, it follows that rank T = 2.

Also we note that R(T) is spanned by $\{T(1,0), T(0,1)\} = \{(1,1,0), (1,-1,1)\}$.

7.5.12 SAQ: Let T_1, T_2 be two operators on $\mathbb{R}^2$ defined by $T_1(x,y)=(y,x), T_2(x,y)=(x,0)$. Show that $T_1T_2 \neq T_2T_1$.

7.5.13 SAQ : If T is a linear operator on $\mathbb{R}^2$ defined by $T(x,y)=(x-y,y)$ then find $T^2(x,y)$.

7.5.14 SAQ: Give an example of two linear operators T and S on $\mathbb{R}^2$ such that $TS=O$ but $ST \neq O$.

7.6 Answers to SAQ's:

7.4.6 SAQ: $T:\mathbb{R}^2 \to \mathbb{R}^2$. $T(a_1,a_2)=(2a_1+a_2,a_1) \forall (a_1,a_2) \in \mathbb{R}^2$

Let $\alpha=(\alpha_1,\alpha_2), \beta=(\beta_1,\beta_2) \in \mathbb{R}^2$ and $a,b \in \mathbb{R}$.

$$T(a\alpha+b\beta)=T(a\alpha_1+b\beta_1, a\alpha_2+b\beta_2)$$

$$=(2(a\alpha_1+b\beta_1)+a\alpha_2+b\beta_2, a\alpha_1+b\beta_1)$$

$$=a(2\alpha_1+\alpha_2,\alpha_1)+b(2\beta_1+\beta_2,\beta_1)$$

$$=aT(\alpha)+bT(\beta) \forall \alpha,\beta \in \mathbb{R}^2 \text{ and } \forall a,b \in \mathbb{R}$$

So T is a linear transformation.

7.4.14 SAQ: $T:\mathbb{R}^2 \to \mathbb{R}^2$. $T(a_1,a_2)=(a_1,-a_2) \forall (a_1,a_2) \in \mathbb{R}^2$

Let $\alpha=(\alpha_1,\alpha_2), \beta=(\beta_1,\beta_2) \in \mathbb{R}^2$ and $a,b \in \mathbb{R}$.

$$T(a\alpha+b\beta)=T(a\alpha_1+b\beta_1, a\alpha_2+b\beta_2)$$

$$=(a\alpha_1+b\beta_1, -a\alpha_2-b\beta_2)$$

$$=a(\alpha_1,-\alpha_2)+b(\beta_1,-\beta_2)$$

$$=aT(\alpha)+bT(\beta) \forall \alpha,\beta \in \mathbb{R}^2 \text{ and } \forall a,b \in \mathbb{R}$$

$\therefore$ T is a linear transformation.

7.4.15 SAQ : $T:\mathbb{R}^2 \to \mathbb{R}^2$. $T(a_1,a_2)=(a_1,0) \forall (a_1,a_2) \in \mathbb{R}^2$

Let $\alpha=(\alpha_1,\alpha_2), \beta=(\beta_1,\beta_2) \in \mathbb{R}^2$ and $a,b \in \mathbb{R}$.

$$\therefore T(a\alpha+b\beta)=T(a\alpha_1+b\beta_1, a\alpha_2+b\beta_2)$$

$= (a\alpha_1 + b\beta_1, 0) = a(\alpha_1, 0) + b(\beta_1, 0)$

$= aT(\alpha) + bT(\beta) \forall a, b \in \mathbb{R}$ and $\forall \alpha, \beta \in \mathbb{R}^2$

$\therefore$ T is a linear transformation.

7.5.12 SAQ: $T_1, T_2 : \mathbb{R}^2 \to \mathbb{R}^2$, $T_1(x, y) = (y, x)$

$$T_2(x, y) = (x, 0) \forall (x, y) \in \mathbb{R}^2$$

$T_1T_2(x, y) = T_1\left[T_2(x, y)\right] = T_1(x, o) = (o, x)$

$T_2T_1(x, y) = T_2\left[T_1(x, y)\right] = T_2(y, x) = (y, o)$

$T_1T_2(x, y) \neq T_2T_1(x, y)$, if $x \neq o$ or $y \neq o$.

$\therefore T_1T_2 \neq T_2T_1$

7.5.13 SAQ: $T : \mathbb{R}^2 \to \mathbb{R}^2, T(x, y) = (x - y, y)$

$T^2(x, y) = T\left(T(x, y)\right) = T(x - y, y)$

$= (x - 2y, y) \forall (x, y) \in \mathbb{R}^2$

7.5.14 SAQ: Give an example to show that $TS = O$ but $ST \neq O$

Define $T : \mathbb{R}^2 \to \mathbb{R}^2$ by $T(x, y) == (1, 0) \forall (x, y) \in \mathbb{R}^2$

$S : \mathbb{R}^2 \to \mathbb{R}^2$ by $S(x, y) = (0, x) \forall (x, y) \in \mathbb{R}^2$

$TS(x, y) = T\left[S(x, y)\right] = T(0, x) = (1, 0) \in \mathbb{R}^2$

$ST(x, y) = S\left[T(x, y)\right] = S(1, 0) = (0, 1)$

Here $TS(x, y) \neq ST(x, y) \therefore TS \neq ST$

7.7 Summary:
Linear transformation and properties, null space, range, rank, nullity are discussed.

7.8 Technical Terms:
Linear transformation, range, null space, rank, nullity.

7.9 Exercise

7.9.1 Let $V = F^{n \times n}; M \in V$. Define $T : V \to V$ by $T(A) = AM + MA \forall A \in V$. Show that T is linear.

7.9.2 Give an example of a linear operator T on $\mathbb{R}^3$ such that $T \neq O, T^2 \neq O$ but $T^3 = O$

7.9.3 Let $V = F^{2\times 2}$. Let $P = \begin{pmatrix} 1 & -1 \\ -2 & 2 \end{pmatrix} \in V$. Define $T : V \to V$ by $T(A) = PA \; \forall \; A \in V$ Find nullity T

7.9.4 Let $T : \mathbb{R}^4 \to \mathbb{R}^3$ be defined by $T(e_1) = (1,1,1)$

$T(e_2) = (1,1,1), T(e_3) = (1,0,0), T(e_4) = (1,0,1)$

Verify that $\rho(T) + \upsilon(T) = \dim \mathbb{R}^4$.

7.10 Answers to Exercises:

7.9.2 $T : \mathbb{R}^3 \to \mathbb{R}^3, T(a,b,c) = (0,a,b) \; \forall \; (a,b,c) \in \mathbb{R}^3$

7.9.3 2

7.11 Model Examination Questions:

7.11.1 Define a linear transformation. Prove that $T(\overline{0}) = \overline{0}, T(-\alpha) = -T(\alpha)$

7.11.2 Define range and null space of a linear transformation.

Show that range and null space are subspaces.

7.11.3 Define rank and nullity. Prove that $\rho(T) + \upsilon(T) = \dim U$

7.11.4 Show that $L(U,V)$ is a vector space.

7.11.5 Show that $\dim L(U,V) = mn$.

7.12. Reference Books:

1. Stephen H: Fried berg and others - Linear Algebra - Prentice Hall India Pvt. Ltd., New Delhi.

2. K. Hoffman and Kunze - Linear Algebra.

2nd Edition - Prentice Hall, New Jersey - 1971.

- A. Satyanarayana Murty

LESSON - 8

MATRIX REPRESENTATION OF A LINEAR TRANSFORMATION

8.1 Objective of the Lesson:

In Chapter 7, we discussed linear transformation from an n-dimensional vector space to an m-dimenstional vector space. In this chapter we represent such a linear transformation as an m x n matrix and study some properties and discuss some problems.

8.2. Structure of the Lesson:

This lesson lhas the following components.

8.3 Introduction:

In this chapter we discuss matrix of a linear transformation from a vector space to a vector space relative to ordered bases. We also study the matrix of a product of linear transformations and the invertibility and isomorphism of linear transformations.

8.4 Matrix Representation:

First we define an ordered basis of an n-dimensional vector space V(F) as a basis for V endowed with a specific order.

For example in F^3, the standard basis $B=\{e_1,e_2,e_3\}$ is considered as an ordered basis. Also $B_1=\{e_2,e_3,e_1\}$ is an ordered basis of F^3, but $B_1 \neq B$ as ordered bases. For $F^n, \{e_1,e_2...,e_n\}$ is considered as the standard ordered basis, where $e_1=(1,0,0,..,0), e_2=(0,1,0,...0)$ etc.

8.4.1 Definition: If $B=\{\alpha_1,\alpha_2,...,\alpha_n\}$ is an ordered basis of an n-dimensional vector space $V(F)$ and $\alpha \in V$, then α can be uniquely expressed as $\alpha = \sum_{in}^{n} a_i\alpha_i$, where $a_i \in F$.

We define the coordinate vector of α relative to the ordered basis B as $[\alpha]_B = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$

8.4.2 Example: For $V_3(\mathbb{R})$, $B=\{e_1,e_2,e_3\}$ is the standard ordered basis; where $e_1=(1,0,0), e_2=(0,1,0), e_3=(0,0,1)$

If $\alpha=(4,-2,3)$, then $[\alpha]_B = \begin{bmatrix} 4 \\ -2 \\ 3 \end{bmatrix}$, since

$$\alpha=(4,-2,3)=4(1,0,0)-2(0,1,0)+3(0,0,1)$$

8.4.3 Example: For $V_3(\mathbb{R})$, we know that $B_1=\{\alpha_1,\alpha_2,\alpha_3\}$ is on ordered basis where

$\alpha_1=(1,0,0), \alpha_2=(1,1,0), \alpha_3=(1,1,1)$, so $\alpha=(4,3,2) \Rightarrow [\alpha]_{B_1} = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$, since

$$(4,3,2)=1(1,0,0)+1(1,1,0)+2(1,1,1)$$

8.4.4 Definition: Let $U(F)$ and $V(F)$ be vector spaces and $\dim U=n, \dim V=m$. Let $B_1=\{\alpha_1,\alpha_2,...,\alpha_n\}$ be an ordered basis of U and let $B_2=\{\beta_1,\beta_2,...,\beta_m\}$ be an ordered basis of V.

Let $T:U \to V$ be a linear transformation.

Since $T(\alpha_1), T(\alpha_2), ..., T(\alpha_n) \in R(T) \subseteq V$, these vectors can be uniquely expressed as linear combinations of elements of B_2.

For each $j, 1 \le j \le n, \exists$ unique $a_{ij} \in F \ni T(\alpha_j) = \sum_{i=1}^{n} a_{ij}\beta_i$

The matrix $A = \left[a_{ij}\right]_{m \times n}$ is called the matrix of T relative to the ordered bases B_1, B_2 and we write $\left[T; B_1, B_2\right] = A$.

If U = V and $B_1 = B_2$, we write $A = \left[T; B_1, B_1\right] = \left[T\right]_{B_1}$.

8.4.5 Note: $\left[T(\alpha_j)\right]_{B_2} = j$ th column of A.

8.4.6 Example: If I and O are the identity and zero operators an n-dimensional vector space V and $B = \{\alpha_1, \alpha_2, ..., \alpha_n\}$ is an ordered basis of V, then $\left[I\right]_B = I_{n \times n}$ and $\left[O\right]_B = O_{n \times n}$.

since $I(\alpha_j) = \alpha_j = o\alpha_1 + o\alpha_2 + \ldots + 1\alpha_j + \ldots + o\alpha_n$,

$O(\alpha_j) = \bar{0} = o\alpha_1 + o\alpha_2 + \ldots + o\alpha_n$ for each $j = 1, 2, ..., n$.

8.4.7 Theorem: Let U(F), V(F) be vector spaces, $\dim U = n, \dim V = m$ and let B_1, B_2 be ordered bases of U and V respectively. If $T_1, T_2 \in L(U, V)$, then $\left[T_1 + T_2; B_1, B_2\right] = \left[T_1; B_1, B_2\right] + \left[T_2; B_1, B_2\right]$ and $\left[aT_1; B_1, B_2\right] = a\left[T_1; B_1, B_2\right]$, where $a \in F$.

Proof: Let $B_1 = \{\alpha_1, \alpha_2, ..., \alpha_n\}$, $B_2 = \{\beta_1, \beta_2, ..., \beta_m\}$

Let $\left[T_1; B_1, B_2\right] = \left[a_{ij}\right]_{m \times n}$ and $\left[T_2; B_1, B_2\right] = \left[b_{ij}\right]_{m \times n}$

$\Rightarrow T_1(\alpha_j) = \sum_{i=1}^{n} a_{ij}\beta_i$, $T_2(\alpha_j) = \sum_{i=1}^{m} b_{ij}\beta_i$, for j = 1,2,.....n.

Now $(T_1 + T_2)(\alpha_j) = T_1(\alpha_j) + T_2(\alpha_j) = \sum_{i=1}^{m} a_{ij}\beta_i + \sum_{i=1}^{m} b_{ij}\beta_i = \sum_{i=1}^{m} (a_{ij} + b_{ij})\beta_i$

$\therefore \left[T_1 + T_2; B_1, B_2\right] = \left[a_{ij} + b_{ij}\right]_{m \times n} = \left[a_{ij}\right]_{m \times n} + \left[b_{ij}\right]_{m \times n} = \left[T_1; B_1, B_2\right] + \left[T_2; B_1, B_2\right]$

Let $a \in F$ N.ow $(aT_1)(\alpha_j) = aT_1(\alpha_j) = a\sum_{i=1}^{m} a_{ij}\beta_i = \sum_{i=1}^{m} aa_{ij}\beta_i, \forall j = 1, 2, ...n$

$\therefore [aT_1; B_1, B_2] = [aa_{ij}]_{m\times n} = a[a_{ij}]_{m\times n} = a[T_1; B_1, B_2]$

8.4.8 Note: We defined the matrix of a linear transformation $T : U \to V$ w.r.t. the ordered bases $B_1 = \{\alpha_1, \alpha_2, ..., \alpha_n\}$, $B_2 = \{\beta_1, \beta_2, ..., \beta_n\}$ of U and V respectively as $A = [a_{ij}]_{m\times n}$, where

$T(\alpha_j) = \sum_{i=1}^{m} a_{ij}\beta_i$, for $j = 1, 2, ...n$

So, for each linear transformation from an n-dimensional vector space to an m-dimensional vector space, $\exists$ an m x n matrix $A = [a_{ij}]$ which is = $[T, B_1, B_2]$

Now if U, V are vector spaces, $B_1 = \{\alpha_1, \alpha_2, ..., \alpha_n\}$, $B_2 = \{\beta_1, \beta_2, ..., \beta_m\}$ are ordered bases of U, V respectively and $A = [a_{ij}]_{m\times n}$, then $\exists$ a linear transformation $T : U \to V \ni [T, B_1, B_2] = A$

If we define $T(\alpha_j) = \sum_{i=1}^{m} a_{ij}\beta_i$, $j = 1, 2, ...n$, then T is a uniquely defermined linear transformation and $[T, B_1, B_2] = A$ (If we write $\sum_{i=1}^{m} a_{ij}\beta_i = \delta_j, j = 1, 2, ..., n$, then $\delta_1, ..., \delta_n$ are n vectors in V and $\exists$ a unique linear transformation $T : U \to V \ni T(\alpha_j) = \delta_j, 1 \le j \le n$. This result was proved in chapter 7 - see theorem 7.4.18).

8.4.9 Theorem: Let $A \in F^{m\times n}$. Define $T : F^{n\times 1} \to F^{m\times 1}$ by $T(X) = A(X)$, for all $X \in F^{n\times 1}$. Then T is a linear transformation. If B_1, B_2 are standard ordered bases of $F^{n\times 1}$ and $F^{m\times 1}$ respectively, then $[T, B_1, B_2] = A$.

Proof: We proved that T is a linear transformation. (See theorem 7.4.13).

Let $B_1 = \{e_1, e_2, ...e_n\}$, $B_2 = \{f_1, f_2, ..., f_m\}$ be standard ordered bases of $F^{n\times 1}, F^{m\times 1}$ respectively. ($e_i \in F^{n\times 1}; i^{\text{th}}$ component $e_i = 1$ and other components = 0).

$$T(e_j) = Ae_j = [a_{ik}]_{m\times n} \begin{bmatrix} o \\ o \\ \vdots \\ o \\ 1 \\ o \\ \vdots \\ o \end{bmatrix} = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix} = a_{1j} f_1 + a_{2j} f_2 + ... + a_{ij} f_2 + ... + a_{mj} f_m$$

$$= \sum_{i=1}^{m} a_{ij} f_i \ .$$

$$\therefore [T : B_1, B_2] = [a_{ij}]_{m\times n} = A.$$

8.4.10 Note: T is called the left multiplication transformation and is denoted by L_A.

8.4.11 Definition: Let A be an $m \times n$ matrix over F. The linear transformation $L_A : F^{n\times 1} \to F^{m\times 1}$ by $L_A(X) = A(X) \forall X \in F^n$ (or $F^{n\times 1}$) is called the left multiplication transformation.

8.4.12 Theorem: Let $U(F), V(F)$ be vector spaces, $\dim U = n, \dim V = m$ and B_1, B_2 be ordered bases of U and V respectively.

Let $T : U \to V$ be a linear transformation and $\alpha \in U$. Then $[T; B_1, B_2][\alpha]_{B_1} = [T(\alpha)]_{B_2}$, where $[\alpha]_{B_1}$ is the coordinate matrix of α w.r.t. the ordered basis B_1.

Proof: Let $B_1 = \{\alpha_1, \alpha_2, ..., \alpha_n\}$, $B_2 = \{\beta_1, \beta_2, ..., \beta_m\}$

Let $A = [a_{ij}]_{m\times n} = [T; B_1, B_2]$

For $\alpha_j \in B_1$, we have $T(\alpha_j) = \sum_{i=1}^{m} a_{ij}\beta_i$

Since $\alpha \in U$, $\exists$ unique $b_1, b_2, ...b_n \in F \ni \alpha = \sum_{j=1}^{n} b_j \alpha_j$

$$\Rightarrow T(\alpha) = T\left[\sum_{j=1}^{n} b_j \alpha_j\right] = \sum_{j=1}^{n} b_j T(\alpha_j)$$

$$= \sum_{j=1}^{n} b_j \sum_{i=1}^{m} a_{ij}\beta_j = \sum_{i=1}^{m}\left(\sum_{j=1}^{n} b_j a_{ij}\right)\beta_i = \sum_{j=1}^{n} c_i\beta_i \text{ where } c_i = \sum_{j=1}^{m} b_j a_{ij}$$

$$\therefore [T(\alpha)]_{B_2} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{bmatrix}$$

$$= \begin{bmatrix} b_1 a_{11} + b_2 a_{12} + b_3 a_{13} + ...+ b_n a_{1n} \\ b_1 a_{21} + b_2 a_{22} + b_3 a_{23} + ...+ b_n a_{2n} \\ \vdots \\ b_1 a_{m1} + b_2 a_{m2} + b_3 a_{m3} + ...+ b_n a_{mn} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11} & a_{12} & a_{13}........a_{1n} \\ a_{21} & a_{22} & a_{23}........a_{2n} \\ \cdots & .... & .... \\ a_{m1} & a_{m2} & a_{m3}.........a_{mn} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

$$= [T ; B_1, B_2][\alpha]_{B_1}$$

Hence the Theorem.

8.4.13 Corollary: If $T : V \to V$ is a linear operator, $\dim V = n$, B is an ordered basis of V and $\alpha \in V$, then $[T(\alpha)]_B = [T]_B [\alpha]_B$

8.5 Composition of linear transformation and matrix multiplication:

We defined the composition of two linear transformation as : $TS(\alpha) = T[S(\alpha)] \forall \alpha \in U$.

8.5.1 Theorem: Let $U(F), V(F), W(F)$ be finite dimensional vector space and let B_1, B_2, B_3 be ordered bases of U, V, W respectively.

Let $T : U \to V, S : V \to W$ be linear transformations. then $[ST; B_1, B_3] = [S; B_2, B_3][T; B_1, B_2]$

Proof: We know that $ST : U \to W$ is defined by $(ST)(\alpha) = S[T(\alpha)], \forall \alpha \in U$ is a linear transformation.

Let $B_1 = \{\alpha_1, \alpha_2, ..., \alpha_p\}$, $B_2 = \{\beta_1, \beta_2, ..., \beta_n\}$, $B_3 = \{\gamma_1, \gamma_2, ..., \gamma_m\}$ be ordered bases of U,V,W respectively.

Let $[T; B_1, B_2] = [b_{kj}]_{n \times p}$

$$[S; B_2, B_3] = [a_{ik}]_{m \times n}$$

$$\Rightarrow \quad T(\alpha_j) = \sum_{k=1}^{n} b_{kj} \beta_k, \ j = 1, 2, ..., p$$

$$S(\beta_k) = \sum_{i=1}^{n} a_{ik} \gamma_i, k = 1, 2, ..., n$$

Now for each $j = 1, 2, ..., p$, we have

$$(ST)(\alpha_j) = S[T(\alpha_j)]$$

$$= S\left(\sum_{k=1}^{n} b_{kj} \beta_k\right) = \sum_{k=1}^{n} b_{kj} S(\beta_k)$$

$$= \sum_{k=1}^{n} b_{kj} \sum_{i=1}^{n} a_{ik} \gamma_i$$

$$= \sum_{i=1}^{n} \left(\sum_{k=1}^{n} a_{ik} b_{kj}\right) \gamma_i$$

$$= \sum_{i=1}^{n} c_{ij} \gamma_i \text{, where } c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$$

$$\therefore [ST; B_1, B_3] = [c_{ij}]_{m \times p} = \left[\sum_{k=1}^{n} a_{ik} b_{kj}\right]_{m \times p} = [a_{ik}]_{m \times n} [b_{kj}]_{n \times p}$$

$$= [S; B_2, B_3][T; B_1, B_2]$$

8.5.2 Corollary: If T, S are linear operators on an n-dimensional vector space V and B is an ordered basis of V, then $[ST]_B = [S]_B [T]_B$.

8.5.3 Theorem: Let $T : F^{n\times1} \to F^{m\times1}$ be a linear transformation. Then $\exists A \in F^{m\times n} \ni T = L_A$.

Proof: Let B_1, B_2 be the standard ordered bases of $F^{n\times1}$ and $F^{m\times1}$ respectively.

Then $\exists$ a unique matrix $A \in F^{m\times n} \ni [T; B_1, B_2] = A$.

$\therefore$ For every $X \in F^{n\times1}$, we have

$$[T(X)]_{B_2} = [T; B_1, B_2][X]_{B_1}$$

$$= A[X]_{B_1}$$

$$\therefore T(X) = AX = L_A(X) \forall X \in F^{n\times1}$$

$$\therefore T = L_A$$

We observe that ;

8.5.4 Theorem: If $A, B \in F^{m\times n}$, then

1) $L_{A+B} = L_A + L_B$

2) $L_{aA} = aL_A, a \in F$

3) $L_{AB} = L_A o L_B$

Proof: $L_{A+B}(X) = (A+B)X = AX + BX = L_A(X) + L_B(X) \forall X \in F^{n\times1}$

$$\Rightarrow L_{A+B} = L_A + L_B$$

$$L_{aA}(X) = (aA)X = a(AX) = aL_A(X) \forall X \in F^{n\times1}$$

$$\Rightarrow L_{aA} = aL_A$$

$$L_{AB}(X) = (AB)X = A(BX) = L_A(BX) = L_A[L_B(X)]$$

$$\Rightarrow L_{AB}(X) = L_A o L_B(X) \forall X \in F^{n\times1} \Rightarrow L_{AB} = L_A o L_B.$$

8.5.5 Theorem: Let $U(F), V(F)$ be vector spaces, dim $U = n \quad \dim V = m$.

Then $L(U,V) \cong F^{m\times n}$

Proof: We know that $L(U,V)$ and $F^{m\times n}$ are vector spaces over F. Let $B_1 = \{\alpha_1, \alpha_2, ..., \alpha_n\}$, $B_2 = \{\beta_1, \beta_2, ..., \beta_m\}$ be ordered bases of U and V respectively.

Let $T_1, T_2 \in L(U,V)$ $\quad [T_1; B_1, B_2] = A = [a_{ij}]_{m\times n}$

and $[T_2; B_1, B_2] = B = [b_{ij}]_{m\times n}$

$\therefore$ For $j = 1, 2, ..., n$, we have $T_1(\alpha_j) = \sum_{i=1}^{m} a_{ij}\beta_i$ and $T_2(\alpha_j) = \sum_{i=1}^{m} b_{ij}\beta_i$

Define $\phi : L(U,V) \to F^{m\times n}$ by

$\phi(T) = [T; B_1, B_2] \forall T \in L(U,V)$

We show that ϕ is an isomorphism.

a) Let $a, b \in F, T_1, T_2 \in L(U,V)$

$\therefore \phi(aT_1 + bT_2) = [aT_1 + bT_2; B_1, B_2]$

$= [aT_1; +B_1, B_2] + [bT_2; +B_1, B_2]$

$= a[T_1; B_1, B_2] + b[T_2; B_1, B_2]$

$= a\phi(T_1) + b\phi(T_2) \forall T_1, T_2 \in L(U,V)$ and $\forall a, b \in F$

$\therefore \phi$ is a linear transformation.

b) Let $T_1, T_2 \in L(U,V)$ and $\phi(T_1) = \phi(T_2)$

$\Rightarrow [T_1; B_1, B_2] = [T_2; B_1, B_2]$

$\Rightarrow A = B \Rightarrow [a_{ij}] = [b_{ij}] \Rightarrow a_{ij} = b_{ij} \forall i, j, 1 \le i \le m, 1 \le j \le n$

$\therefore$ for each $\alpha_j \in B_1$, we have $T_1(\alpha_j) = \sum_{i=1}^{m} a_{ij}\beta_i = \sum_{i=1}^{m} b_{ij}\beta_i = T_2(\alpha_j); \forall \alpha_j \in B$

$\therefore T_1, T_2$ agree an a basis B_1 of $U \Rightarrow T_1 = T_2$ on U $\therefore \phi$ is one-one.

c) Let $A \in F^{m\times n}$ and $A = \left[a_{ij}\right]_{m\times n}$. Write $\delta_j = \sum_{i=1}^{m} a_{ij}\beta_i \forall j = 1, 2, ..., n$.

So $\delta_1, \delta_2, ...\delta_n$ are n vectors in V.

$\{\alpha_1, ..., \alpha_n\}$ is a basis of U and $\delta_1, ...\delta_n$ are n vectors in V.

$\therefore \exists$ a unique linear transformation $T: U \to V \ni T(\alpha_j) = \delta_j, j = 1, 2, ..., n$.

$$\Rightarrow T(\alpha_j) = \sum_{i=1}^{m} a_{ij}\beta_1, \forall j = 1, 2, ..., n$$

$$\Rightarrow [T; B_1, B_2] = A \Rightarrow \phi(T) = A$$

$\therefore \phi$ is onto $\therefore \phi$ is an isomorphism $\therefore L(U,V) \cong F^{m\times n}$

Hence the theorem.

We observe that $\dim F^{m\times n}$ is = $mn = \dim L(U,V) = mn$. (We proved this independently)

8.6 Invertibility and isomorphism.

We know that a function $f: A \to B$ is said to be invertible if there exists a function $g: B \to A$ such that $gof = I_A$ and $fog = I_B$ and g is called inverse of f and we write $g = f^{-1}$. Further, f and g are bijections and $f(x) = y \Leftrightarrow f^{-1}(y) = x$. Since linear transformations are also functions, it is natural to expect that the inverse of a linear transformation is also a linear transformation.

8.6.1 Definition: Let $U(F)$ and $V(F)$ be vector spaces. Let $T: U \to V$ be a bijective mapping. Then the mapping $S: V \to U$ defined by $S(B) = \alpha \Leftrightarrow T(\alpha) = \beta, \alpha \in U, \beta \in V$, is called the inverse of T and it is denoted by T^{-1}. (If such a mapping S exists, then we say that T is invertible).

8.6.2 Note: 1. $T^{-1}(\beta) = \alpha \Leftrightarrow T(\alpha) = \beta, \alpha \in U, \beta \in V$

2. T is invertible iff T is a bijection.

3. T is a bijection iff T^{-1} is a bijection.

4. If $T: U \to V$, $S: V \to W$ are bijections, then $ST: U \to W$ is also a bijection and $(ST)^{-1} = T^{-1}S^{-1}, (T^{-1})^{-1} = T$

8.6.3 Theorem: Let $U(F)$, $V(F)$ be vector spaces and $T:U\to V$ be a bijective linear transformation. Then $T^{-1}:V\to U$ is also a linear transformation.

Proof: Since $T:U\to V$ is a bijection, $T^{-1}:V\to U$ is also a bijection.

To prove that T^{-1} is a linear transformation, let $\alpha,\beta\in V, a,b\in F$.

Since $T:U\to V$ is onto, $\alpha,\beta\in V\Rightarrow\exists\alpha_1,\beta_1\in U\ni T(\alpha_1)=\alpha$ and $T(\beta_1)=\beta$. Since T is a bijection, we have

$T(\alpha_1)=\alpha\Rightarrow T^{-1}(\alpha)=\alpha_1$ and $T(\beta_1)=\beta\Rightarrow T^{-1}(\beta)=\beta_1$

Now $a\alpha+b\beta=aT(\alpha_1)+bT(\beta_1)=T(a\alpha_1+b\beta_1)$

Since T is a bijection, $T^{-1}(a\alpha+b\beta)=a\alpha_1+b\beta_1=aT^{-1}(\alpha)+bT^{-1}(\beta)$ and this is true for all $\alpha,\beta\in V$ and for all $a,b\in F$. Hence T^{-1} is a linear transformation.

8.6.4 Note: 1) If $T:U\to V$ is an isomorphism iff $T^{-1}:V\to U$ is an isomorphism.

2) If inverse of $T:U\to V$ exists, then T is said to be an ivertible linear transformation and $ToT^{-1}=I_V, T^{-1}oT=I_U$.

8.6.5 Definition: If there is an isomorphism (i.e. one-one onto linear transformation) from a vector space U(F) to a vector space V(F), then we say that U and V are isomorphic and we write $U\cong V$.

8.6.6 Theorem: Let U(F), V(F) be finite dimensional vector spaces. Then $U\cong V\Leftrightarrow\dim U=\dim V$.

Proof: 1. Suppose $U\cong V$

$\Rightarrow\exists$ a one-one onto linear transformation $T:U\to V$

Let $\dim U=n$ and $S=\{\alpha_1,\alpha_2,.....,\alpha_n\}$ be a basis of U.

Let $S^1=\{T(\alpha_1),T(\alpha_2),.....,T(\alpha_n)\}$. We show that S^1 is a basis of V.

a) S^1 is L. I: Let $a_1T(\alpha_1)+a_2T(\alpha_2)+......+a_nT(\alpha_n)=\overline{0}$

$\Rightarrow T(a_1\alpha_1+a_2\alpha_2+......+a_n\alpha_n=\overline{0}=T\left(\overline{0}\right)$

$\Rightarrow a_1\alpha_1+a_2\alpha_2+......+a_n\alpha_n=\overline{0}$ ($\because T$ is one-one)

$\Rightarrow a_1=0,a_2=0,....,a_n=0$ ($\because S$ is L1)

$\therefore S^1$ is L.I.

b) $L(S^1) = V$: Let $\beta \in V$. Since $T : U \to V$ is onto, $\exists \alpha \in U \ni T(\alpha) = \beta$

Since $\alpha \in U$, $\exists a_1, a_2, ..., a_n \in F \ni \alpha = a_1\alpha_1 + a_2\alpha_2 + ... + a_n\alpha_n$

$\Rightarrow \beta = T(\alpha) = T(a_1\alpha_1 + a_2\alpha_2 + ... + a_n\alpha_n)$

$= a_1T(\alpha_1) + a_2T(\alpha_2) + ... + a_nT(\alpha_n)$

$\Rightarrow \beta \in L(S^1)$. $\therefore V \subseteq L(S^1)$

Also since $L(S^1) \subseteq V$, we have $L(S^1) = V$

So, S^1 is a basis of V so that $\dim V = n$. Hence $\dim U = \dim V$.

2. Suppose $\dim U = \dim V = n$, Let $S = \{\alpha_1, \alpha_2, ..., \alpha_n\}$ and $S^1 = \{\beta_1, \beta_2, ..., \beta_n\}$ be bases of U and V respectively.

Let $\alpha \in U \Rightarrow \exists a_1, a_2, ..., a_n \in F \ni \alpha = a_1\alpha_1 + a_2\alpha_2 + ... + a_n\alpha_n$

Define $T : U \to V$ by $T(\alpha) = T(a_1\alpha_1 + a_2\alpha_2 + ... + a_n\alpha_n) = a_1\beta_1 + a_2\beta_2 + ... + a_n\beta_n$

for all $\alpha = a_1\alpha_1 + a_2\alpha_2 + ... + a_n\alpha_n \in U$

Let $a, b \in F$ and $\alpha = a_1\alpha_1 + a_2\alpha_2 + ... + a_n\alpha_n, \beta = b_1\alpha_1 + b_2\alpha_2 + ... + b_n\alpha_n \in U$.

$$\therefore T(a\alpha + b\beta) = T\left[\sum_{i=1}^{n}(aa_i + bb_i)\alpha_i\right] = \sum_{i=1}^{n}(aa_i + bb_i)\beta_i = \sum_{i=1}^{n} aa_i\beta_i + \sum_{i=1}^{n} bb_i\beta_1$$

$$= a\sum_{i=1}^{n} a_i\beta_1 + b\sum_{i=1}^{n} b_i\beta_i = aT(\alpha) + bT(\beta) \forall \alpha, \beta \in U \text{ and } \forall \alpha, \beta \in F.$$

$\therefore T$ is a linear transformation.

T is one-one : Let $\alpha, \beta \in U$ and $T(\alpha) = T(\beta)$

$$\alpha, \beta \in U \Rightarrow \exists a_i, b_i \in F \ni \alpha = \sum_{i=1}^{n} a_i\alpha_i, \beta = \sum_{i=1}^{n} b_i\alpha_i$$

Now $T(\alpha) = T(\beta) \Rightarrow \sum_{i=1}^{n} a_i\beta_i = \sum_{i=1}^{n} b_i\beta_i$

$$\Rightarrow \sum_{i=1}^{n}(a_i - b_i)\beta_i = \overline{0}$$

$\Rightarrow a_i - b_i = 0 \;\forall\; i = 1, 2, ..., n$ since S^1 is L.I.

$$\Rightarrow a_i = b_i \;\forall\; i = 1, 2, ..., n$$

$\therefore \alpha = \sum_{i=1}^{n} a_i\alpha_i = \sum_{i=1}^{n} b_i\alpha_i = \beta$ $\quad\therefore T$ is one-one.

T is onto : Let $\beta \in V \Rightarrow \exists a_i \in F \ni \beta = \sum_{i=1}^{n} a_i\beta_i$

Write $\alpha = \sum_{i=1}^{n} a_i\alpha_i \Rightarrow \alpha \in U$ and

$T(\alpha) = \sum_{i=1}^{n} a_i\beta_i = \beta$ $\quad\therefore T$ is onto $\Rightarrow T$ is a bijection.

Hence T is an isomorphism.

$\therefore U \cong V$.

8.6.7. Corollary: If $T: U \to V$ is an isomorphism between finite dimensional vector spaces and B is a basis of U, then T(U) is a basis of V.

8.6.8 Theorem: If $\dim U(F) = n$, then $U \cong F^n$

Proof: Let $S = \{\alpha_1, \alpha_2, ..., \alpha_n\}$ be a basis of U.

Let $\alpha \in U \Rightarrow \exists a_i \in F \ni \alpha = \sum_{i=1}^{n} a_i\alpha_i$

Define $T: U \to F^n$ by $T(\alpha) = (a_1, a_2, ..., a_n) \;\forall\; \alpha = \sum_{i=1}^{n} a_i\alpha_i \in U$.

We show that T is a linear transformation: Let $\alpha, \beta \in U, a, b \in F$

$$\Rightarrow \exists a_i, b_i \in F \ni \alpha = \sum_{i=1}^{n} a_i\alpha_i, \beta = \sum_{i=1}^{n} b_i\alpha_i$$

$\therefore T(a\alpha + b\beta) = T\left(\sum_{i=1}^{n}(aa_i + bb_i)\alpha_1 = (aa_1 + bb_1, aa_1 + bb_2, ..., aa_n + bb_n\right)$

$= a(a_1, a_2, ..., a_n) + b(b_1, b_2, ..., b_n) = aT(\alpha) + bT(\beta) \forall \alpha, \beta \in U$ and $\forall a, b \in F$.

$\therefore T$ is a linear transformation.

T is one-one : Let $\alpha, \beta \in U$ and $T(\alpha) = T(\beta)$, Let $\alpha = \sum_{i=1}^{n} a_i\alpha_i, \beta = \sum_{i=1}^{n} b_i\alpha_i$

$\therefore T(\alpha) = T(\beta) \Rightarrow (a_1, a_2, ..., a_n) = (b_1, b_2, ..., b_n) \Rightarrow a_i = b_i \forall i = 1, 2, ...n$

$\therefore \alpha = \sum_{i=1}^{n} a_i\alpha_i = \sum_{i=1}^{n} b_i\alpha_i = \beta$ $\therefore T$ is one - one.

T is onto : Let $(a_1, a_2, ..., a_n) \in F^n \Rightarrow \alpha = \sum_{i=1}^{n} a_i\alpha_i \in U$ and $T(\alpha) = (a_1, a_2, ..., a_n)$

$\therefore T$ is onto : T is a bijection.

$\therefore T$ is an somorphism.

Hence $U \cong F^n$.

8.6.9 Fundamental Theorem of Homomorphism:

Theorem: Let $U(F)$ and $V(F)$ be vector spaces and $T : U \rightarrow V$ be an onto linear transformation with null space N. Then $\frac{U}{N} \cong V$.

Proof: Since N is the null space of T, we have $N \subseteq U$ and $\frac{U}{N} = \{N + \alpha / \alpha \in U\}$ is the quotient space and $(N + \alpha) + (N + \beta) = N + (\alpha + \beta)$ and $a(N + \alpha) = N + a\alpha \forall N + \alpha, N + \beta \in \frac{U}{N}$ and $a \in F$

Define $\phi : \frac{U}{N} \rightarrow V$

$\phi(N + \alpha) = T(\alpha) \forall N + \alpha \in \frac{U}{N}$

ϕ is well-defined : Let $N+\alpha, N+\alpha^1 \in \frac{U}{N}$ and $N+\alpha = N+\alpha^1$

$$\Rightarrow \alpha-\alpha^1 \in N = \ker T$$

$$\Rightarrow T(\alpha-\alpha^1) = \overline{0} \Rightarrow T(\alpha)-T(\alpha^1) = \overline{0}$$

$$\Rightarrow T(\alpha) = T(\alpha^1) \Rightarrow \phi(N+\alpha) = \phi(N+\alpha^1)$$

$\therefore\ \phi$ is well-defined.

ϕ is a linear transformation : Let $a,b \in F, N+\alpha, N+\beta \in \frac{U}{N}$

$$\therefore \phi\left[a(N+\alpha)+b(N+\beta)\right] = \phi\left[(N+a\alpha)+(N+b\beta)\right] = \phi(N+a\alpha+b\beta)$$

$$= T(a\alpha+b\beta) = aT(\alpha)+bT(\beta)$$

$$= a\phi(N+\alpha)+b\phi(N+\beta)\ \forall\ N+\alpha, N+\beta \in \frac{U}{N} \text{ and } \forall\ a,b \in F.$$

$\therefore \phi$ is linear transformation.

ϕ is one - one : Let $N+\alpha, N+\beta \in \frac{U}{N}$ and $\phi(N+\alpha) = \phi(N+\beta)$

$$\Rightarrow T(\alpha) = T(\beta) \Rightarrow T(\alpha-\beta) = \overline{0}$$

$$\Rightarrow \alpha-\beta \in N \Rightarrow N+\alpha = N+\beta$$

$\therefore \phi$ is one - one.

ϕ is onto : Let $\beta \in V$. Since $T: U \to V$ is onto, $\exists \alpha \in U \ni T(\alpha) = \beta$

$$\therefore \beta = T(\alpha) = \phi(N+\alpha)$$

$$\therefore \beta \in V \Rightarrow \exists N+\alpha \in \frac{U}{N} \ni \phi(N+\alpha) = \beta$$

$\therefore \phi$ is on to.

$\therefore \phi$ is an isomorphism $\quad \therefore \frac{U}{N} \cong V$

8.6.10 Note: Even though T is not onto, the above theorem is true if V is replaced by T(U).

8.6.11 Definition: A linear transformation $T: U(F) \to V(F)$ is said to be singular if $N(T) \neq \{\overline{0}\}$ and it is said to be non-singular if $N(T) = \{\overline{0}\}$.

8.6.12 Theorem: Let $U(F)$, $V(F)$ be two vector spaces and $T: U \to V$ be a linear transformation. then T is non-singular iff "S is $L.I \Rightarrow T(S)$ is $L.I$ ".

Proof: 1. Suppose T is non-singular and suppose $S = \{\alpha_1, \alpha_2, ..., \alpha_n\}$ is $L.I$

Let $S^1 = \{T(\alpha_1), T(\alpha_2), ..., T(\alpha_n)\}$

Let $a_1, a_2, ..., a_n \in F$ and $a_1T(\alpha_1) + a_2T(\alpha_2) + ... + a_nT(\alpha_n) = \overline{0}$

$\Rightarrow T(a_1\alpha_1 + a_2\alpha_2 + ... + a_n\alpha_n) = \overline{0}$

$\Rightarrow a_1\alpha_1 + a_2\alpha_2 + ... + a_n\alpha_n = \overline{0}$, since T is non-singular, $N(T) = \{\overline{0}\}$.

$\Rightarrow a_1 = 0, a_2 = 0, ...a_n = 0$ ($\because$ S is $L.I$)

$\therefore S^1$ is $L.I \Rightarrow T(S)$ is $L.I$

Conversely suppose S is $L.I \Rightarrow T(S)$ is $L.I$.

Let $\alpha \in U$ and $\alpha \neq \overline{0}$ $\therefore S = \{\alpha\}$ is $L.I$

$\Rightarrow T(S) = \{T(\alpha)\}$ is $L.I$

$\Rightarrow T(\alpha) \neq \overline{0}$

$\therefore \alpha \in U, \alpha \neq \overline{0} \Rightarrow T(\alpha) \neq \overline{0}$

$\therefore T$ is non-singular.

8.6.13 Theorem: Let $U(F)$, $V(F)$ be vector spaces and $T: U \to V$ be a linear transformation. Then T is non-singular iff T is one-one.

Proof: Suppose T is non-singular $\Rightarrow N(T) = \{0\}$

Suppose $\alpha, \beta \in U$ and $T(\alpha) = T(\beta) \Rightarrow T(\alpha) - T(\beta) = \overline{0}$

$\Rightarrow T(\alpha - \beta) = \overline{0}$

$\Rightarrow \alpha - \beta = \overline{0}$ $\because N(T) = \{\overline{0}\}$

$\Rightarrow \alpha = \beta$ $\therefore T$ is one - one.

Conversely suppose that T is one-one.

Suppose $T(\alpha) = \overline{0} \Rightarrow T(\alpha) = T\left(\overline{0}\right)$

$\Rightarrow \alpha = \overline{0}$, Since T is one - one.

$\therefore N(T) = \{\overline{0}\}$ $\therefore T$ is non-singular.

Hence the Theorem.

8.6.14 Note: $T : U \to V$ is non-singular if either

$N(T) = \{\overline{0}\}$ (or) $T(\alpha) = \overline{0} \Rightarrow \alpha = \overline{0}$ (or) $\alpha \neq \overline{0} \Rightarrow T(\alpha) \neq \overline{0}$.

Solved Problems

8.6.15 Let $T : \mathbb{R}^3 \to \mathbb{R}^2$ be a linear transformation defined by

$T(x, y, z) = (2x + y - z, 3x - 2y + 4z), \forall (x, y, z) \in \mathbb{R}^3$

Obtain the matrix of T relative to the ordered bases.

$B_1 = \{(1,1,1), (1,1,0), (1,0,0)\}$, $B_2 = \{(1,3), (1,4)\}$

Solution: We have

$T(1,1,1) = (2,5)$

$T(1,1,0) = (3,1)$

$T(1,0,0) = (2,3)$

Let $(a,b) = x(1,3) + y(1,4) = (x + y, 3x + 4y)$

$\therefore x + y = a$ .... (1)

$3x + 4y = b$

(1) x 3 : $3x + 3y = 3a$

Subtracting, we get : $y = b - 3a$

$\therefore x + y - 3a = a \Rightarrow x = 4a - b$

$\therefore (a,b) \in \mathbb{R}^2 \Rightarrow (a,b) = (4a-b)(1,3) + (b-3a)(1,4)$

$\therefore T(1,1,1) = (2,5) = (5-6)(1,3) + (8-5)(1,4)$

$= (3)(1,3) + (-1)(1,4)$

$T(1,1,0) = (3,1) = 11(1,3) + (-8)(1,4)$

$T(1,0,0) = (2,3) = 5(1,3) + (-3)(1,4)$

$$\therefore [T; B_1, B_2] = \begin{bmatrix} 3 & 11 & 5 \\ -1 & -8 & -3 \end{bmatrix}$$

8.6.16 SAQ: Let $T : \mathbb{R}^3 \to \mathbb{R}^2$ be the linear transformation defined by $T(x,y) = (x+y, 2x-y, 7y)$. Find $[T; B_1, B_2]$, where B_1, B_2 are the standard ordered bases of $\mathbb{R}^2$ and $\mathbb{R}^3$ respectively.

8.6.17 SAQ: Let $T : \mathbb{R}^3 \to \mathbb{R}^2$ be the linear transformation defined by

$$T(x,y,z) = (3x + 2y - 4z, x - 5y + 3z)$$

Find $[T; B_1, B_2]$, where $B_1 = \{(1,1,1), (1,1,0), (1,0,0)\}$ and $B_2 = \{(1,3), (2,5)\}$ are the ordered bases of $\mathbb{R}^3$ and $\mathbb{R}^2$ respectively.

Solved Problem:

8.6.18 If A and B are subspaces of a vector space V over a field F, then show that $\frac{A+B}{B} \cong \frac{A}{A \cap B}$.

Solution: We know that A, B are subspaces $\Rightarrow A+B, A \cap B$ are subspacs.

Also $B \subseteq A + B$ $\therefore \frac{A+B}{B}$ is a vector space over F.

Also $A \cap B$ is a subspace of $A \Rightarrow \frac{A}{A \cap B}$ is a vector space over F.

Any element of $\frac{A+B}{B}$ is of the form $B + \alpha + \beta$, where $\alpha \in A$ and $\beta \in B$ i.e. $B + \beta + \alpha = B + \alpha$, since $\beta \in B \Rightarrow B + \beta = B$.

So any element of $\frac{A+B}{B}$ is of the form $B + \alpha$ for some $\alpha \in A$.

Define a mapping $T : A \to \frac{A+B}{B}$ by $T(\alpha) = B + \alpha \;\forall\, \alpha \in A$

Clearly T is well defined ($\because \alpha_1 = \alpha_2 \Rightarrow B + \alpha_1 = B + \alpha_2 \Rightarrow T(\alpha_1) = T(\alpha_2)$)

Let $\alpha, \beta \in A, a, b \in F$

$$\therefore T(a\alpha + b\beta) = B + (a\alpha + b\beta) = (B + a\alpha) + (B + b\beta)$$

$$= a(B + \alpha) + b(B + \beta)$$

$$= aT(\alpha) + bT(\beta)$$

$\forall \alpha, \beta \in A$ and $\forall a, b \in F$

$\therefore T$ is a linear transformation.

Any element of $\frac{A+B}{B}$ is of the form $B + \alpha$ for some $\alpha \in A$

$$\therefore B + \alpha \in \frac{A+B}{B} \Rightarrow \exists \alpha \in A \ni T(\alpha) = B + \alpha$$

$\therefore T$ is onto.

$\therefore$ By the Fundamental Theorem of homomorphisms, we have

$$\frac{A}{\ker T} \cong \frac{A+B}{B}$$

But $\ker T = \{\alpha \in A / T(\alpha) = B\} = \{\alpha \in A / B + \alpha = B\}$

$$= \{\alpha \in A / \alpha \in B\} = A \cap B$$

$$\therefore \frac{A}{A \cap B} \cong \frac{A+B}{B} \text{ (or) } \frac{A+B}{B} \cong \frac{A}{A \cap B}$$

8.7 Answers to SAQ's:

8.6.16 SAQ: $T : \mathbb{R}^2 \to \mathbb{R}^3$, $T(x, y) = (x + y, 2x - y, 7y)$.

$B_1 = \{\alpha_1 = (1,0), \alpha_2 = (0,1)\}$

$B_2 = \{\beta_1 = (1,0,0), \beta_2 = (0,1,0), \beta_3 = (0,0,1)\}$

are the standard bases of $\mathbb{R}^2$ and $\mathbb{R}^3$ respectively.

$T(\alpha_1) = T(1,0) = (1,2,0) = 1(1,0,0) + 2(0,1,0) + 0(0,0,1)$

$T(\alpha_2) = T(0,1) = (1,-1,7) = 1(1,0,0) - 1(0,1,0) + 7(0,0,1)$

$$\therefore [T, B_1, B_2] = \begin{bmatrix} 1 & 1 \\ 2 & -1 \\ 0 & 7 \end{bmatrix}$$

8.6.17 SAQ: $T : \mathbb{R}^3 \to \mathbb{R}^2$, $T(x,y) = (3x+2y-4z, x-5y+3z)$.

$B_1 = \{\alpha_1 = (1,1,1), \alpha_2 = (1,1,0), \alpha_3 = (1,0,0)\}$

$B_2 = \{\beta_1 = (1,3), \beta_2 = (2,5)\}$ are ordered bases of $\mathbb{R}^3$ and $\mathbb{R}^2$ respectively.

$T(\alpha_1) = (1,-1)$

$T(\alpha_2) = (5,-4)$

$T(\alpha_3) = (3,1)$

Let $(a,b) = x\beta_1 + y\beta_2 \Rightarrow (a,b) = x(1,3) + y(2,5)$

$$\Rightarrow \left.\begin{matrix} x+2y=a \\ 3x+5y=b \end{matrix}\right\} \left.\begin{matrix} 3x+5y=b \\ 3x+6y=3a \end{matrix}\right\} \Rightarrow y = 3a-b$$

$$\therefore x = a - 2y = a - 6a + 2b = 2b - 5a$$

$\therefore (a,b) = (2b-5a)\beta_1 + (3a-b)\beta_2$

$\therefore T(\alpha_1) = (1,-1) = -7\beta_1 + 4\beta_2$

$T(\alpha_2) = (5,-4) = -33\beta_1 + 19\beta_2$

$T(\alpha_3) = (3,1) = -13\beta_1 + 8\beta_2$

$$\therefore [T, B_1, B_2] = \begin{bmatrix} -7 & -33 & -13 \\ 4 & 19 & 8 \end{bmatrix}$$

8.8 Summary:

In this chapter, matrix of a linear transformation from a FDVS to a FDVS relative to ordered bases is discussed. The concept of invertibility of isomorphisms of vector spaces is discussed and some problems are discussed.

8.9 Technical Terms:

Matrix representation of a linear transformation, isomorphism, invertibility.

8.10 Exercises:

8.10.1 Let $T:\mathbb{R}^2 \to \mathbb{R}^2$ be defined by $T(x,y)=(4x-2y, 2x+y)$

Find $[T;B]$, where $B=\{(1,1),(-1,0)\}$

8.10.2 Let $T:\mathbb{R}^3 \to \mathbb{R}^3$ be defined by $T(x,y,z)=(2y+z, x-4y, 3x)$

Find $[T:B]$, where $B=\{(1,0,0),(0,1,0),(0,0,1)\}$

8.10.3 If the matrix of T on $\mathbb{R}^2$ relative to the standard ordered basis is $\begin{bmatrix} 2 & -3 \\ 1 & 1 \end{bmatrix}$, find the matrix of T relative to the basis $\{(1,1),(1,-1)\}$.

8.10.4 Find $T:\mathbb{R}^3 \to \mathbb{R}^4$ as a linear transformation whose range is spanned by $(1,-1,2,3)$ and $(2,3,-1,0)$.

8.10.5 Let V be a vector space of 2 x 2 matrices over $\mathbb{R}$.

Let P be the fixed matrix in V, $P=\begin{bmatrix} 1 & -1 \\ -2 & 2 \end{bmatrix}$ and $T:V\to V$ be defined by $T(A)=PA, \forall A\in V$. Find nullity T.

8.11 Answers to exercises:

8.11.1 $\begin{bmatrix} 3 & -2 \\ 1 & 2 \end{bmatrix}$

8.11.2 $\begin{bmatrix} 0 & 2 & 1 \\ 1 & -4 & 0 \\ 3 & 0 & 0 \end{bmatrix}$

8.11.3 $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & -1 \end{bmatrix}$

8.11.4 $T(x,y,z)=(x+2y,-x+3y,2x-y,3x)\ \forall (x,y,z)\in\mathbb{R}^3$

8.11.5 2

8.12 Model Examination Questions:

8.12.1 Explain the concept of a matrix of a linear transformation.

8.12.2 Define the invertibility of a linear operator.

8.12.3 State and prove the fundamental theorem of homomorphisms of vector spaces.

8.12.4 Find the matrix of the linear transformation

$T:\mathbb{R}^3\to\mathbb{R}^2$ defined by $T(x,y,z)=(x+y,2z-x)\ \forall (x,y,z)\in\mathbb{R}^3$

relative to the ordered bases $B_1=\{(1,0,-1),(1,1,1),(1,0,0)\}$ and $B_2=\{(0,1),(1,0)\}$

8.13 Reference Books:

1. Stephen H. Fried berg and others : Linear Algebra, Prentice Hall India Pvt. Ltd., New Delhi.

2. Hoffman and Kunz; Linear Algebra, 2nd Edition, Prentice Hall; New Jersey - 197.

- A. Satyanarayana Murty

LESSON - 9

MATRICES AND DETERMINANTS

9.1 Objective of the Lesson:

In this chapter, we define elementary operations that are used to obtain simple computational methods for determining the rank of a linear transformation and the solution of a system of linear equations. There are two types of elementary matrix operations - row operations and column operations.

9.2. Structure of the Lesson:

This lesson has the following components.

9.3 Introduction:

In this chapter, we discuss the elementary operations and elementary matrices. We also discuss determinants.

9.4 Elementary Matrix Operations and Elementary Matrices:

9.4.1 Definition: Let A be an m x n matrix. Any one of the following three operations on the rows (Columns) of A is called an elementary row (Column) operation:

1. R_{ij} ; Interchange of i^{th} and j^{th} rows.

2. $R_i(k)$: Multiplying every element of i^{th} row by k.

3. $R_{ij}(k)$: Multiplying every element of j^th row by k and adding to the corresponding element of i^th row.

Similarly we have column operations $C_{ij}, C_i(k), C_{ij}(k)$.

9.4.2 Elementary Matrix:

Definition: A Matrix obtained from a unit matrix I_n by subjecting the unit matrix to any one of the elementary transformations is called an elementary matrix.

E_{ij} : Elementary matrix obtained by interchanging i^th and j^th rows (column) in I_n.

$E_i(k)$: Elementary matrix obtained by multiplying every element of i^th row (column) with k in I_n.

$E_{ij}(k)$: Elementary matrix obtained by multiplying every element of j^th row with k and then adding them to the corresponding elements of i^th row in I_n.

Similarly E_{ij}^1, $E_i^1(k)$, $E_{ij}^1(k)$ denote the elementary matrices obtained by applying the corresponding elementary column operations on I.

9.4.3 Note: 1. $|E_{ij}| = -1 = |E_{ij}^1|$

2. $|E_i(k)| = k = |E_i^1(k)|, k \neq 0$

3. $|E_{ij}^1(k)| = 1 = |E_{ij}^1(k)|$

4. Every elmentary matrix is non-singular and hence it is invertible.

9.4.4. Theorem: Every elementary row (column) transformation of a matrix can be obtained by pre-multiplication (post multiplication) with the corresponding elementary matrix.

Proof: First we prove that every elementary row (column) operation of a product C = AB can be effected by subjecting the prefactor A (Post factor B) to the same row (column) operation.

Let A be m x n and B be n x p matrices. The AB is of order m x p.

We can write $A = \begin{bmatrix} R_1 \\ R_2 \\ \vdots \\ R_m \end{bmatrix}_{m\times 1}$, $B = \begin{bmatrix} c_1 & c_2 & \cdots & c_p \end{bmatrix}_{1\times p}$, where $R_1, R_2, ..., R_m$ are rows (of order 1 x n) and $C_1, C_2, ..., C_p$ are columns of B (of order n x 1)

$$\therefore AB = \begin{bmatrix} R_1C_1 & R_1C_2...... & R_1C_p \\ R_2C_1 & R_2C_2...... & R_2C_p \\ \cdots & \cdots & \cdots \\ R_mC_1 & R_mC_2..... & R_mC_p \end{bmatrix}$$

This shows that if the rows $R_1, R_2, ...R_m$ of A are subjected to an elementary row operation, then the rows of AB are subjected to the same elementay row operation.

Similarly, if the columns of B i.e. $C_1, C_2, ...C_p$ are subjected to an elementary column operation, then the columns of AB are subjected to the same elementary column operation.

Now to prove the theorem let A be an m x n matrix and I_m be a unit matrix so that $A = I_m A$. Let R be any elementary row operation applied on A.

Then $RA = R(I_m A) = (RI_m)A = EA,$ where E is the elementary matrix corresponding to the row operation R.

Again let I_n be unit matrix so that $A = AI_n$.

Let C be any elementary column operation applied on A.

Then $CA = C(AI_n) = A(CI_n) = AE,$ where E is the elementary matrix corresponding to the column operation C.

9.4.5 Theorem: $E_{ij}^{-1} = E_{ij}, \left[E_i(k)\right]^{-1} = E_i\left(\frac{1}{k}\right), k \neq 0, \left[E_{ij}(k)\right]^{-1} = E_{ij}(-k)$ if $c \neq 0$.

Proof: a) E_{ij} is the E-matrix (elementary marix) obtained from I_n by applying R_{ij}. If we again apply R_{ij} on E_{ij}, we get I

$$\therefore E_{ij}E_{ij} = I \quad \therefore E_{ij} \text{ is invertible and } \left(E_{ij}\right)^{-1} = E_{ij}.$$

b) $E_i(k)$ is the E - matrix obtained from I by applying $R_i(k)$. If we again apply $R_i\left(\frac{1}{k}\right)$ on E_i, we get I.

$\therefore E_i\left(\frac{1}{k}\right)E_i(k) = I \Rightarrow E_i(k)$ is invertible and $\left[E_{ij}(k)\right]^{-1} = E_i\left(\frac{1}{k}\right)$

c) $E_{ij}(k)$ is the E - matrix obtained from I by applying $R_{ij}(k)$. If we again apply $R_{ij}(-k)$, we get I

$\therefore E_{ij}(-k)E_{ij}(k) = I \Rightarrow E_{ij}(k)$ is invertible and $\left[E_{ij}(k)\right]^{-1} = E_{ij}(-k)$

Similarly, we can show that

$$\left(E_{ij}^{\ 1}\right)^{-1} = E_{ij}^{\ 1}, \left[E_i^1(k)\right]^{-1} = E_i^1\left(\frac{1}{k}\right), k \neq 0$$

$$\left[E_{ij}^{\ 1}(k)\right]^{-1} = E_{ij}^{\ 1}(-k)$$

We note that every E - matrix is nonsingular and the inverse of an E - matrix is also non-singular.

9.4.6 Definition: A matrix A is said to be equivalent to B, if B is obtained from A by a finite number of E - operations.

We write $A \sim B$.

In the set M of all m x n matrices, $\sim$ is an equivalence relation.

Since a) $A \sim A \Rightarrow \sim$ is reflexive

b) $A \sim B \Rightarrow B \sim A$ so that $\sim$ is symmetric.

c) $A \sim B, B \sim C \Rightarrow A \sim C$ so that $\sim$ is transitive.

9.4.7 Definition: A matrix A is said to be row (column) equivalent to B, denoted by $A \overset{R}{\sim} B \left(A \overset{C}{\sim} B\right)$, if it is possible to obtain B from A by a finite number of E - row (column) operations.

Solved Problems:

9.4.8 Compute the matrixes $E_{23}, E_{34}(-1), E_2(-2), E_{12}$ for I_4.

Solution $E_{23} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} i, \left(\because I_4 \overset{R_{23}}{\sim} E_{23}\right)$

$$I_4 \overset{R_{34}(-1)}{\sim} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix} = E_{34}(-1)$$

$$I_4 \overset{R_2(-2)}{\sim} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = E_2(-2)$$

$$I_4 \overset{R_{12}}{\sim} \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = E_{12}$$

9.4.9 Show that $\begin{bmatrix} 1 & 2 & 1 \\ -1 & 0 & 2 \\ 2 & 1 & -3 \end{bmatrix} \sim I_3$

Solution: Let $A = \begin{bmatrix} 1 & 2 & 1 \\ -1 & 0 & 2 \\ 2 & 1 & -3 \end{bmatrix} \underset{R_{31}(-2)}{\overset{R_{21}(1)}{\sim}} \begin{bmatrix} 1 & 2 & 1 \\ 0 & 2 & 3 \\ 0 & -3 & -5 \end{bmatrix}$

$$\underset{\substack{R_2(1/2) \\ R_3(-1/3)}}{\overset{R_{12}(-1)}{\sim}} \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 3/2 \\ 2 & 1 & 5/3 \end{bmatrix} \overset{R_{32}(-1)}{\sim} \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 3/2 \\ 0 & 0 & +1/6 \end{bmatrix}$$

$$\overset{R_3(6)}{\sim} \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 3/2 \\ 0 & 0 & 1 \end{bmatrix} \underset{R_{23}(-3/2)}{\overset{R_{13}(2)}{\sim}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I_3$$

$\therefore\ A \sim I_3$

9.4.10 Express $\begin{bmatrix} 1 & 2 & 3 \\ 1 & 0 & 1 \\ 1 & -1 & 1 \end{bmatrix}$ as a product of E - matrices.

Solution: Let $A = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 0 & 1 \\ 1 & -1 & 1 \end{bmatrix} \underset{R_{31}(-1)}{\overset{R_{21}(-1)}{\sim}} \begin{bmatrix} 1 & 2 & 3 \\ 0 & -2 & -2 \\ 0 & -3 & -2 \end{bmatrix} \underset{R_3(-\frac{1}{3})}{\overset{R_2(-\frac{1}{2})}{\sim}} \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 1 \\ 0 & 0 & \frac{2}{3} \end{bmatrix} \underset{R_{32}(-1)}{\overset{R_{12}(-2)}{\sim}} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & -\frac{1}{3} \end{bmatrix} \overset{R_3(-3)}{\sim}$

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \underset{C_{32}(-1)}{\overset{C_{31}(-1)}{\sim}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I_3$$

$$\therefore\ A \underset{\substack{R_3(-\frac{1}{3}) \\ R_{12}(-2), R_{32}(-1), R_3(-3) \\ C_{31}(-1), C_{32}(-1)}}{\overset{R_2(-\frac{1}{2})}{\sim}} I_3$$

$$\therefore I_3 = E_3(-3)E_{32}(-1)E_{12}(-2)E_3(-\tfrac{1}{3})E_2(-\tfrac{1}{2})E_{31}(-1)E_{21}(-1)AE^1_{32}(-1)E^1_{31}(-1)$$

$$A = [E_{21}(-1)]^{-1}[E_{31}(-1)]^{-1}[E_2(-\tfrac{1}{2})]^{-1}[E_3(-\tfrac{1}{3})]^{-1}[E_{12}(-2)]^{-1}[E_{32}(-1)]^{-1}[E_3(-3)]^{1} I_3 [E^1_{32}(-1)]^{-1}[E^1_{31}(-1)]^{-1}$$

$$= E_{21}(1)E_{31}(1)E_2(2)E_3(3)E_{12}(2)E_{32}(1)E_3(-\tfrac{1}{3})E^1_{32}(1)E^1_{31}(1)$$

9.4.11 Express $A = \begin{bmatrix} 1 & 3 & 3 \\ 1 & 4 & 3 \\ 1 & 3 & 4 \end{bmatrix}$ as a product of E - matrices.

Solution: $A = \underset{R_{31}(-1)}{\overset{R_{21}(-1)}{\sim}} \begin{bmatrix} 1 & 3 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \underset{C_{31}(-3)}{\overset{C_{21}(-3)}{\sim}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I_3$

$$I_3 = E_{31}(-1)E_{21}(-1)AE_{21}^1(-3)E_{31}^1(-3)$$

$$\Rightarrow A = [E_{21}(-1)]^{-1}[E_{31}(-1)]_{I_3}^{-1}[E_{31}^1(-3)]^{-1}[E_{21}^1(-3)]^{-1}$$

$$\Rightarrow A = E_{21}(1)E_{31}(1)E_{31}(3)E_{21}(3)$$

$\therefore A$ is a product of the E - matrices.

9.5 Determinants:

Determinants of matrices of order 2 x 2, 3 x 3 were studied at the Intermediate Level. Here we attempt to define the determinant of a square matrix of order n x n and discuss the properties although we deal with determinants of order 2 x 2, 3 x 3.

9.5.1 Definition: Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ be a 2 x 2 matrix over F.

We define the determinant of A, denoted by $\det A$ or $|A|$ as the scalar *ad - bc.*

If $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, we define adjoint of A as adj $A = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$.

9.5.2 Note: We observe that $\det(A+B) \neq \det A + \det B$, since

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, B = \begin{bmatrix} 3 & 2 \\ 6 & 4 \end{bmatrix} \Rightarrow A+B = \begin{bmatrix} 4 & 4 \\ 9 & 8 \end{bmatrix}$$

and $\det A = -2, \det B = 0$, $\det(A+B) = -4$

9.5.3 Note: The function $\det : F^{2\times 2} \to F$ is not a linear transformation.

9.5.4 Theorem: The function $\det : F^{2\times 2} \to F$ is a linear function of each row of a 2 x 2 matrix A when the second row is fixed i.e. if $u, v, w \in F^2$ and $k \in F$, then

$$\det \begin{bmatrix} u + kv \\ w \end{bmatrix} = \det \begin{bmatrix} u \\ w \end{bmatrix} + k \det \begin{bmatrix} v \\ w \end{bmatrix}$$

Proof: Let $u = (a_1, b_1), v = (a_2, b_2), w = (a_3, b_3) \in F^2$ and $k \in F$, then

$$\det \begin{bmatrix} u \\ w \end{bmatrix} + k \det \begin{bmatrix} v \\ w \end{bmatrix} = \det \begin{bmatrix} a_1 & b_1 \\ a_3 & b_3 \end{bmatrix} + k \det \begin{bmatrix} a_2 & b_2 \\ a_3 & b_3 \end{bmatrix}$$

$$= (a_1b_3 - a_3b_1) + k(a_2b_3 - a_3b_2) = (a_1 + ka_2)b_3 - (b_1 + kb_2)a_3$$

$$= \det\begin{bmatrix} a_1 + ka_2 & b_1 + kb_2 \\ a_3 & b_3 \end{bmatrix} = \det\begin{bmatrix} u + kv \\ w \end{bmatrix}$$

Similarly we can show that $\det\begin{bmatrix} w \\ u \end{bmatrix} + k\det\begin{bmatrix} w \\ v \end{bmatrix} = \det\begin{bmatrix} w \\ u + kv \end{bmatrix}$

9.5.5. Theorem: Let $A \in F^{2\times 2}$. Then $\det A \neq 0 \Leftrightarrow A$ is invertible.

(Recall that A is invertible iff $\exists\, B \ni AB = BA = I$)

If $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ then $A^{-1} = \dfrac{1}{\det A}\begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}$

Proof: Suppose $\det A \neq 0$. Let $C = \dfrac{1}{\det A}\begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}$

$$\therefore AC = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \frac{1}{\det A}.\begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}$$

$$= \frac{1}{\det A}\begin{bmatrix} a_{11}a_{22} - a_{12}a_{21} & -a_{11}a_{12} + a_{11}a_{22} \\ a_{21}a_{22} - a_{21}a_{22} & -a_{21}a_{12} + a_{22}a_{11} \end{bmatrix}$$

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

Similarly we can show that $CA = I$. So A is invertible and $A^{-1} = C$.

Conversely suppose A is invertible. So rank of A is = 2. (For definition of rank, see lesson 10). (Since rank of an invertible n x n matrix is n)

$\therefore a_{11} \neq 0$ or $a_{21} \neq 0$. Suppose $a_{11} \neq 0$. Multiply R_1 with $\dfrac{-a_{21}}{a_{11}}$ and add to R_2 so that we get :

$$\begin{bmatrix} a_{11} & a_{12} \\ 0 & a_{22} - \dfrac{a_{12}a_{21}}{a_{11}} \end{bmatrix}$$

The rank of this is = 2 so that $a_{22} - \frac{a_{12}a_{21}}{a_{11}} \neq 0$.

$\Rightarrow a_{11}a_{22} - a_{12}a_{21} \neq 0 \Rightarrow \det A \neq 0$

Similarly we can prove that $\det\ A \neq 0$ when $a_{21} \neq 0$.

So in any case, we have $\det\ A \neq 0$. Hence the Theorem.

Now we define determinants of order n.

9.5.6 Definition: Let $A \in F^{n\times n}$

If $n = 1$, then $A = \left[a_{ij}\right]$, we define $\det\ A = a_{11}$.

If $n \geq 2$, we define det A recursively as follows:

$$\det A = \sum_{j=1}^{n} (-1)^{1+j}{}_{a_{ij}} . \det \overline{A}_{ij} .$$

(This is called the determinant of A and is denoted by $|A|$).

Here $\overline{A}_{ij}$ is the $(n-1)\times(n-1)$ matrix obtained from A by deleting i[th] row and j[th] column of A

If we write $c_{ij} = (-1)^{1+j}\ \det \overline{A}_{ij}$, then

$$\det\ A = a_{11}c_{11} + a_{12}c_{12} + ... + a_{1n}c_{1n}$$

c_{ij} is called the cofactor of a_{ij}.

9.5.7 Note: det A = Sum of the products of each entry (in R_1 of A) and the corresponding cofactor.

9.5.8 Example: Find det A using cofactor expansion along the first row of the matrix

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

Solution: $\det A = (-1)^{1+1} A_{11}\left|\overline{A}_{11}\right| + (-1)^{1+2} A_{12}\left|\overline{A}_{12}\right| + (-1)^{1+3} A_{13}\left|\overline{A}_{13}\right|$

$$= (-1)^2 .1(30-8) + (-1)^3 .3(18+8) + (-1)^4 (-3)(-12-20)$$

$$= 22 - 78 + 96 = 40$$

9.5.9 Theorem: $|I_n| = 1$

Proof: (by induction on n)

If $n = 1$, then $I_1 = [1] \Rightarrow |I_1| = 1$. So the theorem is true when $n = 1$.

Assume the truth of the theorem for $n-1$.

Expanding I_n along the first row, we get $|I_n| = 1\det\ I_{n-1} + 0 + 0 ... + 0 = 1(1) = 1$. The theorem is true for n. So by Mathematical induction, the Theorem is true for all positive integers n.

Hence the theorem.

The determinant of a square matrix can be evaluated by cofactor expansion along any row i.e. if $A \in F^{n\times n}$, then $\det A = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det \overline{A}_{ij}$ where $\overline{A}_{ij}$ is the $(n-1)\times(n-1)$ matrix obtained by deleting R_i and C_j from A.

9.5.10 Theorem: If $A \in F^{n\times n}$ and B is a matrix obtained from A by interchanging two rows of A, then $\det B = -\det A$.

Proof: Let $A \in F^{n\times n}$ and $a_1, a_2, ..., a_n$ be the rows of A so that $A = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$

Let B be the matrix obtained from A by interchanging rth row and sth row, r < s.

$\therefore B = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_s \\ \vdots \\ a_r \\ \vdots \\ a_n \end{bmatrix}$. Now consider the matrix whose rth and sth rows are replaced by $a_r + a_s$.

$$\therefore \text{ we have } \det\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_r + a_s \\ \vdots \\ a_r + a_s \\ \vdots \\ a_n \end{bmatrix} = \det\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_r \\ \vdots \\ a_r + a_s \\ \vdots \\ a_n \end{bmatrix} + \det\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_s \\ \vdots \\ a_r + a_s \\ \vdots \\ a_n \end{bmatrix}$$

$$\Rightarrow 0 = \det\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_r \\ \vdots \\ a_r \\ \vdots \\ a_n \end{bmatrix} + \det\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_r \\ \vdots \\ a_s \\ \vdots \\ a_n \end{bmatrix} + \det\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_s \\ \vdots \\ a_r \\ \vdots \\ a_n \end{bmatrix} + \det\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_s \\ \vdots \\ a_s \\ \vdots \\ a_n \end{bmatrix} = 0 + \det A + \det B + 0$$

$\Rightarrow \det B = -\det A$

9.5.11 Theorem: If $A, B \in F^{n\times n}$, then $\det AB = \det A \det B$.

Proof: First we prove the theorem when A is an elementary matrix. If A is a matrix obtained from I by interchanging two rows of I, then $\det A = -1$.

By Theorem (9.4.4), AB is a matrix obtained by interchanging two rows of B. So, by Theorem (9.5.10), $\det AB = -\det B = \det A \det B$

Similarly we can prove the theorem, when A is an elementary matrix of other types.

If A is an $n\times n$ matrix with rank < n, then $\det A = 0$.

Since rank $AB \le$ rank $A \le n$, we have $\det AB = 0$

$\therefore \det AB = 0 = \det A . \det B$

If rank $A = n$, then A is invertible and hence it is a product of elementary matrices.

Let $A = E_m E_{m-1} ... E_2 E_1$

$$\det AB = \det(E_m E_{m-1} ... E_2 E_1 B)$$

$$= \det E_m . \det(E_{m-1} ... E_2 E_1 B)$$

$$= \det E_m . \det(E_{m-1} ... E_2 E_1 B)$$

= ....

$$= \det E_m \det E_{m-1} ...... \det E_2 \det E_1 . \det B$$

$$= \det(E_m E_{m-1} ... E_2 E_1) \det B$$

$$= \det A \det B$$

9.5.12 Theorem: A matrix $A \in F^{n \times n}$ is invertible iff $\det A \neq 0$.

Further A is invertible $\Rightarrow \det A^{-1} = \dfrac{1}{\det A}$

Proof: Suppose A is not invertible $\Rightarrow$ Rank of A is $< n \Rightarrow \det A = 0$

Suppose A is invertible $\Rightarrow A^{-1}$ exists and $AA^{-1} = I$

$$\therefore 1 = \det I = \det AA^{-1} = \det A \det A^{-1}$$

$\therefore \det A \det A^{-1} = 1 \therefore \det A \neq 0$ and $\det A^{-1} = \dfrac{1}{\det A}$

9.5.13 Theorem: Let $A \in F^{n \times n}$ and let B be a matrix obtained by adding a multiple of one row to another row of A. Then $\det B = \det A$.

Proof: Suppose B is the n x n matrix obtained from A by adding k times r[th] row to s[th] row where $r \neq s$.

Let $A = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}, B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$, then $b_i = a_i$ for $i \neq s$ and $b_s = a_s + ka_r$

Let C be the matrix obtained from A by replacing a_s with a_r.

$\therefore$ By theorem (9.5.4), we get

$\det B = \det A + k \det C = \det A,$ since $\det C = 0$.

9.5.14 Theorem: If $A \in F^{n\times n}$ and $\rho(A) < n$, then $\det A = 0$

Proof: If $\rho(A) < n$, then the rows $a_1, a_2, ..., a_n$ of A are linearly dependent. So $\exists a_r \ni a_r = al.c$ of other rows.

$\therefore \exists c_1, c_2, ...c_{r-1}, ..., c_n \ni a_r = c_1a_1 + c_2a_2 + ... + c_{r-1}a_{r-1} + c_{r+1}a_{r+1} + ... + c_na_n$

Let B be the matrix obtained from A by adding $-c_i$ times row i to row r for each $i \neq r$

The r[th] row of B consists of zeros only and so $\det B = 0$.

But by theorem (9.5.14) $\det B = \det A$. Hence $\det A = 0$.

9.5.15 Note: If $A \in F^{n\times n}$, then $\det kA = k^n \det A, \det(-A) = (-1)^n \det A$ and det $A = 0$, if two rows are identical.

9.5.16 Theorem: For any $A \in F^{n\times n}, \det A^T = \det A$

Proof: If A is not invertible, then $\rho(A) < n$ and $\det A = 0$

But we know that $\rho(A) = \rho(A^T) \Rightarrow \rho(A^T) < n$

$\therefore A^T$ is not invertible.

$\therefore \det A^T = 0$

$\therefore \det A^T = \det A = 0$

Suppose A is invertible. Then A is a product of elementary matrices.

Suppose $A = E_mE_{m-1...}E_2E_1$

$\therefore \det A^T = \det(E_mE_{m-1...}E_2E_1)^T = \det(E_1^TE_2^T \dots E_m^T)$

$= \det(E_1^T)\det(E_2^T)\dots\det(E_m^T)$

$= \det E_1 \det E_2 \dots \det E_m$

$= (\det E_m)(\det E_{m-1})\dots(\det E_2)(\det E_1)$

$$= \det(E_m E_{m-1} \dots E_1)$$

$$= \det A$$

Hence the Theorem.

Solved Problem:

9.5.17 Evaluate the determinant of the matrix $A = \begin{bmatrix} 2 & 0 & 0 & 1 \\ 0 & 1 & 3 & -3 \\ -2 & -3 & -5 & 2 \\ 4 & -4 & 4 & -6 \end{bmatrix}$

Solution:

$$A \underset{R_{41}(-2)}{\overset{R_{31}(1)}{\sim}} \begin{bmatrix} 2 & 0 & 0 & 1 \\ 0 & 1 & 3 & -3 \\ 0 & -3 & -5 & 3 \\ 0 & -4 & 4 & -8 \end{bmatrix} \underset{R_{42}(4)}{\overset{R_{32}(3)}{\sim}} \begin{bmatrix} 2 & 0 & 0 & 1 \\ 0 & 1 & 3 & -3 \\ 0 & 0 & +4 & -6 \\ 0 & 0 & 16 & -20 \end{bmatrix}$$

$$\therefore \overset{R_{43}(-4)}{\sim} \begin{bmatrix} 2 & 0 & 0 & 1 \\ 0 & 1 & 3 & -3 \\ 0 & 0 & 4 & -6 \\ 0 & 0 & 0 & 4 \end{bmatrix} = B(say)$$

B is an upper triangular matrix.

$\therefore \det A =$ Product of diagonal elements $= 2(1)(4)(4) = 32$

9.5.18 SAQ: If w is a complex cube root of unity, then show that $\begin{vmatrix} 1 & w & w^2 \\ w & w^2 & 1 \\ w^2 & 1 & w \end{vmatrix} = 0$

9.5.19 SAQ: Show that $\begin{vmatrix} 1^2 & 2^2 & 3^2 & 4^2 \\ 2^2 & 3^2 & 4^2 & 5^2 \\ 3^2 & 4^2 & 5^2 & 6^2 \\ 4^2 & 5^2 & 6^2 & 7^2 \end{vmatrix} = 0$

9.5.20 SAQ: Show that $\begin{vmatrix} b-c & c-a & a-b \\ c-a & a-b & b-c \\ a-b & b-c & c-a \end{vmatrix} = 0$

Solved Problems:

9.5.21 Show that $\begin{vmatrix} 1+a & 1 & 1 \\ 1 & 1+b & 1 \\ 1 & 1 & 1+c \end{vmatrix} = abc\left[1+\frac{1}{a}+\frac{1}{b}+\frac{1}{c}\right]$

Solution: $LHS = abc\begin{vmatrix} 1+\frac{1}{a} & \frac{1}{a} & \frac{1}{a} \\ \frac{1}{b} & 1+\frac{1}{b} & \frac{1}{b} \\ \frac{1}{c} & \frac{1}{c} & 1+\frac{1}{c} \end{vmatrix}$

$$\underset{R_{13}(1)}{\overset{R_{12}(1)}{\sim}} abc\begin{vmatrix} 1+\frac{1}{a}+\frac{1}{b}+\frac{1}{c} & 1+\frac{1}{a}+\frac{1}{b}+\frac{1}{c} & 1+\frac{1}{a}+\frac{1}{b}+\frac{1}{c} \\ \frac{1}{b} & 1+\frac{1}{b} & \frac{1}{b} \\ \frac{1}{c} & \frac{1}{c} & 1+\frac{1}{c} \end{vmatrix}$$

$$= abc\left[1+\frac{1}{a}+\frac{1}{b}+\frac{1}{c}\right]\begin{vmatrix} 1 & 1 & 1 \\ \frac{1}{b} & 1+\frac{1}{b} & \frac{1}{b} \\ \frac{1}{c} & \frac{1}{c} & 1+\frac{1}{c} \end{vmatrix} \underset{C_{31}(-1)}{\overset{C_{21}(-1)}{\sim}} = abc\left[1+\frac{1}{a}+\frac{1}{b}+\frac{1}{c}\right]\begin{vmatrix} 1 & 0 & 0 \\ \frac{1}{b} & 1 & 0 \\ \frac{1}{c} & 0 & 1 \end{vmatrix}$$

$$\underset{C_{31}(-1)}{\overset{C_{21}(-1)}{\sim}} \begin{vmatrix} 1 & 0 & 0 \\ \frac{1}{b} & 1 & 0 \\ \frac{1}{c} & 0 & 1 \end{vmatrix} \Rightarrow$$ Expanding with R_1, we get

$\therefore \det$ of given matrix is = $abc\left[1+\frac{1}{a}+\frac{1}{b}+\frac{1}{c}\right]$ .1

9.5.22 Show that $\begin{vmatrix} b+c & c+a & a+b \\ c+a & a+b & b+c \\ a+b & b+c & c+a \end{vmatrix} = 2\begin{vmatrix} a & b & c \\ b & c & a \\ c & a & b \end{vmatrix}$

Solution: Applying $C_{12}(1)$ and $C_{13}(1)$, we get

$$LHS = \begin{vmatrix} 2(a+b+c) & c+a & a+b \\ 2(a+b+c) & a+b & b+c \\ 2(a+b+c) & b+c & c+a \end{vmatrix}$$

$$= 2\begin{vmatrix} a+b+c & c+a & a+b \\ a+b+c & a+b & b+c \\ a+b+c & b+c & c+a \end{vmatrix}$$

$$\underset{C_{31}(-1)}{\overset{C_{21}(-1)}{\sim}} 2\begin{vmatrix} a+b+c & -b & -c \\ a+b+c & -c & -a \\ a+b+c & -a & -b \end{vmatrix}$$

$$\overset{C_{12}(1)}{\sim} 2\begin{vmatrix} a+c & -b & -c \\ a+b & -c & -a \\ b+c & -a & -b \end{vmatrix}$$

$$\overset{C_{13}(1)}{\sim} 2\begin{vmatrix} a & -b & -c \\ b & -c & -a \\ c & -a & -b \end{vmatrix}$$

$$\underset{C_3(-1)}{\overset{C_2(-1)}{\sim}} 2\begin{vmatrix} a & b & c \\ b & c & a \\ c & a & b \end{vmatrix}$$

9.6 Answers to SAQ's:

9.5.18 SAQ: Applying $C_{12}(1)$, we get $\begin{vmatrix} 1+w & w & w^2 \\ w+w^2 & w^2 & 1 \\ w^2+1 & 1 & w \end{vmatrix}$

$$\overset{C_{13}(1)}{\sim} \begin{vmatrix} 1+w+w^2 & w & w^2 \\ 1+w+w^2 & w^2 & 1 \\ 1+w+w^2 & 1 & w \end{vmatrix} = \begin{vmatrix} 0 & w & w^2 \\ 0 & w^2 & 1 \\ 0 & 1 & w \end{vmatrix} = 0$$

since $1+w+w^2=0, w$ being cube root of unity.

9.5.19 SAQ: Applying $C_{32}(-1), C_{41}(-1)$, we get

$$\begin{vmatrix} 1^2 & 2^2 & 5.1 & 5.3 \\ 2^2 & 3^2 & 7.1 & 7.3 \\ 3^2 & 4^2 & 9.1 & 9.3 \\ 4^2 & 5^2 & 11.1 & 11.3 \end{vmatrix} \sim 3\begin{vmatrix} 1^2 & 2^2 & 5 & 5 \\ 2^2 & 3^2 & 7 & 7 \\ 3^2 & 4^2 & 9 & 9 \\ 4^2 & 5^2 & 11 & 11 \end{vmatrix} = 0$$

9.5.20 SAQ: Applying $C_{12}(1)$, we get $\begin{vmatrix} b-a & c-a & a-b \\ c-b & a-b & b-c \\ a-c & b-c & c-a \end{vmatrix}$

$$\overset{C_{13}(1)}{\sim}\begin{vmatrix} 0 & c-a & a-b \\ 0 & a-b & b-c \\ 0 & b-c & c-a \end{vmatrix} = 0$$

9.7 Summary:

In this lesson, we discussed elementary transformations and applied the techniques in determinants.

9.8 Technical Terms:

Elementary row operations, Elementary Matrix, Determinants.

9.9 Exercises:

1. If $\begin{vmatrix} x & x^2 & 1+x^3 \\ y & y^2 & 1+y^3 \\ z & z^2 & 1+z^3 \end{vmatrix} = 0$ and x, y, z are different,

show that $xyz = -1$.

2. Show that $\begin{vmatrix} 1 & x & x^2 \\ 1 & y & y^2 \\ 1 & z & z^2 \end{vmatrix} = (x-y)(y-z)(z-x)$

3. Show that $\begin{vmatrix} 0 & c & b \\ c & 0 & a \\ b & a & 0 \end{vmatrix}^2 = \begin{vmatrix} b^2+c^2 & ab & ac \\ ab & c^2+a^2 & bc \\ ac & bc & a^2+b^2 \end{vmatrix}$

4. Show that $\begin{vmatrix} 2bc-a^2 & c^2 & b^2 \\ c^2 & 2ac-b^2 & a^2 \\ b^2 & a^2 & 2ab-c^2 \end{vmatrix} = \begin{vmatrix} a & b & c \\ b & c & a \\ c & a & b \end{vmatrix}^2 = (a^3+b^3+c^3-3abc)^2$

5. Find the value of the determinant $\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix}$

6. Evaluate $\begin{vmatrix} a+b+2c & a & b \\ c & b+c+2a & b \\ c & a & c+a+2b \end{vmatrix}$

9.10 Answers to Exercises:

5. 0

6. $2(a+b+c)^3$

9.11 Model Examination Questions:

1. Explain the concept of elementary row/column operations.

2. Explain the concept of determinants of order $2\times 2, n\times n$

3. Show that $\begin{vmatrix} 1+a & 1 & 1 & 1 \\ 1 & 1+b & 1 & 1 \\ 1 & 1 & 1+c & 1 \\ 1 & 1 & 1 & 1+d \end{vmatrix} = abcd\left[1+\frac{1}{a}+\frac{1}{b}+\frac{1}{c}+\frac{1}{d}\right]$

4. Show that $\begin{vmatrix} a+x & a & a & a \\ b & b+y & b & b \\ c & c & c+z & c \\ d & d & d & d+w \end{vmatrix} = xyzw\left[1+\frac{a}{x}+\frac{b}{y}+\frac{c}{z}+\frac{d}{w}\right]$

9.12 Reference Books:

1. Stephen H. Friedberg and others - Linear Algebra Prentice Hall India Pvt. Ltd - New Delhi.

2. K. Hoffman and Kunze - Linear Algebra - Prentice Hall, New Jersey.

- A. Satyanarayana Murty

LESSON - 10

RANK OF A MATRIX

10.1 Objective of the Lesson:

In this lesson, we define the rank of a matrix and use elementary operations to compute the rank of a matrix. We also discuss the procedure for computing the inverse of an invertible matrix.

10.2. Structure of the Lesson:

This lesson has the following components.

10.3 Introduction:

In this lesson, the concept of a rank of the matrix is introduced. We use elementary operations to compute the rank of a matrix and the rank of a linear transformation. We also introduce a procedure for computing the inverse of an invertible matrix using elementary transformations.

10.4 Rank of a Matrix:

10.4.1 Definition: Let $A \in F^{m \times n}$. We define the rank of A as the rank of the linear transformation

$L_A : F^n \to F^m$.

We recall the definition of $L_A : F^n \to F^m$ as:

$L_A(X) = AX, \forall X \in F^n$

We observe that many results about the rank of a matrix can be obtained from the corresponding results about linear transformation.

We know that every matrix A is the matrix representation of the linear transformation L_A w.r.t. appropriate standard ordered bases.

i.e. if $A \in F^{m\times n}$, then $\exists$ standard bases B_1, B_2 of F^n and F^m respectively

such that $A = [L_A; B_1, B_2]$.

We also observe that $L_A = L_B \Leftrightarrow A = B, L_{A+B} = L_A + L_B, L_{kA} = kL_A$

for all $k \in F, A \in F^{m\times n}, B \in F^{n\times p} \Rightarrow L_{AB} = L_A L_B; L_{I_n} = I_{F^n}$

We also know that : if $T \in L(U,V), U(F), V(F)$ are finite dimensional vector spaces and B_1, B_2 are ordered bases of U and V respectively, then

Rank $T = Rank[T; B_1, B_2]$

So the problem of finding rank of a linear transformation is reduced to that of finding rank of a matrix.

Now we prove

10.4.2 Theorem: Let A be an m x n matrix. If P and Q are invertible $m\times m, n\times n$ matrices respectively, then

(a) Rank AQ = Rank A, (b) Rank PA = Rank A, (c) Rank PAQ = Rank A.

Proof: We have $L_A : F^n \to F^m, L_Q : F^n \to F^n$

Since Q is non-singular, we observe that L_Q is onto.

($\because y \in F^n$ (codomain) $\Rightarrow \exists X = Q^{-1}y \ni L_Q(X) = QQ^{-1}y = y$)

Now $R(L_{AQ}) = R(L_A L_Q) = L_A L_Q(F^n) = L_A(F^n)$ ($\because L_Q$ is onto)

= $R(L_A)$ (Here $R(L_A)$ means range of L_A)

$\therefore$ Rank L_{AQ} = $\dim R(L_{AQ})$

$= \dim R(L_A)$

= Rank L_A

Similarly we can prove that rank PA = rank A,

(Since Rank PA = Rank $(PA)^T$ = Rank $A^T P^T$ = Rank A^T) (by (a))

= Rank A, by Theorem 10.4.17

Rank PAQ = Rank (PA)Q = Rank PA = Rank A.

10.4.3 Corollary: Elementary row (column) operations on a matrix are rank preserving.

Proof: Suppose B is a matrix obtained from A by an elementary row operation. So $\exists$ an elementary matrix E such that B = EA.

Since E is invertible, we have $\rho(B) = \rho(A)$

10.4.4 Note: Pre-multiplication (Post multiplication) of a matrix by an elementary matrix and hence by a finite number of elementary matrices (by a finite number of elementary operations) does not alter the rank of the matrix.

10.4.5 Theorem: The rank of a matrix is equal to the maximum number of its linearly independent columns i.e. the rank of a matrix is the dimension of the subspace generated by its columns.

Proof: Let $A \in F^{m \times n} \Rightarrow$ Rank A = Rank $L_A = \dim(R(L_A))$

Let B be the standard basis of F^n

$\therefore B$ spans F^n.

$\therefore R(L_A) =$ Span $L_A(B) =$ Span $\{L_A(e_1), L_A(e_2), \ldots . L_A(e_n)\}$

But we know that $L_A(e_j) = Ae_j = (a_1 a_2 \ldots a_j \ldots a_n) \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$

$= a_j$, the j^{th} column of A.

$\therefore R(L_A) = \text{Span } \{a_1, a_2 \ldots a_n\}$

$\therefore$ Rank A = Rank $L_A = \dim R(L_A)$

$= \dim$ Span $\{a_1, a_2 \ldots a_n\}$

10.4.6 SAQ: Find the rank of the matrix $A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}$

10.4.7 SAQ : Find the rank of $A = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 0 & 3 \\ 1 & 1 & 2 \end{bmatrix}$

Reduction to Normal Form:

10.4.8 Theorem: Every non-zero matrix A can be reduced to the form $\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$ by a finite number of elementary operations, where I_r is the unit matrix of order r and r is the rank of A.

Proof: Let $A = \left[a_{ij}\right]_{m \times n}, \rho(A) = r; A \neq O$

Since $A \neq O$, $\exists$ atleast one element $a_{ij} = k \neq 0$

Interchanging R_i with R_1 and C_j with C_1, we obtain a matrix B with leading element $k(\neq 0)$.

Now multiplying R_1 of B with $\frac{1}{k}$, we get a matrix C with leading element 1 so that

$$C = \begin{bmatrix} 1 & c_{12} & c_{13} \ldots\ldots c_{1n} \\ c_{21} & c_{22} & c_{23} \ldots\ldots c_{2n} \\ \cdots & \cdots & \cdots \\ c_{n1} & c_{n2} & c_{n3} \ldots\ldots c_{nn} \end{bmatrix}$$

Adding suitable multiples of the column C_1 of C to the other columns of C and adding suitable multiples of the first row to the remaining rows of C, we get a matrix D in which all elements of R_1 and C_1 of D, except the leading element (=1) are zeros.

$$\therefore D = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & d_{22} & d_{23} & \dots & d_{2n} \\ \vdots & \vdots & & & \\ 0 & d_{m2} & d_{m3} & \dots & d_{mn} \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{bmatrix}_{m\times n}$$, where A_1 is a $(m-1)\times(n-1)$ matrix.

If $A_1 = O$, then $A \sim \begin{bmatrix} I_1 & O \\ O & O \end{bmatrix}$ and is this case, there is nothing to prove.

If $A_1 \neq O$, we proceed with A_1 as we did with A.

The elementary operations applied on A_1 do not alter the elements of either 1st row or 1st column of D.

Proceedings like this, we get a matrix $P \ni P = \begin{bmatrix} I_P & O \\ O & O \end{bmatrix}$

Now $\rho(P) = p$. But P is obtained from A by elementary operations and hence $\rho(A)$ is unaltered.

$\therefore p = r$. Hence A can be reduced to the form $\begin{bmatrix} I_r & O \\ O & O \end{bmatrix}$

10.4.9 Note: 1. Using elementary operations, a matrix A of rank r can be reduced to the form $I_r, [I_r 0], \begin{bmatrix} I_r \\ 0 \end{bmatrix} \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$, called its normal form.

10.4.10 Note: 1) To reduce A to normal form, sometimes, both row operations and column operations are to be applied.

2) If $\rho(A) = r$ and A is $m\times n$, then $r \leq m$ and $r \leq n; r \leq \min\{m, n\}$

3) $\rho(A) = r$ means, every $(r+1)^{th}$ minor = 0 and $\exists$ an rth minor of A, which is $\neq 0$.

10.4.11 Theorem: If A is an m x n matrix of rank r, then $\exists$ non-singular matrices P and $Q \ni PAQ = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$

Proof: Given that $\rho(A) = r$

$\therefore A$ can be reduced to the normal form $\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$

For getting this, let the number of row operations used be s and let the number of column operations used be t. We also know that every elementary row (column) operation on A is equivalent to pre-(post) multiplication of A by a suitable elementary matrix.

$\therefore \exists$ elementary matrices $P_1, P_2, \ldots, P_s$ and $Q_1, Q_2, \ldots Q_t \ni$

$$P_s P_{s-1} \ldots P_2 P_1 A Q_1, Q_2 \ldots Q_t = \begin{bmatrix} I_r & O \\ O & O \end{bmatrix}$$

We know that each elementary matrix is non-singular and the product of non-singular matrices is non-singular.

Let $P_s P_{s-1} \ldots P_2 = P$ and $Q_1, Q_2 \ldots Q_t = Q$.

$\Rightarrow P, Q$ are non-singular.

$\therefore PAQ = \begin{bmatrix} I_r & O \\ O & O \end{bmatrix}$

10.4.12 Note: If P and Q are non-singular, then $\rho(A) = \rho(PAQ)$.

10.4.13 Theorem: Every invertible n x n matrix is a product of elementary matrices.

Proof:- Let A be an invertible matrix $\Rightarrow A$ is nonsingular $\Rightarrow |A| \neq 0 \Rightarrow \rho(A) = n$.

$\therefore A$ can be reduced to I_n by a finite number of row, column operations.

We know that elementary row (column)operation is equivalent to pre (post) multiplication of A by a suitable elementary matrix.

$\therefore \exists$ elementary matrices $P_1, P_2, \ldots, P_s; Q_1, Q_2 \ldots, Q_t \ni$

$$P_s P_{s-1} \ldots P_2 P_1 A Q_1, Q_2 \ldots Q_t = I_n$$

Since each of P_i, Q_j is non-singular, we have $A = P_1^{-1} P_2^{-1} \ldots P_s^{-1} Q_t^{-1} \ldots Q_1^{-1}$.

$\Rightarrow A = a$ product of elementary matrices.

10.4.14 Note: $|A| \neq 0$, the A can be expressed as a product of elementary matrices in many ways.

10.4.15 Theorem: If $A \in F^{m\times n}, P \in F^{n\times n}$ and P is non-singular, then $\rho(PA) = \rho(A)$ and $\rho(AP) = \rho(A)$.

Proof: Since P is non-singular, $\exists$ elementary matrices

$$P_1, P_2, \dots, P_s \ni P = P_1P_2, \dots P_s$$

$$\therefore PA = P_1P_2, \dots P_sA$$

Pre-multiplication by s elementary matrices is equivalent to s elementary operations on A.

But elementary operations do not alter the rank.

$$\therefore \rho(PA) = \rho(A)$$

Similarly we can prove that $\rho(AP) = \rho(A)$

10.4.16 Theorem: If $A \sim B$, then $\rho(A) = \rho(B)$

Proof: If $A \sim B$, then B is obtained from A by a finite number of elementary operations.

$$\therefore \rho(A) = \rho(B)$$

10.4.17 Theorem: If A is m x n matrix, then $\rho(A^1) = \rho(A)$

Proof: $\exists$ non-singular matrices $P_{m\times m}$ and Q $_{n\times n} \ni PAQ = D = \begin{bmatrix} I_r & O \\ O & O \end{bmatrix}, \rho(A) = \rho(D) = r$

$$\therefore D^1 = (PAQ)^1 = Q^1A^1P^1$$

Since P and Q are invertible, we have P^1 and Q^1 are invertible.

$$\therefore \rho(A^1) = \rho(Q^1A^1P^1) = \rho(D^1)$$

Let $\rho(A) = r$. Since D^1 is an n x m matrix, $D^1 = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}^1$ so that $\rho(D^1) = r$.

$$\therefore \rho(A^1) = \rho(D^1) = r = \rho(A)$$

Hence the theorem.

10.4.18 Theorem: Let U, V, W be finite dimensional vector spaces and $T : U \to V, S : V \to W$ be linear transformations. Let A, B be matrices such that AB is defined. Then

(i) $\rho(ST) \le \rho(S)$

(ii) $\rho(AB) \leq \rho(A)$

(iii) $\rho(AB) \leq \rho(B)$

iv) $\rho(ST) \leq \rho(T)$

Proof: We have $R(ST) = ST(U) = S\left[T(U)\right] = S\left[R(T)\right]$

$$\subseteq S(V) (\because R(T) \subseteq V)$$

$$= R(S)$$

$\therefore \rho(ST) = \dim R(ST) \leq \dim R(S) = \rho(S)$

$\therefore \rho(ST) \leq \rho(S)$

$\rho(AB) = \rho(L_{AB}) =$ rank $(L_A L_B) \leq$ rank L_A (by (i)) = rank A

$\therefore \rho(AB) \leq \rho(A)$

(iii) $\rho(AB) = \rho(AB)^T = \rho(B^T A^T)$

$$\leq \rho(B^T)$$

$$= \rho(B)$$

$\therefore \rho(AB) \leq \rho(B)$

(iv) Let B_1, B_2, B_3 be ordered bases of U, V, W respectively.

Let $A_1 = \left[T; B_1, B_2\right], A_2 = \left[S; B_2, B_3\right]$. Then $A_2 A_1 = \left[ST; B_1, B_3\right]$

$\therefore$ Rank ST = Rank $A_2 A_1 \leq$ Rank A_1 by (iii)

= Rank T

$\therefore$ Rank ST $\leq$ Rank T.

10.4.19 Note: $\rho(AB) \leq \min\left\{\rho(A), \rho(B)\right\}, A, B$ are matrices of suitable orders.

$\rho(ST) \leq \min\left\{\rho(S), \rho(T)\right\}, T, S$ are linear transformations $\ni$ ST exists.

10.5 The Inverse of a Matrix:

We have remarked that an n x n matrix is invertible iff its rank is n. Since we know how to compute the rank of any matrix, we can always test a matrix to determine whether it is invertible. We now provide a simple technique for computing the inverse of a matrix using elementary row operations.

10.5.1 Definition: Let A, B be $m\times n, m\times p$ matrices respectively. By the argumented matrix $[A \quad B]$ or $[A/B]$ or $[A/B]$ we mean $m\times(n+p)$ matrix whose first n columns are the columns of A and whose last p columns are the columns of B.

10.5.2 Theorem:

If A is an invertible n x n matrix, it is possible to transform the matrix (A/I_n) into the matrix (I_n/A^{-1}) by a finite number of elementary row operations.

Proof: Suppose A is an n x n invertible matrix. Consider the n x 2n argumented matrix $C=[A/I_n]$.

We have $A^{-1}C=[A^{-1}A/A^{-1}I_n]=[I_n/A^{-1}]$ ----------- (1)

We know that A^{-1} is the product of elementary matrices, say

$A^{-1}=E_pE_{p-1}...E_1$. Then (1) becomes:

$E_pE_{p-1}...E_1[A/I_n]=A^{-1}C=[I_n/A^{-1}]$

Since multiplication on the left by an elementary matrix transforms the matrix by an elementary row operation, it is possible to transform the matrix $[A/I_n]$ into the form $[I_n/A^{-1}]$ by finite number of elementary row operations.

10.5.3 Theorem: If $A_{n\times n}$ is invertible matrix and if by a finite number of elementary row operations, the matrix $[A/I_n]$ is transformed into a matrix of the form $[I_n/B]$, then $B=A^{-1}$

Proof: Suppose A is an n x n invertible matrix.

For some n x n matrix, suppose that $[A/I_n]$ is transformed into $[I_n/B]$ by a finite number of elementary row operations. Let $E_{1,}E_2...E_p$ be the elementary matrices corresponding to these elementary row operations.

$\therefore E_pE_{p-1}...E_1[A/I_n]=[I_n/B]$

Let $M=E_pE_{p-1}...E_1$

$[M\,A/M]=M[A/I_n]=[I_n/B]$

Clearly $MA = I_n$ and $M = B$ and hence we have $M = A^{-1}$

$\therefore B = A^{-1}$

10.5.4 Note: We write $IA = A$.

By applying row operations of I on LHS, the same row operation should be applied on A. This should be continued till we get $BA = I$. Then $B = A^{-1}$.

Solved Problems:

10.5.5 Find the rank of (a) $A = \begin{bmatrix} 1 & 2 & 1 & 1 \\ 1 & 1 & -1 & 1 \end{bmatrix}$, (b) $A = \begin{bmatrix} 1 & 3 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 3 & 0 & 0 \end{bmatrix}$

Solution: (a) One row is not a multiple of other so that the two rows are L.I.

$\therefore \rho(A) = 2$

(b) $A \overset{R_{21}(-1)}{\sim} \begin{bmatrix} 1 & 3 & 1 & 1 \\ 0 & -3 & 0 & 0 \\ 0 & 3 & 0 & 0 \end{bmatrix} \overset{R_{32}(1)}{\sim} \begin{bmatrix} 1 & 3 & 1 & 1 \\ 0 & -3 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \overset{R_2(-\frac{1}{3})}{\sim} \begin{bmatrix} 1 & 3 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = B$ (say)

A is reduced to B.

$\therefore \rho(A) = \rho(B) = 2$, since B has two independent rows.

10.5.6 Find the rank of $A = \begin{bmatrix} 1 & 2 & 3 & 1 \\ 2 & 1 & 1 & 1 \\ 1 & -1 & 1 & 0 \end{bmatrix}$

Solution:

$$A \underset{R_{31}(-1)}{\overset{R_{21}(-2)}{\sim}} \begin{bmatrix} 1 & 2 & 3 & 1 \\ 0 & -3 & -5 & -1 \\ 0 & -3 & -4 & -1 \end{bmatrix} \overset{R_{32}(-1)}{\sim} \begin{bmatrix} 1 & 2 & 3 & 1 \\ 0 & -3 & -5 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \overset{R_2(-\frac{1}{3})}{\sim} \begin{bmatrix} 1 & 2 & 3 & 1 \\ 0 & 1 & 5/3 & 1/3 \\ 0 & 0 & 1 & 0 \end{bmatrix} = B$$, the Echelon form of A.

$\rho(A)$ = number of nonzero rows in the Echelon form = 2.

10.5.7 Find the rank of $A=\begin{bmatrix}2 & 3 & -1 & -1\\ 1 & -1 & -2 & -4\\ 3 & 1 & 3 & -2\\ 6 & 3 & 0 & -7\end{bmatrix}$

Solution:

$$A\overset{R_{12}}{\sim}\begin{bmatrix}1 & -1 & -2 & -4\\ 2 & 3 & -1 & -1\\ 3 & 1 & 3 & -2\\ 6 & 3 & 0 & -7\end{bmatrix}\underset{\substack{R_{31}(-3)\\ R_{41}(-6)}}{\overset{R_{21}(-2)}{\sim}}\begin{bmatrix}1 & -1 & -2 & -4\\ 0 & 5 & 3 & 7\\ 0 & 4 & 9 & 10\\ 0 & 9 & 12 & 17\end{bmatrix}$$

$$A\overset{R_{23(-1)}}{\sim}\begin{bmatrix}1 & -1 & -2 & -4\\ 0 & 1 & -6 & -3\\ 0 & 4 & 9 & 10\\ 0 & 9 & 12 & 17\end{bmatrix}\underset{R_{42}(-9)}{\overset{R_{32}(-4)}{\sim}}\begin{bmatrix}1 & -1 & -2 & -4\\ 0 & 1 & -6 & -3\\ 0 & 0 & 33 & 22\\ 0 & 0 & 66 & 44\end{bmatrix}$$

$$\overset{R_{43}(-2)}{\sim}\begin{bmatrix}1 & -1 & -2 & -4\\ 0 & 1 & -6 & -3\\ 0 & 4 & 33 & 22\\ 0 & 0 & 0 & 0\end{bmatrix}\overset{R_3(\frac{1}{33})}{\sim}\begin{bmatrix}1 & -1 & -2 & -4\\ 0 & 1 & -6 & -3\\ 0 & 0 & 1 & \frac{2}{3}\\ 0 & 0 & 0 & 0\end{bmatrix}=B(say)$$

B is the Echelon form $\therefore \rho(A)=3$. (= number of nonzero rows of B)

10.5.8 Reduce the matrix $A=\begin{bmatrix}1 & -1 & 2 & -3\\ 4 & 1 & 0 & 2\\ 0 & 3 & 0 & 4\\ 0 & 1 & 0 & 2\end{bmatrix}$ to the normal form and hence find the rank.

Solution:

$$A \overset{R_{21}(-4)}{\sim} \begin{bmatrix} 1 & -1 & 2 & -3 \\ 0 & 5 & -8 & 14 \\ 0 & 3 & 0 & 4 \\ 0 & 1 & 0 & 2 \end{bmatrix} \overset{R_{24}}{\sim} \begin{bmatrix} 1 & -1 & 2 & -3 \\ 0 & 1 & 0 & 2 \\ 0 & 3 & 0 & 4 \\ 0 & 5 & -8 & 14 \end{bmatrix}$$

$$\underset{C_{31}(-2)}{\overset{C_{21}(1)}{\sim}} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 2 \\ 0 & 3 & 0 & 4 \\ 0 & 5 & -8 & 14 \end{bmatrix} \underset{R_{42}(-5)}{\overset{R_{32}(-3)}{\sim}} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & -2 \\ 0 & 0 & -8 & 4 \end{bmatrix}$$

$$\overset{C_{42}(-2)}{\sim} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -2 \\ 0 & 0 & -8 & 4 \end{bmatrix} \overset{R_{34}}{\sim} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -8 & 4 \\ 0 & 0 & 0 & -2 \end{bmatrix}$$

$$\underset{R_4(-\frac{1}{2})}{\overset{R_3(-\frac{1}{8})}{\sim}} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -\frac{1}{2} \\ 0 & 0 & 0 & 1 \end{bmatrix} \overset{R_{34}(\frac{1}{2})}{\sim} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = I_4$$

$$\therefore A \sim I_4 \therefore \rho(A) = 4$$

10.5.9 Reduce the matrix $A = \begin{bmatrix} 0 & 2 & 3 \\ 2 & 4 & 0 \\ 3 & 0 & 1 \end{bmatrix}$ using row operations to a matrix B and obtain the rank.

Solution:

$$A \overset{R_{12}}{\sim} \begin{bmatrix} 2 & 4 & 0 \\ 0 & 2 & 3 \\ 3 & 0 & 1 \end{bmatrix} \overset{R_1(\frac{1}{2})}{\sim} \begin{bmatrix} 1 & 2 & 0 \\ 0 & 2 & 3 \\ 3 & 0 & 1 \end{bmatrix} \overset{R_{31}(-3)}{\sim} \begin{bmatrix} 1 & 2 & 0 \\ 0 & 2 & 3 \\ 0 & -6 & 1 \end{bmatrix}$$

$$\overset{R_{32}(3)}{\sim}\begin{bmatrix}1&2&0\\0&2&3\\0&0&10\end{bmatrix}\overset{R_3(\frac{1}{10})}{\sim}\begin{bmatrix}1&2&0\\0&2&3\\0&0&1\end{bmatrix}\overset{R_{23}(-3)}{\sim}\begin{bmatrix}1&2&0\\0&2&0\\0&0&1\end{bmatrix}$$

$$\overset{R_{12}(-1)}{\sim}\begin{bmatrix}1&0&0\\0&2&0\\0&0&1\end{bmatrix}\overset{R_2(1/2)}{\sim}\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}=I_3$$

$$\therefore A\sim I_3 . B=I_3 \therefore \rho(A)=3$$

10.5.10 Find non-singular matrices P and Q such that PAQ is of the form $\begin{bmatrix}I_r&O\\O&O\end{bmatrix}$, where

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

Solution: Write $A=IAI$

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

Apply : $R_{21}(-1):\begin{bmatrix}1&1&2\\0&1&1\\0&-1&-1\end{bmatrix}=\begin{bmatrix}1&0&0\\-1&1&0\\0&0&1\end{bmatrix}A\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}$

Apply : $C_{21}(-1),C_{31}(-2):\begin{bmatrix}1&0&0\\0&1&1\\0&-1&-1\end{bmatrix}=\begin{bmatrix}1&0&0\\-1&1&0\\0&0&1\end{bmatrix}A\begin{bmatrix}1&-1&-2\\0&1&0\\0&0&1\end{bmatrix}$

Apply $R_{32}(1):\begin{bmatrix}1&0&0\\0&1&1\\0&0&0\end{bmatrix}=\begin{bmatrix}1&0&0\\-1&1&0\\-1&+1&1\end{bmatrix}A\begin{bmatrix}1&-1&-2\\0&1&0\\0&0&1\end{bmatrix}$

Apply $C_{32}(-1)$: $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & 1 & 1 \end{bmatrix} A \begin{bmatrix} 1 & -1 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}$

If $P = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & 1 & 1 \end{bmatrix}, Q = \begin{bmatrix} 1 & -1 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}$, then

$$PAQ = \begin{bmatrix} I_2 & 0_{2\times1} \\ 0_{1\times2} & 0_{1\times1} \end{bmatrix}$$

10.5.11 Reduce $A = \begin{bmatrix} 5 & 3 & 14 & 4 \\ 0 & 1 & 2 & 1 \\ 1 & -1 & 2 & 0 \end{bmatrix}$ to Echelon form and hence find the rank.

Solution:

$$A \overset{R_{13}}{\sim} \begin{bmatrix} 1 & -1 & 2 & 0 \\ 0 & 1 & 2 & 1 \\ 5 & 3 & 14 & 4 \end{bmatrix} \overset{R_{31}(-5)}{\sim} \begin{bmatrix} 1 & -1 & 2 & 0 \\ 0 & 1 & 2 & 1 \\ 0 & 8 & 4 & 4 \end{bmatrix}$$

$$\overset{R_{32}(-8)}{\sim} \begin{bmatrix} 1 & -1 & 2 & 0 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & -12 & -4 \end{bmatrix} \overset{R_3(\frac{-1}{12})}{\sim} \begin{bmatrix} 1 & -1 & 2 & 0 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & \frac{1}{3} \end{bmatrix} = B$$

B is in Echelon form, and hence the rank of A is 3.

10.5.12 Verify whether the matrix $A = \begin{bmatrix} 0 & 2 & 4 \\ 2 & 4 & 2 \\ 3 & 3 & 1 \end{bmatrix}$ is invertible. If so, find its inverse.

Solution: Given $A=\begin{bmatrix}0 & 2 & 4\\ 2 & 4 & 2\\ 3 & 3 & 1\end{bmatrix}$

The argumented matrix $[A/I_A]=\left[\begin{array}{ccc|ccc}0 & 2 & 4 & 1 & 0 & 0\\ 2 & 4 & 2 & 0 & 1 & 0\\ 3 & 3 & 1 & 0 & 0 & 1\end{array}\right]$

We convert the matrix into the form $[I/B]$.

$$[A/I]\overset{R_{12}}{\sim}\left[\begin{array}{ccc|ccc}2 & 4 & 2 & 0 & 1 & 0\\ 2 & 4 & 2 & 1 & 0 & 0\\ 3 & 3 & 1 & 0 & 0 & 1\end{array}\right]$$

$$\overset{R_1(\frac{1}{2})}{\sim}\left[\begin{array}{ccc|ccc}1 & 2 & 1 & 0 & \frac{1}{2} & 0\\ 0 & 2 & 4 & 1 & 0 & 0\\ 3 & 3 & 1 & 0 & 0 & 1\end{array}\right]$$

$$\overset{R_{31}(-3)}{\sim}\left[\begin{array}{ccc|ccc}1 & 2 & 1 & 0 & \frac{1}{2} & 0\\ 0 & 2 & 4 & 1 & 0 & 0\\ 0 & -3 & -2 & 0 & \frac{-3}{2} & 1\end{array}\right]$$

$$\overset{R_2(\frac{1}{2})}{\sim}\left[\begin{array}{ccc|ccc}1 & 2 & 1 & 0 & \frac{1}{2} & 0\\ 0 & 1 & 2 & \frac{1}{2} & 0 & 0\\ 0 & -3 & -2 & 0 & \frac{-3}{2} & 1\end{array}\right]$$

$$\underset{R_{32}(3)}{\overset{R_{12}(-2)}{\sim}}\left[\begin{array}{ccc|ccc}1 & 0 & -3 & -1 & \frac{1}{2} & 0\\ 0 & 1 & 2 & \frac{1}{2} & 0 & 0\\ 0 & 0 & 4 & \frac{3}{2} & \frac{-3}{2} & 1\end{array}\right]\overset{R_3(\frac{1}{4})}{\sim}\left[\begin{array}{ccc|ccc}1 & 0 & -3 & -1 & \frac{1}{2} & 0\\ 0 & 1 & 2 & \frac{1}{2} & 0 & 0\\ 0 & 0 & 1 & \frac{3}{8} & \frac{-3}{8} & \frac{1}{4}\end{array}\right]$$

$$\overset{\substack{R_{13}(3)\\R_{23}(-2)}}{\sim}\left[\begin{array}{ccc|ccc} 1 & 0 & 1 & +\frac{1}{8} & -\frac{5}{8} & \frac{3}{4} \\ 0 & 1 & 0 & -\frac{1}{4} & \frac{3}{4} & -\frac{1}{2} \\ 0 & 0 & 1 & \frac{3}{8} & -\frac{3}{8} & \frac{1}{4} \end{array}\right]$$

$\therefore A$ is invertible and

$$A^{-1}=\begin{bmatrix} +\frac{1}{8} & -\frac{5}{8} & \frac{3}{4} \\ -\frac{1}{4} & \frac{3}{4} & -\frac{1}{2} \\ \frac{3}{8} & -\frac{3}{8} & \frac{1}{4} \end{bmatrix}$$

10.5.13 Using elementary operations, verify whether $A=\begin{bmatrix} 1 & 2 & 1 \\ 2 & 1 & -1 \\ 1 & 5 & 4 \end{bmatrix}$ is invertible.

Solution: Consider the argumented matrix $[A/I]=\left[\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 2 & 1 & -1 & 0 & 1 & 0 \\ 1 & 5 & 4 & 0 & 0 & 1 \end{array}\right]$

We verify whether $[A/I]$ can be converted into the form $[I/B]$ using elementary operations.

Now $[A/I]\underset{R_{31}(-1)}{\overset{R_{21}(-2)}{\sim}}=\left[\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & -3 & -3 & -2 & 1 & 0 \\ 0 & 3 & 3 & -1 & 0 & 1 \end{array}\right]$

$$\overset{R_{32}(1)}{\sim}\left[\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & -3 & -3 & -2 & 1 & 0 \\ 0 & 0 & 0 & -3 & 1 & 1 \end{array}\right]$$

$[A/I]$ cannot be converted into the form $[I/B]$, since the third row is a zero row.

Hence A is not invertible.

10.5.14 SAQ : Find the rank of $A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$

10.5.15 SAQ : Find the rank of $A = \begin{bmatrix} 1 & 1 & 1 \\ -1 & -1 & -1 \\ 1 & 1 & 1 \end{bmatrix}$

10.6 Answers to SAQ's:

10.4.6 SAQ: $A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix} \Rightarrow \rho(A) = ?$

We observe that R_1, R_3 are identical $\Rightarrow |A| = 0$

But $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \neq 0.$ $\quad \therefore \rho(A) = 2$

10.4.7 SAQ: $A = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 0 & 3 \\ 1 & 1 & 2 \end{bmatrix} \underset{R_{31}(-1)}{\overset{R_{21}(-1)}{\sim}} \begin{bmatrix} 1 & 2 & 1 \\ 0 & -2 & 2 \\ 0 & -1 & 1 \end{bmatrix} \overset{R_2(-\frac{1}{2})}{\sim} \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix}$

$$\sim \begin{bmatrix} & 2 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix} \therefore \rho(A) = 2$$

10.5.14 SAQ : Every 2nd order minor is 0.

$\therefore$ Rank A = 1.

10.5.15 SAQ : $A \underset{R_{31}(-1)}{\overset{R_{21}(1)}{\sim}} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \therefore \rho(A) = 1$

10.7 Summary:

Rank of an m x n matrix is discussed. Using elementary row/column operations, we obtained normal form. We also explaind the procedure for finding inverse.

10.8 Technical Terms:

Rank, Elementary operations, Normal form.

10.9 Exercises:

10.9.1 Find the rank of the matrix $A=\begin{bmatrix}1 & 2 & 3 & 0\\ 2 & 4 & 3 & 2\\ 3 & 2 & 1 & 3\\ 6 & 8 & 7 & 5\end{bmatrix}$

10.9.2 Find the rank of the matrix $\begin{bmatrix}-2 & -1 & -3 & -1\\ 1 & 2 & 3 & -1\\ 1 & 0 & 1 & 1\\ 0 & 1 & 1 & -1\end{bmatrix}$

10.9.3 Find the rank of the matrix $A=\begin{bmatrix}2 & -1 & 3 & 1\\ 1 & -8 & 6 & 8\\ 1 & 2 & 0 & -2\end{bmatrix}$

10.9.4 Find the rank of the matrix by reducing to normal form $A=\begin{bmatrix}1 & 1 & 1 & -1\\ 1 & 2 & 3 & 4\\ 3 & 4 & 5 & 2\end{bmatrix}$

10.9.5 Find two non singular matrices P and Q such that PAQ is in the normal form, where

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

10.9.6 Find two non-singular matrices P and Q such that PAQ is in the normal form, where

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

10.9.7 Find the inverse of A using elementary operations: $A=\begin{bmatrix}-1 & -3 & 3 & -1\\ 1 & 2 & -1 & 0\\ 2 & -5 & 2 & -3\\ -1 & 1 & 0 & 1\end{bmatrix}$

10.10 Answers to Exercises:

10.9.1 3

10.9.2 3

10.9.3 2

10.9.4 2

10.9.5 $P=\begin{pmatrix}1 & 0 & 0\\ -1 & 1 & 0\\ -1 & 1 & 1\end{pmatrix}, Q=\begin{pmatrix}1 & -1 & 1\\ 0 & 1 & -1\\ 0 & 0 & 1\end{pmatrix}$

10.9.6 $P=\begin{pmatrix}1 & 0 & 0 & 0\\ 3 & -1 & 0 & 0\\ 5 & -1 & -1 & 0\\ -1 & -1 & 1 & 1\end{pmatrix}, Q=\begin{pmatrix}1 & 0 & 0 & -1\\ 0 & 1 & -2 & -3\\ 0 & 0 & 0 & 1\end{pmatrix}$

10.9.7 $A^{-1}=\begin{pmatrix}0 & 2 & 1 & 3\\ 1 & 1 & -1 & -2\\ 1 & 2 & 0 & 1\\ -1 & 1 & 2 & 6\end{pmatrix}$

10.11 Model Examination Questions:

1. Expain the concept of rank of a matrix.

2. Using elementary operations reduce $A = \begin{pmatrix} 1 & -1 & 2 & -3 \\ 4 & 1 & 0 & 2 \\ 0 & 3 & 0 & 4 \\ 0 & 1 & 0 & 2 \end{pmatrix}$ to normal form and hence find rank.

3. Find nonsingular matrices P and Q such that PAQ is in the normal form, where $A = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 2 & 3 \\ 0 & -1 & -1 \end{bmatrix}$

4. Using elementary operations, find the inverse of $A = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 2 & 3 \\ 3 & 1 & 1 \end{bmatrix}$

10.12 Reference Books:

1. Hoffman and Kunze, Linear Algebra, 2nd edition - Prentice Hall.

2. Stephen H. Friedberg and others - Linear Algebra, Prentice Hall India Pvt. Ltd - New Delhi.

- A Satyanarayana Murty

LESSON - 11

SYSTEMS OF LINEAR EQUATIONS

11.1 Objective of the Lesson:

In this lesson, we study the systems of linear equations and find solutions, when exist, using elementary operations. The elementary operations are used to provide a computational method for finding all solutions to such systems.

11.2. Structure of the Lesson:

This lesson has the following components.

11.3 Introduction:

In this lesson, we study the systems of linear equations. "A System of n linear equations in n unknowns has a solution" - This statement, sometimes, may be incorrect, because several possibilities including no solution may arise.

11.4 System of Linear Equations - Theoretical Aspects:

The equation $b = a_1x_1 + a_2x_2 + \ldots + a_nx_n$ .... (1)

expressing b in term of the variables $x_1, x_2 \ldots, x_n$ and the scalars $a_1, a_2, \ldots, a_n$ is called a linear equation.

For a given b, we must find $x_1, x_2 \ldots, x_n$ satisfafying (1)

A solution to a linear equation (1) is an ordered collection of n scalars $y_1, y_2 \ldots, y_n \ni$ (1) is satisfied when $x_1 = y_1, x_2 = y_2 \ldots, x_n = y_n$ are substituted in (1).

11.4.1 Definition: A system of m linear equations in n unknowns over a field F or simply a linear system, is a set of m linear equations, each in n unknowns. A linear system can be denoted by

$$\left.\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m \end{array}\right\} \quad \ldots. \text{(S)}$$

where a_{ij} and $b_i \in F, 1 \le i \le m, 1 \le j \le n$ and $x_1, x_2, \ldots, x_n$ are variables taking values in F.

If we write $A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$, then

the system can be represented as:

$$AX = B$$

A is called coefficient matrix.

A solution of the system (S) is an n - tuple $s = \begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{bmatrix} \in F^n \ni$

As = B

The set of all solutions of a linear system is called the solution set of the system. A system of linear equations is said to be consistent if it has a solution. Otherwise, the system is said to be inconsistent.

11.4.2 System of Homogeneous linear equations:

Consider a system of m homogeneous linear equations in n unknowns namely

$$\left.\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = 0 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = 0 \end{array}\right\} \ldots\ldots\ldots\ (I)$$

where $a_{ij} \in F$. This system can be written as $AX = O$

11.4.3 Note: $x_1 = 0, x_2 = 0, ..., x_n = 0$ is a solution of $AX = O$.

i.e. $X = O$ is a solution of $AX = O$. This is called the trivial solution or zero solution. Any other solution is called a nonzero solution.

$\therefore AX = O$ is always consistent.

A system $AX = B$ of m linear equations in n unknows is said to be

(a) homogeneous if $B = O$

(b) non-homogeneous if $B \neq O$

Any homogeneous system has atleast one solution, namely the zero solution.

11.4.4 Theorem: Let $AX = O$ be a homogeneous system of m linear equations in n unknowns over a field F. Let K denote the set of all solutions of $AX = O$. Then $K = N(L_A)$

K is a subspace of $F^n, \dim K = n$ - rank $L_A = n$ - rank A.

Proof: We have $K = \left\{X \in F^n / AX = O\right\}$

We have proved that K is a subspace of F^n and it is also equal to $N(L_A)$, the nullspace of L_A i.e. $K = N(L_A)$.

$\therefore$ We know that rank L_A + nullity $L_A = \dim F^n = n (\because L_A : F^n \to F^m)$

$\Rightarrow$ rank A + nullity $L_A = n$

$\Rightarrow$ nullity $L_A = n - r$

11.4.5 Note: The system $AX = O$ has $n - r$ $L.I$ solutions.

11.4.6 Corollary: If $m < n,$, the system $AX = O$ has a non-zero solution.

Proof: Suppose $m < n$

$\therefore$ Rank A = Rank $L_A \leq m$

$\therefore \dim K = n - \rho(L_A) > n - m,$ when $K = N(L_A)$

Since $n - m > 0$, we have $\dim K > 0$, since $K \neq O$.

$\therefore \exists$ a nonzero solution $s \in K$.

$\therefore s$ is a nonzero solution to $AX = O$.

i.e. $AX = O$ has a nonzero solution.

We know that the solution set S of (I) i.e. $AX = O$ is a subspace of F^n. [Since $AO = O \Leftrightarrow O \in S \therefore \phi \neq S \subseteq F^n$.Let X_1, X_2 be solutions of $AX = O$

$$\therefore AX_1 = O, AX_2 = O \Rightarrow A(X_1 + X_2) = 0$$

Also $A(kX_1) = (kAX_1) = k.O = O$

or $A(kX_1 + lX_2) = A(kX_1) + A(lX_2) = kAX_1 + lAX_2 = k.O + l.O = O \Rightarrow kX_1 + lX_2 \in S$]

We observe that if $AX = O$ has a nonzero solution, then it has an infinite number of nonzero solutions, since $X_1 \neq O$ is a solution $\Rightarrow kX_1$ is also a solution for every $K \in \mathbb{R}$.

11.4.7 Note: A system of m homogeneous equations in n unknows with $m < n$, has a nonzero solution.

$$(\therefore \rho(A) = r \leq m < n \Rightarrow n - r > 0)$$

11.4.8 System of non-homogeneous equations : The equations

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

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

. . . .

$$a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_n$$

can be written as $AX = B$, where $A = (a_{ij})_{m\times n}, X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$

The system $AX = O$ is called the homogeneous system corresponding to $AX = B$.

11.4.9 Theorem: Let K be the solution set of the system $AX = B$ and let K_H be the solution set of the corresponding homogeneous system $AX = O$. Then for any solution S to $AX = B$,

$$K = \{S\} + K_H = \{S + K / K \in K_H\}.$$

Proof: Let S be any solution to $AX = B$. We show that $K = \{S\} + K_H \Rightarrow AS = B$.

Let $W \in K. \Rightarrow AW = B. \Rightarrow A(W - S) = AW - AS = B - B = 0$

$$\Rightarrow W - S \in K_H$$

$$\Rightarrow \exists K_O \in K_H \ni W - S = K_O$$

$$\therefore W = S + K_o \in \{S\} + K_H \; \therefore K \subseteq \{S\} + K_H$$

Now suppose $W \in \{S\} + K_H \Rightarrow \exists K_0 \in K_H \ni W = S + K_0$

$$\therefore AW = A(S + K_0) = AS + AK_0$$

$$= B + O = B \Rightarrow W \in K$$

$$\therefore \{S\} + K_H \subseteq K$$

$$\therefore K = \{S\} + K_H$$

11.4.10 Theorem: Let $AX = B$ be a system of linear equations. Then the system is consistent iff $\rho(A) = \rho(A/B)$

Proof: Suppose $AX = B$ is a system of linear equations.

Here A is the coeffiencient matrix and $[A/B]$ is the augmented matrix.

We know that $L_A : F^n \to F^m, L_A(X) = AX. \forall X \in F^n$.

$\therefore AX = B$ has a solution means that $\exists X_1 \in F^n \ni AX_1 = B$

$$\Rightarrow B = L_A(X_1)$$

$$\Rightarrow B \in R(L_A)$$

Let $a_1, a_2, \ldots, a_n$ be the columns of A.

We know that $R(L_A)$ = Span $\{a_1, a_2, \ldots, a_n\}$

$\therefore AX = B$ is consistent

$\Leftrightarrow AX = B$ has a solution.

$\Leftrightarrow B \in$ Span $\{a_1, a_2, \ldots, a_n\}$

But $B \in$ Span $\{a_1, a_2, \ldots, a_n\}$ $\Leftrightarrow$ Span $\{a_1, a_2, \ldots, a_n\}$ = Span $\{a_1, a_2, \ldots, a_n, B\}$

$$\Leftrightarrow \dim\{a_1, a_2, \ldots, a_n\} = \dim\{a_1, a_2, \ldots, a_n, B\}$$

$\Leftrightarrow$ rank A = rank $= [A/B]$. Hence the theorem.

Note: If $\rho(A) \neq \rho[A/B]$, the system is inconsistent (i.e.) system has no solution).

11.4.11 Theorem: If A is a non-singular matrix of order n, then the system $AX = B$ in n un-knowns has a unique solution.

Proof: Since A is nonsingular matrix of order n, $|A| \neq 0$.

$\therefore \rho(A) = n, \rho[A/B] = n$

$\therefore \rho(A) = \rho[A/B] \Rightarrow AX = B$ is consistent.

$\Rightarrow AX = B$ has a solution.

Since A is nonsingular, A^{-1} exists $\therefore AX = B \Rightarrow A^{-1}(AX) = A^{-1}B$

$$\Rightarrow (A^{-1}A)X = A^{-1}B \Rightarrow IX = A^{-1}B$$

$$\Rightarrow X = A^{-1}B$$

$\therefore X = A^{-1}B$ is a solution of $AX = B$.

Let X_1, X_2 be solutions of $AX = B \Rightarrow AX_1 = B, AX_2 = B \Rightarrow AX_1 = AX_2$

$\therefore A^{-1}(AX_1) = A^{-1}(AX_2) \Rightarrow (A^{-1}A)X_1 = (A^{-1}A)X_2 \Rightarrow Iy_1 = Iy_2 \Rightarrow X_1 = X_2$

Hence the uniqueness and hence the theorem.

11.4.12 Theorem: Let $AX = B$ be a system of n linear equations in n unknowns. If A is invertible, then the system has exactly one solution $A^{-1}B$. Conversely if the system has exactly one solution, then A is invertible.

Proof: Suppose A is inversible $\Rightarrow A^{-1}$ exists.

Write $s = A^{-1}B$.

$\therefore As = (AA^{-1})B = (A^{-1}A)B = IB = B$

$\therefore s = A^{-1}B$ is a solution of $AX = B$.

Let s_1 be a solution of $AX = B \Rightarrow As_1 = B$

Now $As = B, As_1 = B \Rightarrow As = As_1$,

$$\therefore A^{-1}(As_1) = A^{-1}(As) \Rightarrow (A^{-1}A)s_1 = (A^{-1}A)s$$

$$\Rightarrow Is_1 = Is \Rightarrow s_1 = s$$

$\therefore s = A^{-1}B$ is the only solution of $AX = B$.

Conversely suppose that the system $AX = B$ has exactly one solution namely s. $\therefore As = B$

Let K_H denote the solution set for the corresponding homogeneous system $AX = O$.

$$\therefore \{s\} = \{s\} + K_H \Rightarrow K_H = \{O\}$$

$\therefore N(L_A) = \{O\} \Rightarrow L_A$ is non-singular

$\Rightarrow A$ is invertible [$L_A : F^n \to F^n$ is one-one $\Rightarrow L_A$ is onto]

Hence the Theorem.

Solved Problems:

11.4.13 Solve $x - 2y + z = 0$ (System containing one equation in 3 unknowns)

Solution: $A = \begin{bmatrix} 1 & -2 & 1 \end{bmatrix}$ is the coefficient matrix $\Rightarrow \rho(A) = 1$.

If K is the solution set, then dim $K = 3-1 = 2$.

We observe that $\alpha_1 = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}, \alpha_2 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$ are independent vectors in $K; \{\alpha_1, \alpha_2\}$ is a basis of K.

$K = \{a_1\alpha_1 + a_2\alpha_2 / a_1, a_2 \in \mathbb{R}\}$ is the solution set.

11.4.14 Solve: $x + 2y + z = 0, x - y - z = 0$ (System of 2 equations in 3 unknows)

Solution: The coefficient matrix is $A = \begin{bmatrix} 1 & 2 & 1 \\ 1 & -1 & -1 \end{bmatrix}$

We observe that $\rho(A) = 2$

If K is the solution set, the dim $K = 3-2 = 1$

We observe that $\alpha = \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix}$ is a basis of K.

Solution set is $= \{a\alpha / a \in \mathbb{R}\}$

11.4.15 Solve : $x + 2y + 3z = 0, 3x + 4y + 4z = 0, x + 10y + 12z = 0$

Solution: The coefficient matrix is $A = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 4 & 4 \\ 7 & 10 & 12 \end{bmatrix}$

$$A \underset{R_{31}(-7)}{\overset{R_{21}(-3)}{\sim}} \begin{bmatrix} 1 & 2 & 3 \\ 0 & -2 & -5 \\ 0 & -4 & -9 \end{bmatrix} \overset{R_2(-\frac{1}{2})}{\sim} \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 5/2 \\ 0 & -4 & -9 \end{bmatrix} \overset{R_{32}(4)}{\sim} \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 5/2 \\ 0 & 0 & 1 \end{bmatrix} = B$$

$\therefore A \sim B \therefore B$ is the echelon form.

$\therefore \rho(A) = 3$ = number of unknowns.

$\therefore K = \left\{ \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \right\}$. The system has zero solution only.

11.4.16 Problem: Solve : $x + y + z = 0, 3x + 4y + 5z = 0, 2x + 3y + 4z = 0$

Solution: The coefficient matrix is $A = \begin{bmatrix} 1 & 1 & 1 \\ 3 & 4 & 5 \\ 2 & 3 & 4 \end{bmatrix}$

$$A \underset{R_{31}(-2)}{\overset{R_{21}(-3)}{\sim}} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 1 & 2 \end{bmatrix} \overset{R_{32}(-1)}{\sim} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix} = B$$

$\therefore A \sim B, \rho(B) = 2 \therefore \rho(A) = 2$

If K is the solution set, then dim $K = 3 - 2 = 1$.

$\therefore$ The system has only one L.I solution.

The given system is equivalent to $x + y + z = 0$

$$y + 2z = 0$$

$$\Rightarrow \quad y = -2z$$

$$\therefore x - 2z + z = 0 \Rightarrow x = z$$

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

If $z = k,$ then $x = k, y = -2k, z = k$

$$\therefore K = \{k\alpha / k \in \mathbb{R}\}, \alpha = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}.$$

11.4.17 Solve $x+y-z+t=0, x-y+2z-t=0, 3x+y+t=0$

Solution: The coefficient matrix is $A=\begin{bmatrix}1 & 1 & -1 & 1\\ 1 & -1 & 2 & -1\\ 3 & 1 & 0 & 1\end{bmatrix}$

Now $A \underset{R_{31}(-3)}{\overset{R_{21}(-1)}{\sim}} \begin{bmatrix}1 & 1 & -1 & 1\\ 0 & -2 & 3 & -2\\ 3 & -2 & 3 & -2\end{bmatrix} \overset{R_{32}(-1)}{\sim} \begin{bmatrix}1 & 1 & -1 & 1\\ 0 & -2 & 3 & -2\\ 0 & 0 & 0 & 0\end{bmatrix} = B$ (say)

$\rho(B)=2 \therefore \rho(A)=2$

Let K be the solution set.

$\therefore \dim K = n-r = 4-2 = 2$

The given system is equivalent to $x+y-z+t=0$

$$-2y+3z-2t=0$$

Let $z=k_1, t=k_2$ $\quad \therefore 2y = 3k_1 - 2k_2 \Rightarrow y = \frac{3}{2}k_1 - k_2$

$$\therefore x = -y+z-t = \frac{3}{2}k_1 + k_2 + k_1 - k_2 = -\frac{1}{2}k_1$$

$$\therefore \alpha = \begin{bmatrix}x\\ y\\ z\\ t\end{bmatrix} = \begin{bmatrix}-\frac{1}{2}k_1\\ \frac{3}{2}k_1 - k_2\\ k_1\\ k_2\end{bmatrix} = k_1\begin{bmatrix}-\frac{1}{2}\\ \frac{3}{2}\\ 1\\ 0\end{bmatrix} + k_2\begin{bmatrix}0\\ -1\\ 0\\ 1\end{bmatrix}$$

$\therefore$ Solution set $K = \left\{ \begin{bmatrix}-1/2\\ 3/2\\ 1\\ 0\end{bmatrix} k_1 + k_2 \begin{bmatrix}0\\ -1\\ 0\\ 1\end{bmatrix} \middle/ k_1, k_2 \in \mathbb{R} \right\}$

11.5 System of linear equations - computationel aspects: In this section, we use elementary row operations to find one solution and using that all the solutions to the given non-homogeneous system (when the system is consistent).

11.5.1 Definition: If two systems of linear equations have the same solution set, then the systems are said to be equivalent.

11.5.2 Theorem: Let $AX = B$ be a system of m linear equations in n unknowns and C be an invertible m x m matrix.

Then the system $(CA)X = CB$ is equivalent to $AX = B$.

Proof: Let K be the solution set for $AX = B$ and let

K^1 be the solution set for $(CA)X = CB$.

Let $W \in K$. Then $AW = B \Rightarrow C(AW) = CB \Rightarrow (CA)W = CB$

$$\therefore W \in K^1 \qquad \therefore K \subseteq K^1$$

Let $W \in K^1 \Rightarrow (CA)W = CB \Rightarrow C(AW) = CB$

Hence $AW = C^{-1}(CAW) = C^{-1}(CB) = (C^{-1}C)B = IB = B$

$\Rightarrow W \in K \therefore K^1 \subseteq K \therefore K = K^1$

Hence the theorem.

11.5.3 Corollary: Let $AX = B$ be a system of m linear equations in n unknowns. If $\left[A^1/B^1\right]$ is a matrix obtained from $[A/B]$ by a finite number of elementary row operations, then the system $A^1X = B^1$ is equivalent to the original system.

Proof: Let (A/B) be the augmented matrix of the system $AX = B$.

Let $\left(A^1/B^1\right)$ be obtained from (A/B) by elementary row operations.

This is equivelant to pre-multiplication of $[A/B]$ by the elementary matrices (of order m x m), $E_1, E_2 ..., E_p$. Let $C = E_p E_{p-1} ..., E_2 E_1$.

$\therefore \left(A^1/B^1\right) = C[A/B] = [CA/CB]$. Now C is invertible, since each E_i is invertible.

Hence the system $A^1X = B^1$ is equivalent to $AX = B$.

11.5.4 Echelon form of a matrix:

If all the elements in a row of a matrix are zeros, then it is called a zero row and if there is atleast one nonzero element in a row, then it is called a nonzero row.

11.5.5 Definition: A matrix is said to be in Echelon form if it has the following properties:

i) Zero rows, if any, must follow nonzero rows.

ii) The first nonzero element in each nonzero row is 1.

iii) The number of zeros before the first nonzero element in a row is less than the number of zeros before the nonzero element of the next row.

11.5.6 Note: 1) Some authors ignore the property (ii) to consider a matrix to be in Echelon form.

2) The rank of a matrix in Echelon form is equal to the number of nonzero rows of the matrix.

11.5.7 Gaussian Elimination:

We now explain a procedure for solving any system of linear equations by using the following example. This procedure is called Gaussian elimination.

Consider the system of linear equations : $3x_1 + 2x_2 + 3x_3 - 2x_4 = 1$

$$x_1 + x_2 + x_3 = 3$$

$$x_1 + 2x_2 + x_3 - x_4 = 2$$

The augmented matrix is $[A/B]$ is $\begin{bmatrix} 3 & 2 & 3 & -2 & 1 \\ 1 & 1 & 1 & 0 & 3 \\ 1 & 2 & 1 & -1 & 2 \end{bmatrix}$

By performing elementary row operations, the augmented matrix is transformed into an upper triangular matrix in which the first nonzero entry of each row is 1 and one occurs in a column to the right of the first nonzero entry of each preceding row.

($A = [a_{ij}]_{n \times n}$ is upper triangular if $a_{ij} = 0$ if $i > j$)

1. To get 1 as the first row first column element, we interchange R_1 and R_3.

$$[A \quad B] \overset{R_{13}}{\sim} \left[\begin{array}{cccc|c} 1 & 2 & 1 & -1 & 2 \\ 1 & 1 & 1 & 0 & 3 \\ 3 & 2 & 3 & -2 & 1 \end{array}\right]$$

2. Using type 3 row operations, we use R_1 to get zeros in the remaining positions of C_1.

By applying $R_{21}(-1), R_{31}(-3),$ we get $\begin{bmatrix} 1 & 2 & 1 & -1 & 2 \\ 0 & -1 & 0 & 1 & 1 \\ 0 & -4 & 0 & 1 & -5 \end{bmatrix}$

3. We get 1 in the next row in the left most possible column, without using previous rows. In this example, C_2 is the left most possible column.

By applying $R_2(-1)$, we get 1 in $(2,2)$ position, we get $\begin{bmatrix} 1 & 2 & 1 & -1 & 2 \\ 0 & -1 & 0 & -1 & -1 \\ 0 & -4 & 0 & 1 & -5 \end{bmatrix}$

4. Now use Type 3 - elmentary row operations to get zeros below 1. In this example, we apply

$R_{32}(4), \begin{bmatrix} 1 & 2 & 1 & -1 & 2 \\ 0 & 1 & 0 & -1 & -1 \\ 0 & 0 & 0 & -3 & -9 \end{bmatrix}$

5. Repeat steps 3, 4 on each succeeding row until no nonzero rows remain.

By applying $R_3(-1/3)$, we get $\begin{bmatrix} 1 & 2 & 1 & -1 & 2 \\ 0 & 1 & 0 & -1 & -1 \\ 0 & 0 & 0 & 1 & 3 \end{bmatrix}$

6. Work upward, begining with last nonzero row and add multiples of each row to the rows above. (so that we get zeros above the first nonzero entry in each row).

By applying $R_{13}(1), R_{23}(1),$ we get $\begin{bmatrix} 1 & 2 & 1 & 0 & 5 \\ 0 & 1 & 0 & 0 & 2 \\ 0 & 0 & 0 & 1 & 3 \end{bmatrix}$

7. Repeat the process described in step 6 for each preceeding row until it is performed with the 2nd row at which the reduction process is complete.

In this example, by applying $R_{12}(-2),$ we get $\begin{bmatrix} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 2 \\ 0 & 0 & 0 & 1 & 3 \end{bmatrix}$

This matrix corresponds to the system : $x_1 + x_3 = 1$

$$x_2 = 2$$

$$x_4 = 3$$

This system is equivalent to the original system.

The equivalent system can easily be solved.

Solved Problems:

11.5.8 Show that the equations $x + y + z = 4, 2x + 5y - 2z = 3$, $x + 7y - 7z = 5$ are inconsistent.

Solution: The augmented matrix is

$$[A/B] = \begin{bmatrix} 1 & 1 & 1 & 4 \\ 2 & 5 & -2 & 3 \\ 1 & 7 & -7 & 5 \end{bmatrix}$$

$$\underset{R_{31}(-1)}{\overset{R_{21}(-2)}{\sim}} \begin{bmatrix} 1 & 1 & 1 & 4 \\ 0 & 3 & -4 & -5 \\ 0 & 6 & -8 & -9 \end{bmatrix} \overset{R_{32}(-2)}{\sim} \begin{bmatrix} 1 & 1 & 1 & 4 \\ 0 & 3 & -4 & -5 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\rho(A) = 2, \rho[A/B] = 3$$

Since $\rho(A) \neq \rho[A/B]$, the system is inconsistent.

11.5.9 Solve the System : $x + 2y + z = 2, 3x + y - 2z = 1$,

$$4x - 3y - z = 3$$

$$2x + 4y + 2z = 4$$

Solution: The augmented matrix of the system is :

$$[A/B] = \begin{bmatrix} 1 & 2 & 1 & 2 \\ 3 & 1 & -2 & 1 \\ 4 & -3 & -1 & 3 \\ 2 & 4 & 2 & 4 \end{bmatrix}$$

$$\underset{R_{31}(-4)}{\overset{R_{21}(-3)}{\sim}}\begin{bmatrix}1 & 2 & 1 & 2\\0 & -5 & -5 & -5\\0 & -11 & -5 & -5\\0 & 0 & 0 & 0\end{bmatrix}\overset{R_2(\frac{-1}{5})}{\sim}\begin{bmatrix}1 & 2 & 1 & 2\\0 & 1 & 1 & 1\\0 & -11 & -5 & -5\\0 & 0 & 0 & 0\end{bmatrix}$$

$$\overset{R_{32}(11)}{\sim}\begin{bmatrix}1 & 2 & 1 & 2\\0 & 1 & 1 & 1\\0 & 0 & 6 & 6\\0 & 0 & 0 & 0\end{bmatrix}\overset{R_3(\frac{1}{6})}{\sim}\begin{bmatrix}1 & 2 & 1 & 2\\0 & 1 & 1 & 1\\0 & 0 & 1 & 1\\0 & 0 & 0 & 0\end{bmatrix}$$

$$\rho(A)=3=\rho[A/B]$$

The system is equivalent to $x+2y+z=2$

$$y+z=1$$

$$z=1$$

$$\therefore y=0$$

$$\therefore x+0+1=2$$

$$\therefore x=1$$

$\rho(A)=3 \therefore A$ is invertible and hence the system has a unique solution, namely $\begin{bmatrix}1\\0\\1\end{bmatrix}$.

11.5.10 Problem: Obtain for what values of λ and μ, the equations $x+y+z=6, x+2y+3z=10, x+2y+\lambda z=\mu$ have

(a) no solution (b) a unique solution (c) an infinite number of solutions.

Solution: The augmented matrix of the system is $[A/B]=\left[\begin{array}{ccc|c}1 & 1 & 1 & 6\\1 & 2 & 3 & 10\\1 & 2 & \lambda & \mu\end{array}\right]$

$$\underset{R_{32}(-1)}{\overset{R_{21}(-1)}{\sim}}\left[\begin{array}{ccc|c} 1 & 1 & 1 & 6 \\ 0 & 1 & 2 & 4 \\ 0 & 0 & \lambda-3 & \mu-10 \end{array}\right]$$

If $\lambda \neq 3$, then $\rho(A)=3$ and $\rho[A/B]=3$.

$\therefore$ The system has unique solution.

if $\lambda=3$ and $\mu=10$, then $\rho(A)=2=\rho[A/B]<3$.

$\therefore$ The system has an infinite number of solutions.

If $\lambda=3$ and $\mu=10$, then $[A/B] \sim \left[\begin{array}{ccc|c} 1 & 1 & 1 & 6 \\ 0 & 1 & 2 & 4 \\ 0 & 0 & 0 & \mu-10 \end{array}\right]$

Since $\mu-10 \neq 0$, $\rho(A)=2 \neq \rho[A/B](=3)$.

In this case, the system is inconsistent.

Hence $\lambda=3, \mu \neq 10 \Rightarrow$ system is inconsistent.

$\lambda \neq 3 \quad \Rightarrow$ System has unique solution.

$\lambda=3, \mu=10 \Rightarrow$ System has an infinite number of solutions.

11.5.11 Solve the following system by reducing to reduced row echelon form

$$2x_1+3x_2+x_3+4x_4-9x_5=17$$

$$x_1+x_2+x_3+x_4-3x_5=6$$

$$x_1+x_2+x_3+2x_4-5x_5=8$$

$$2x_1+2x_2+2x_3+3x_4-x_5=14$$

Solution: The augmented matrix is:

$$[A/B]=\begin{bmatrix} 2 & 3 & 1 & 4 & -9 & 17 \\ 1 & 1 & 1 & 1 & -3 & 6 \\ 1 & 1 & 1 & 2 & -5 & 8 \\ 2 & 2 & 2 & 3 & -8 & 14 \end{bmatrix}$$

$$\overset{R_{12}}{\sim}\begin{bmatrix} 1 & 1 & 1 & 1 & -3 & 6 \\ 2 & 3 & 1 & 4 & -9 & 17 \\ 1 & 1 & 1 & 2 & -5 & 8 \\ 2 & 2 & 2 & 3 & -8 & 14 \end{bmatrix}$$

$$\underset{R_{31}(-1)}{\overset{R_{21}(-2)}{\sim}}\begin{bmatrix} 1 & 1 & 1 & 1 & -3 & 6 \\ 0 & 1 & -1 & 2 & -3 & 5 \\ 0 & 0 & 0 & 1 & -2 & 2 \\ 0 & 0 & 0 & 1 & -2 & 2 \end{bmatrix}$$

$$\overset{R_{34}(-1)}{\sim}\begin{bmatrix} 1 & 1 & 1 & 1 & -3 & 6 \\ 0 & 1 & -1 & 2 & -3 & 5 \\ 0 & 0 & 0 & 1 & -2 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$\therefore \rho(A) = 3 = \rho[A/B]$

The system is equivalent to $x_1 + x_2 + x_3 + x_4 - 3x_5 = 6$

$$x_2 - x_3 + 2x_4 - 3x_5 = 5$$

$$x_4 - 2x_5 = 2$$

System has $n - r = 5 - 3 = 2$ L.I solutions:

Let $x_3 = t_1, x_5 = t_2$ $\quad \therefore x_4 = 2 + 2t_2$

$$\therefore x_2 - t_1 + 2(2 + 2t_2) - 3t_2 = 5 \Rightarrow x_2 - t_1 + 4 + t_2 = 5$$

$$\Rightarrow x_2 = 1 + t_1 - t_2$$

$$\therefore x_2 + 1 + t_1 - t_2 + t_1 + 2 + 2t_2 - 3t_2 = 6 \Rightarrow x_1 + 2t_1 - 2t_2 = 3$$

$$\Rightarrow x_1 = 3 - 2t_1 + 2t_2$$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 3-2t_1+2t_2 \\ 1+t_1-t_2 \\ t_1 \\ 2+2t_2 \\ t_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \\ 0 \\ 2 \\ 0 \end{bmatrix} + t_1 \begin{bmatrix} -2 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t_2 \begin{bmatrix} 2 \\ -1 \\ 0 \\ 2 \\ 1 \end{bmatrix}, t_1, t_2 \in \mathbb{R}$$

We observe that $\begin{bmatrix} 3 \\ 1 \\ 0 \\ 2 \\ 0 \end{bmatrix}$ is a particular solution of the system and $\left\{ \begin{bmatrix} -2 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \\ 0 \\ 2 \\ 1 \end{bmatrix} \right\}$ is a basis of the corresponding homogeneous system.

11.5.12 SAQs: Verify whether the system $x + 2y + 3z = 0$

$$3x + 4y + 4z = 0$$

$$7x + 10y + 12z = 0$$

has a non-trivial solution.

11.5.13 SAQ: Verify whether the system $2x + 6y + 11 = 0$

$$6x + 20y - 6z + 3 = 0$$

$$6y - 18z + 1 = 0$$

is consistent.

11.5.14 SAQ : Solve $2x - y + 3z = 8$

$$-x + 2y + z = 4$$

$$3x + y - 4z = 0$$

11.5.15 SAQ: Verify whether the system $x + y + z = -3$

$$3x + y - 2z = -2$$

$$2x + 4y + 7z = 7$$

is consistent.

11.5.16 SAQ: Verify whether the system $x+2y-z=3$

$$3x-y+2z=1$$

$$2x-2y+3z=2$$

$$x-y+z=-1$$

is consistent. If it is consistent, solve.

11.6 Answer to SAQ's:

11.5.12 SAQ: The coefficient matrix is $A=\begin{bmatrix}1 & 2 & 3\\ 3 & 4 & 4\\ 7 & 10 & 12\end{bmatrix}$

$$\therefore A \underset{R_{31}(-7)}{\overset{R_{21}(-3)}{\sim}} = \begin{bmatrix}1 & 2 & 3\\ 0 & -2 & -5\\ 0 & -4 & -9\end{bmatrix} \overset{R_2(\frac{-1}{2})}{\sim} \begin{bmatrix}1 & 2 & 3\\ 0 & 1 & \frac{5}{2}\\ 0 & -4 & -9\end{bmatrix}$$

$$\overset{R_{32}(4)}{\sim} \begin{bmatrix}1 & 2 & 3\\ 0 & 1 & \frac{5}{2}\\ 0 & 0 & 1\end{bmatrix} = B$$

Since $A \sim B$ and B is the echelon form, $\rho(A)=3$

$\therefore r=n$ $\therefore$ The system has trivial solution only.

11.5.13 SAQ: The augmented matrix is

$$[A \quad B] = \begin{bmatrix}2 & 6 & 0 & -11\\ 6 & 20 & -6 & -3\\ 0 & 6 & -18 & -1\end{bmatrix} \overset{R_{21}(-3)}{\sim} \left[\begin{array}{ccc|c}2 & 6 & 0 & -11\\ 0 & 2 & -6 & 30\\ 0 & 6 & -18 & -1\end{array}\right] \overset{R_{32}(-3)}{\sim} \left[\begin{array}{ccc|c}2 & 6 & 0 & -11\\ 0 & 2 & -6 & 30\\ 0 & 0 & 0 & -91\end{array}\right]$$

$$\rho(A)=2, \rho[A/B]=3 \therefore \rho(A) \neq \rho[A/B]$$

$\therefore$ The system is inconsistent.

11.5.14 SAQ: The augmented matrix of the system is

$$[A/B]=\begin{bmatrix}2 & -1 & 3 & 8\\ -1 & 2 & 1 & 4\\ 3 & 1 & -4 & 0\end{bmatrix}\overset{R_{12}}{\sim}\begin{bmatrix}-1 & 2 & 1 & 4\\ 2 & -1 & 3 & 8\\ 3 & 1 & -4 & 0\end{bmatrix}$$

$$\underset{R_{31}(+3)}{\overset{R_{21}(2)}{\sim}}\begin{bmatrix}-1 & 2 & 1 & 4\\ 0 & 3 & 5 & 16\\ 0 & 7 & -1 & 12\end{bmatrix}\overset{R_2(\frac{1}{3})}{\sim}\begin{bmatrix}-1 & 2 & 1 & 4\\ 0 & 1 & \frac{5}{3} & \frac{16}{3}\\ 0 & 7 & -1 & 12\end{bmatrix}$$

$$\overset{R_{32}(-7)}{\sim}\begin{bmatrix}-1 & 2 & 1 & 4\\ 0 & 1 & 5/3 & 16/3\\ 0 & 0 & -3/3 & -76/3\end{bmatrix}\overset{R_3(\frac{-3}{38})}{\sim}\begin{bmatrix}-1 & 2 & 1 & 4\\ 0 & 1 & 5/3 & 16/3\\ 0 & 0 & 1 & 2\end{bmatrix}$$

$$\therefore \rho(A)=\rho[A/B]=3$$

$\therefore$ The system is consistent. It has unique solution.

The equivalent system is $-x+2y+z=4$

$$y+\frac{5}{3}z=\frac{16}{3}$$

$$z=2$$

$$\therefore y+\frac{10}{3}=\frac{16}{3}\Rightarrow y=2$$

$$\therefore x+4+2=6\Rightarrow x=2$$

$\therefore$ solution is $\begin{bmatrix}2\\2\\2\end{bmatrix}$

11.5.15 SAQ : The augmented matrix is

$$[A/B]=\left[\begin{array}{ccc|c}1&1&1&-3\\3&1&-2&-2\\2&4&7&7\end{array}\right]\underset{R_{31}(-2)}{\overset{R_{21}(-3)}{\sim}}\left[\begin{array}{cccc}1&1&1&-3\\0&-2&-5&7\\0&2&5&13\end{array}\right]\overset{R_{43}(-1)}{\sim}\left[\begin{array}{ccc|c}1&2&-1&3\\0&-2&-5&7\\0&0&0&20\end{array}\right]$$

$\therefore \rho(A)=2\neq 3=\rho[A/B]\therefore$ The system is inconsistent.

11.5.16 SAQ: The augmented matrix is $[A/B]=\begin{bmatrix}1&2&-1&3\\3&-1&2&1\\2&-2&3&2\\1&-1&1&-1\end{bmatrix}$

$$[A/B]\underset{\substack{R_{31}(-2)\\R_{41}(-1)}}{\overset{R_{21}(-3)}{\sim}}\begin{bmatrix}1&2&-1&3\\0&-7&5&-8\\0&-6&5&-4\\0&-3&2&-4\end{bmatrix}\overset{R_{23}(-1)}{\sim}\begin{bmatrix}1&2&-1&3\\0&-1&0&-4\\0&-6&5&-4\\0&-3&2&-4\end{bmatrix}$$

$$\overset{R_2(-1)}{\sim}\begin{bmatrix}1&2&-1&3\\0&1&0&4\\0&-6&5&-4\\0&-3&2&-4\end{bmatrix}\underset{R_{42}(3)}{\overset{R_{32}(6)}{\sim}}\begin{bmatrix}1&2&-1&3\\0&1&0&4\\0&0&5&20\\0&0&2&8\end{bmatrix}$$

$$\underset{R_4(1/2)}{\overset{R_3(1/5)}{\sim}}\begin{bmatrix}1&+2&-1&3\\0&1&0&4\\0&0&1&4\\0&0&1&4\end{bmatrix}$$

$$\overset{R_{43}(-1)}{\sim}\left[\begin{array}{ccc|c}1&2&-1&3\\0&1&0&4\\0&0&1&4\\0&0&0&0\end{array}\right]$$

$\therefore \rho(A) = \rho[A/B] = 3 \therefore$ The system is consistent. Equivalent system is

$x + 2y - z = 3, y = 4, z = 4 \therefore x = -1$

$\therefore$ The system has unique solution.

11.7 Summary:

The theory of homogeneous linear equations and the theory of non homogeneous linear equations is discussed.

11.8 Technical Terms:

Linear equations, homogeneous, non-homogeneous linear equations consistency, inconsistency.

11.9 Exercises:

11.9.1 Find the dimension and basis set of $x_1 + 3x_2 = 0, 2x_1 + 6x_2 = 0$

11.9.2 Solve : $x_1 + 2x_2 - x_3 = 0, 2x_1 + x_2 + x_3 = 0$

11.9.3 Examine for consistency $x + 2y - z = 3$

$$3x - y + 2z = 1$$

$$2x - 2y + 3z = 2$$

$$x - y + z = 1$$

11.9.4 Solve : $2x + 2y - 2z = 1$

$$4x + 2y - z = 2$$

$$6x + 6y + \lambda z = 3 \forall \lambda \in \mathbb{R}$$

11.9.5 Determine whether the following system has a solution:

$$x_1 + 2x_2 + 3x_3 = 1$$

$$x_1 + x_2 - x_3 = 0$$

$$x_1 + 2x_2 + x_3 = 3$$

11.10 Answers to Exercises:

11.9.1 $\left\{\begin{bmatrix}-3\\1\end{bmatrix}\right\}$

11.9.2 $\left\{\begin{bmatrix}-1\\1\\1\end{bmatrix}\right\}$

11.9.3 Consistent

11.9.4 If $\lambda \neq 2$, system has unique solution.

$$\lambda = 2 \Rightarrow x = \frac{1}{2} - k, y = k; z = 0, k \in \mathbb{R}$$

11.9.5 System has a solution.

11.11 Model Examination Questions:

1. Prove that the system $AX = O$ has $n - r$ L.I solutions, where r is rank A and n is the number of unknowns.

2. Prove that the system $AX = B$ is consistent iff $\rho(A) = \rho[A/B]$

3. Solve the system $x_1 + 2x_2 + x_3 = 2$

$$3x_1 + x_2 - 2x_3 = 1$$

$$4x_1 - 3x_2 - x_3 = 3$$

$$2x_1 + 4x_2 + x_3 = 4$$

4. Prove that the system $AX = B$ has unique solution iff A is invertible.

5. Show that the system $x + 2y - z = 3, 3x - y + 2z = 1$, $2x - 2y + 3z = 2, x - y + z = 1$ is consistent and solve.

11.12 Reference Books:

1. Hoffmen and Kunze - Linear algebra, 2nd Edition Prentice Hall.

2. Stephen H. Friedberg and others - Linear algebra - Prentice Hall-India Pvt. Ltd. - New Delhi.

- A. Satyanarayana Murty

LESSON - 12

DIAGONALIZATION

12.1 Objective of the Lesson:

This lesson is concerned with diagonalization problem. For a given operator T on a finite dimensional vector space we study about

i) The existance of an ordered basis B for V and

ii) If such basis exists the method of finding it.

A solution of diagonalization problem leads to the concept of eigen values and eigen vectors and so we study about,

iii) Finding eigen values and eigen vectors of linear transformations.

12.2 Structure of the Lesson:

This lesson contains the following items.

12.3 Introduction:

In this lesson we introduce the important notions of eigen values and eigen vectors of a linear operator and a square matrix defined over a field. Using these concepts we discuss the diagonalization and diagonalizability of linear operators and matrices.

12.4 Diagnolizable Linear Operator:

12.4.1 Definition:

1) A linear operator T on a finite dimensional vector space V is called diagonalizable if there is an ordered basis B for V such that $[T]_B$ is a diagonal matrix.

2) A square matrix A is called diagonalizable if L_A (left multiplication transformation by matrix A) is diagonalizable.

12.4.2 Eigen vector and eigen value of a linear operator :

Definition: Let T be a linear operator an a vector space V. A non zero vector $v \in V$ is called an eigen vector of T if there exists a scalar λ such that $T(v) = \lambda V$. The scalar λ is called the eigen value of A corresponding to the eigen vector v.

12.4.3 Definition: Let A be in $M_{n\times n}(F)$. A non zero vector $V \in F^n$ is called an eigen vector of A if v is an eigen vector of L_A; that is if $Av = \lambda v$ for some scalar λ. The scalar λ is called the eigen value of A corresponding to the eigen vector v.

Note: i) The words characteristic vector, Latent Vectors, proper vector, spectral vector are also used in place of eigen vector.

ii) Eigen values are also known as characteristic values, Latent roots, proper values, spectral values.

iii) A vector is an eigen vector of a matrix A if and only if it is an eigen vector of L_A.

iv) A scalar λ is an eigen value of A, if and only if it is an eigen value of L_A.

In order to diagonalize a matrix or a linear operator we have to find a basis of eigen vectors and the corresponding eigen values.

Before continuing our study of diagonalization problem, we give the method of computing eigen values.

12.4.4 Method of Computing eigen Values:

Theorem: Let A be an $n \times n$ matrix in the entries in the field F. Then a scalar λ is an eigen value of A if and only if det $(A - \lambda I_n) = 0$

Proof: A scalar λ is an eigen value of A, if and only if there exists a nonzero vector $v \in F^n$ such that $Av = \lambda v$. That is $(A - \lambda I_n)v = 0$. This is true if and only if $A - \lambda I_n$ is not invertible. However this result is equivalent to the statement that det $(A - \lambda I_n) = 0$.

12.4.5 Characteristic matrix of the given matrix A:

Definition: Let A be a $n \times n$ matrix with entries in the field F. λ be a scalar, then the matrix $(A - \lambda I)$ is called the characteristic matrix of A.

12.4.6 Characteristic polynomial of A:

Definition: Let A be a $n \times n$ matrix with entries in the field F. Then the polynomial $f(\lambda) = \det(A - \lambda I)$ is called the characteristic polynomial of the square matrix A of order n - of degree n in λ. $f(\lambda)$ is the characteristic function of A.

Note: By theorem 12-4-4 it follows that the eigen values of a matrix are the zeros of its characteristic polynomial.

12.4.6A Characteristic polynomial of a linear operator:

Definition: Let T be a linear operator on a n - dimensional vector space V with an ordered basis B. We define the characteristic polynomial $f(\lambda)$ of T to be the characteristic polynomial of $A = [T]_B$. i.e. $f(\lambda) = \det(A - \lambda I_n)$.

Note: The characteristic polynomial of an operator T is defined by det $(T - \lambda I)$.

12.4.6B Characteristic Equation:

The equation $f(\lambda) = |A - \lambda I| = 0$ is called the characteristic equation of A.

Note : λ is a characteristic value of the matrix A if and only if det $(A - \lambda I) = 0$.

12.4.7 Definition: Spectrum: The set of all characteristic values of A is called the spectrum of A.

12.4.8 Short cut method:

Working procedure to find the characteristic polynomial of a matrix A.

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

The characteristic polynomial of A corresponding to the characteristic root λ is given by $\lambda^3 - tr(A).\lambda^2 + (A_{11} + A_{22} + A_{33})\lambda - \det A$. Where A_{11}, A_{22}, A_{33} denotes the cofactors of the diagonal elements a_{11}, a_{22} and a_{33} respectively. Observe that the coefficients of the characteristic polynomial of a 3×3 square matrix A are with alternating signs as follows

$S_1 = tr(A) =$ Sum of the terms of the principal diagonal.

$S_2 = A_{11} + A_{22} + A_{33}$ where

$$A_{11} = (-1)^{1+1}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}; A_{22} = (-1)^{2+2}\begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix}; A_{33} = (-1)^{3+3}\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$$

here each S_k is the sum of all principal minors of A of order k.

Note: i) If A is a n x n square matrix then its characteristic polynomial is

$$\lambda^n - S_1\lambda^{n-1} + S_2\lambda^{n-1} + ... + (-1)^n S_n$$

Where S_k is the sum of principal minors of order k.

ii) For a diagonal element of a square matrix, its minor and cofactor are the same.

12.4.9 Show that every square matrix need not posses eigen values.

Solution: Consider the matrix $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ over the field of reals. Its characteristic equation is $|A - \lambda I| = 0$.

i.e. det $\left\{\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} - \lambda\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right\} = 0$

$$\Rightarrow \det\begin{vmatrix}-\lambda & 1\\ -1 & -\lambda\end{vmatrix}=0 \Rightarrow \lambda^2+1=0$$

Which has no solution is the field of real numbers. So A has no characteristic value and hence no characteristic vector over the field of reals.

However if A is regarded as a complex matrix then its characteristic equations namely $\lambda^2+1=0$ has two distinct roots $i\,,\,-i$ over the field of complex numbers and consequently A has two distinct eigen vectors.

12.4.10 Theorem: Let $A \in M_{n\times n}(F)$. Then show that

i) The characteristic polynomial of A is a polynomial of degree n with leading coefficient $(-1)^n$. ii) A has atmost n distinct eigen values.

Proof: Let $A=\left[a_{ij}\right]_{n\times n}$ where the entries of A belongs to the field F.

The characteristic polynomial of A is given by $|A-\lambda I|$.

i.e. the characteristic polynomial of A is

$$\begin{vmatrix} a_{11}-\lambda & a_{12}\ldots\ldots\ldots & a_{1n}\\ a_{22} & a_{22-\lambda}\ldots\ldots\ldots & a_{2n}\\ a_{n+1} & a_{11}\ldots\ldots\ldots & a_{nn-\lambda}\end{vmatrix}$$

Expanding the determinant, we get the polynomial as $(-1)^n\left\{\lambda^n+a_1\lambda^{n-1}+a_2\lambda^{n-2}\ldots+a_n\right\}$ so the leading coefficient is $(-1)^n$ and as the polynomial is of degree n; it can not have more than n zeros. So A cannot have more than n eigen values.

Note: If $T:V\rightarrow V$ is a linear operator such that dim $V=n;$ then $[T]_B=A$ is a $n\times n$ matrix. So det $(A-\lambda I)$ is a polynomial of degree n. So A or T can not have more than n distinct eigen values.

1.4.11 Procedure to find eigen values and eigen vectors:

Let $A=\left[a_{ij}\right]_{n\times n}$ be the square matrix of order n.

Step 1:

Write the characteristic equation of A given by $|A-\lambda I|=0$. This is an equation of degree n in λ.

Step 2: Solve the equation $|A-\lambda I|=0$ to get n roots $\lambda_1, \lambda_2 ... \lambda_n$. Which are the eigen values of A.

Step 3: The corresponding eigen vectors of A are given by the nonzero vectors $V=[v_1, v_2, ..., v_n]$ satisfying the equation $AV = \lambda_i V$ or $(A-\lambda_i I)V = 0$ where $i=1,2....n$

12.4.12 Theorem: A linear operator T on a finite dimensional vector space V is diagonalizable if and only if there exists an ordered basis B for V consisting of eigen vectors of T. Further more if T is diagonalizable, $B=\{v_1, v_2, ..., v_n\}$ is an ordered basis of eigen vectors of T and $D=[T]_B$ then D is diagonal matrix and d_{jj} is the eigen value corresponding to V_j for $1 \le j \le n$.

Proof: Let V be a finite dimensional vector space and T be a linear operator on V. Suppose T is diagonalizable. Then there exists an ordered basis $B=\{v_1, v_2, ..., v_n\}$ for V such that $[T]_B$ is a diagonal matrix. Note that if $D=[T]_B$ is a diagonal matrix, then for each vector $v_j \in B$, we have

$$T(v_j) = \sum_{i=1}^{n} d_{ij} v_j = d_{jj} v_j = \lambda_j v_j \text{ where } \lambda_j = d_{jj}$$

Conversely if $B=\{v_1, v_2, ..., v_n\}$ is an ordered basis of V, such that $T(v_j) = \lambda_j v_j$ for some scalars $\lambda_1, \lambda_2 ... \lambda_n$, then $T(v_j) = 0v_1 + 0v_2 + ... + \lambda_j v_j + ... + 0v_n$

$$\text{Then clearly } [T]_B = \begin{bmatrix} \lambda_1 & 0........ & 0 \\ 0 & \lambda_2 & 0 \\ \vdots & \vdots & \vdots \\ 0 & 0 & \lambda_n \end{bmatrix} \text{ which is a diagonal matrix.}$$

In the preceeding paragraph, each vector v in the basis B satisfies the condition $T(v) = \lambda v$ for some scalar λ. Moreover, as v lies in a basis, v is non zero. Hence the theorem.

Note: To diagonalize a matrix or a linear operator we have to find a basis of eigen vectors and the corresponding eigen values.

12.5 Worked Out Exampes:

W.E. 1: Let $A=\begin{bmatrix} 1 & 3 \\ 4 & 2 \end{bmatrix}, B=\{v_1, v_2\}$ where $v_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$, $v_2 = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$ is an ordered basis of R^2. Prove that v_1, v_2 are eigen vectors of A. Find $[L_A]_B$. Show that A and $[L_A]_B$ are diagonalizable.

Solution: Given $A=\begin{bmatrix}1 & 3\\ 4 & 2\end{bmatrix}, v_1=\begin{bmatrix}1\\ -1\end{bmatrix}, v_2=\begin{bmatrix}3\\ 4\end{bmatrix}$

$$L_A(v_1)=\begin{bmatrix}1 & 3\\ 4 & 2\end{bmatrix}\begin{bmatrix}1\\ -1\end{bmatrix}=\begin{bmatrix}1 & -3\\ 4 & -2\end{bmatrix}_{2\times 1}=\begin{bmatrix}-2\\ 2\end{bmatrix}=-2\begin{bmatrix}1\\ -1\end{bmatrix}=-2v_1$$

Where $L_A=T$.

So v_1 is an eigen vector of L_A and hence v_1 is an eigen vector of A. Here $\lambda_1=-2$ is the eigen. value corresponding to v_1.

Further more $L_A(v_2)=\begin{bmatrix}1 & 3\\ 4 & 2\end{bmatrix}\begin{bmatrix}3\\ 4\end{bmatrix}=\begin{bmatrix}3+12\\ 12+8\end{bmatrix}=\begin{bmatrix}15\\ 20\end{bmatrix}=5\begin{bmatrix}3\\ 4\end{bmatrix}=5v_2$

So v_2 is an eigen vector of L_A and so v_2 is an eigen vector of A; with the corresponding eigen value $\lambda_2=5$. Note that $B=\{v_1,v_2\}$ is an ordered basis of R^2 consisting of eigen vectors of both A and L_A and so A and L_A are diagonalizable.

Further more $[L_A]_B=\begin{bmatrix}-2 & 0\\ 0 & 5\end{bmatrix}$ is a diagonal matrix.

W.E. 2: Determine the eigen values and eigen vectors of the matrix $A=\begin{bmatrix}5 & 4\\ 1 & 2\end{bmatrix}$.

Solution: The characteristic equation of A is $|A-\lambda I|=0$.

$$\Rightarrow\begin{vmatrix}5-\lambda & 4\\ 1 & 2-\lambda\end{vmatrix}=0$$

$$\Rightarrow (5-\lambda)(2-\lambda)-4=0\Rightarrow \lambda^2-7\lambda+6=0$$

$$\Rightarrow (\lambda-6)(\lambda-1)=0\Rightarrow \lambda=6,1$$

So the eigen values of A are 6, 1.

To find the eigen vector corresponding to $\lambda=6$:

The eigen vector $v' = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ of A Corresponding to the eigen value 6 are given by the non zero solution of the equation $(A-6I)v' = O$

$$\Rightarrow \begin{bmatrix} 5-6 & 4 \\ 1 & 2-6 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = O$$

$$\Rightarrow \begin{bmatrix} -1 & 4 \\ 1 & -4 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = O = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$R_2 + R_1$ gives

$$\begin{bmatrix} -1 & 4 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$\Rightarrow -v_1 + 4v_2 = 0 \Rightarrow v_1 = 4v_2$ putting $v_2 = 1$ we get $v_1 = 4$.

So $v' = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$ is an eigen vector of A; Corresponding to the eigen value 6.

The set of all eigen values of A corresponding to the eigen value 6 is given by $C_1 v'$ where C_1 is a nonzero scalar.

The eigen value v of A corresponding to the eigen value 1 are given by the non zero solution of the equation $(A-\lambda I)v = O$

$$\Rightarrow (A-1I)v = O$$

$$\Rightarrow \begin{bmatrix} 4 & 4 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$4R_2 - R_1$ gives $\begin{bmatrix} 4 & 4 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

$$\Rightarrow 4v_1 + v_2 = 0 \Rightarrow v_1 = -v_2$$

Let $v_1 = 1$ them $v_2 = -1$

So $v'' = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ is an eigen vector of A corresponding to the eigen value 1. Every non zero multiple of v''. Which is of the form $C_2 \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ where $C_2 \neq 0$ is an eigen vector corresponding to the eigen value 1.

W.E. to 3: $T : R^3(R) \to R^3(R)$ is a linear operator defined by $T(a,b,c) = (7a - 4b + 10c, 4a - 3b + 8c, -2a + b - 2c)$ find the eigen values of T and an ordered basis B for R^3(R) such that $[T]_B$ is a diagonal matrix.

Sol: Gives $T : R_3 \to R^3$ is defined by $T(a,b,c) = (7a - 4b + 10c, 4a - 3b + 8c, -2a + b - 2c)$. Consider the usual ordered basis $B = \{e_1, e_2, e_3\}$ for R^3 where $e_1 = (1,0,0), e_2(0,1,0), e_3(0,0,1)$

$$T(e_1) = T(1,0,0) = (7(1) - 4(0) + 10(0), 4(1) - 3(0) + 8(0) - 2(1) + 0 - 2(0))$$

$$= (7, 4, -2)$$

$$= 7(1,0,0) + 4(0,1,0) - 2(0,0,1)$$

$$= 7e_1 + 4e_2 - 2e_3$$

$$T(e_2) = T(0,1,0) = (-4 - 3, 1)$$

$$= -4(1,0,0) - 3(0,0,0) + 1(0,0,0)$$

$$= -4e_1 - 3e_2 + 1e_3$$

$$T(e_3) = T(0,1,0) = (10; 8, -2)$$

$$= 10e_1 + 8e_2 - 2e_3$$

$$A[T]_B = \begin{bmatrix} 7 & -4 & 10 \\ 4 & -3 & 8 \\ -2 & 1 & -2 \end{bmatrix} \quad \text{........... (1)}$$

If λ is an eigen value of A; then $|A - \lambda I| = 0$

or $\Rightarrow \begin{vmatrix} 7-\lambda & -4 & 10 \\ 4 & -3-\lambda & 8 \\ -2 & 1 & -2-\lambda \end{vmatrix} = 0$

From (1)

trace of $A = 7-3-2 = 2$

$A_{11} = (-1)^{1+1}$ minor of $7 = \begin{vmatrix} -3 & 8 \\ 1 & -2 \end{vmatrix} = 6-8 = -2$

$A_{22} = (-1)^{2+2}$ minor of $(-3) = \begin{vmatrix} 7 & 10 \\ -2 & -2 \end{vmatrix} = -14+20 = 6$

$A_{33} = (-1)^{3+3}$ minor of $(-2) = \begin{vmatrix} 7 & -4 \\ 4 & -3 \end{vmatrix} = -21+16 = -5$

Their Sum $= A_{11} + A_{22} + A_{33} = (-2+6-5) = -1$

det $A = 7(6-8) + 4(-8+16) + 10(4-6)$

$= 7(-2) + 4(8) + 10(-2) = -2$

The characteristic equation is

$\lambda^3 - \lambda^2$ trace of A $+ \lambda(A_{11} + A_{22} + A_{33}) - \det A = 0$

$\lambda^3 - 2\lambda^2 - 1\lambda + 2 = 0 = f(\lambda)$ say

$f(1) = 1^3 - 2(1)^2 - 1 + 2 = 0$ $\quad \therefore \lambda - 1$ is a factor.

$$\begin{array}{r|l} 1 & 1 \; -2 \; -1 \; 2 \\ & \;\;\; 1 \; -1 \; -2 \\ \hline & 1 \; -1 \; -2 \;\; | \; 0 \end{array}$$

The other factor is $\lambda^2 - \lambda - 2 = (\lambda-2)(\lambda+1)$

Hence the characteristic equation is $(\lambda-1)(\lambda-2)(\lambda+1) = 0$

So characteristic values are $-1, 1, 2$.

ii) To find the characteristic vector corresponding to $\lambda = -1$.

$$(A - \lambda I)v = O$$

$$\Rightarrow \begin{bmatrix} 7+1 & -4 & 10 \\ 4 & -3+1 & 8 \\ -2 & 1 & -2+1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 8 & -4 & 10 \\ 4 & -2 & 8 \\ -2 & 1 & -1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$2R_2 - R_1, 4R_3 + R_1$ gives

$$\begin{bmatrix} 8 & -4 & 10 \\ 0 & 0 & 6 \\ 0 & 0 & 6 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$R_3 - R_2$ gives

$$\begin{bmatrix} 8 & -4 & 10 \\ 0 & 0 & 6 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$6v_3 = 0 \Rightarrow v_3 = 0$

$8v_1 - 4v_2 + 10v_3 = 0 \Rightarrow 8v_1 = 4v_2 \Rightarrow v_2 = 2v_1$

put $v_1 = 1$ then $v_2 = 0$

So $v' = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$ and every scalar multiple of it is an eigen vector.

ii) To find the eigen vector corresponding to $\lambda = 1$

$$(A - \lambda I)v = 0$$

$$\Rightarrow \begin{bmatrix} 7-1 & -4 & 10 \\ 4 & -3-1 & 8 \\ -2 & 1 & -2-1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 6 & -4 & 10 \\ 4 & -4 & 8 \\ -2 & 1 & -3 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$R_1 \times \frac{1}{2}; R_2 \times \frac{1}{4}$ gives

$$\begin{bmatrix} 3 & -2 & 5 \\ 1 & -1 & 2 \\ -2 & 1 & -3 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$R_1 - 3R_2, R_3 + 2R_2$ gives

$$\begin{bmatrix} 0 & 1 & -1 \\ 1 & -1 & 2 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$R_3 + R_1$ gives $\begin{bmatrix} 0 & 1 & -1 \\ 1 & -1 & 2 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

$$v_2 - v_3 = 0 \Rightarrow v_2 = v_3$$

$$v_1 - v_2 + 2v_3 = 0 \Rightarrow v_1 + v_3 = 0 \Rightarrow v_3 = -v_1$$

putting $v_1 = 1$, $v_3 = v_2 = -1$

So $v'' = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix}$ and every scalar multiple of it is an eigen vector.

iii) To find the eigen vector corresponding to $\lambda = 2$.

$$(A - \lambda I)v = 0$$

$$\Rightarrow \begin{bmatrix} 7-2 & -4 & 10 \\ 4 & -3-2 & 8 \\ -2 & 1 & -2-2 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 5 & -4 & 10 \\ 4 & -5 & 8 \\ -2 & 1 & -4 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$R_1 \leftrightarrow R_3$ gives $\begin{bmatrix} -2 & 1 & -4 \\ 4 & -5 & 8 \\ 5 & -4 & 10 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

$R_2 + 2R_1, 2R_3 + 5R_1$ gives

$$\begin{bmatrix} -2 & 1 & -4 \\ 0 & -3 & 0 \\ 0 & -3 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$R_3 - R_2$ gives

$$\begin{bmatrix} -2 & 1 & -4 \\ 0 & -3 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Rightarrow -3v_2 = 0 \Rightarrow v_2 = 0$$ and

$$-2v_1 + v_2 - 4v_3 = 0 \Rightarrow 2v_1 = -4v_3 \Rightarrow v_1 = -2v_3$$

Put $v_3 = -1$, then $v_1 = 2, v_2 = 0, v_3 = -1$ and so $v''' = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ -1 \end{bmatrix}$ and every non zero scalar multiple of it is an eigen vector.

The basis for which $[T]_B$ is a diagonal matrix is $B^{'} = \{v^{'}, v^{''}, v^{'''}\}$

i.e. $B^{'} = \{(1,2,0),(1,-1,-1)(2,0,-1)\}$

W.E. 4: Let T be the linear operator on $P_2(R)$ defined by $T[f(x)] = f(x)+(x+1)f^{'}(x)$ find the eigen values of T.

Solution: $B = \{1, x, x^2\}$ is the standard Basis of $P_2(R)$. We are given linear operator on $P_2(R)$ defined by $T[f(x)] = f(x)+(x+1)f^{'}(x)$

$$T(1) = 1+(x+1)(0) = 1 = 1(1)+0x+0x^2$$

$$T(x) = x+(x+1)(1)$$

$$= 2x+1 = 1(1)+2.x+0x^2$$

$$T(x^2) = x^2+(x+1)(2x)$$

$$= 2x+3x^2 = 0(1)+2.x+3.x^2$$

So $A = [T]_B = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 2 \\ 0 & 0 & 3 \end{bmatrix}$

The characteristic equation is $|A-\lambda I| = 0$

$$\Rightarrow \begin{vmatrix} 1-\lambda & 1 & 0 \\ 0 & 2-\lambda & 2 \\ 0 & 0 & 3-\lambda \end{vmatrix} = 0$$

$\Rightarrow (1-\lambda)[(2-\lambda)(3-\lambda)] = 0$ (expanding along first column)

$$\Rightarrow (1-\lambda)(2-\lambda)(3-\lambda) = 0$$

$$\Rightarrow \lambda = 1, 2, 3$$

12.6 Properties of eigen values:

12.6.1 Theorem: Let T be a linear operator on a vector space V; and let λ be an eigen value of T.

A vector $v \in V$ is an eigen vector of T corresponding to λ if and only if $v \neq 0$ and $v \in N(T-\lambda I)$, the null space of the linear operator $(T-\lambda I)$.

Proof: Let v is the characteristic vector corresponding to the characteristic value λ.

So $T(v) = \lambda v \Rightarrow (T-\lambda I)v = 0$

So $v \in$ the null space of $T-\lambda$ i.e. $N(T-\lambda I)$

Conversely Let $v \in$ null space of $T-\lambda I$.

$$\Rightarrow (T-\lambda I)v = 0$$

$$\Rightarrow Tv = \lambda v$$

So v is the characteristic vector corresponding to the characteristic value λ.

Hence the Theorem.

12.6.2 Theorem: Prove that a square matrix A; and its transport A^T have the same set of eigen values.

Proof: Characteristic polynomial of $A = \det(A-\lambda I)$

$= \det(A-\lambda I)^T$ since the determinant of a matrix and its transpose are equal.

$= \det\left[A^T - (\lambda I)^T\right]$

$= \det\left[A^T - \lambda I^T\right]$

$= \det\left[A^T - \lambda I\right]$ since $I^T = I$

= Characteristic polynomial of A^T.

So A and A^T have the same characteristic polynomial and hence the same set of eigen values.

12.6.3 Show that zero is a characteristic root of a matrix if and only if the matrix is singular.

Solution: 0 is a characteristic value of A.

$\Leftrightarrow \lambda = 0$ satisfies the equation $|A-\lambda I| = 0$

$\Leftrightarrow |A-0I| = 0 \Leftrightarrow A$ is singular.

Note : i) If λ is a characteristic root of a non singular matrix then $\lambda \neq 0$.

ii) At least one characteristic root of every singular matrix is zero.

12.6.4 If λ is a characteristic root of a matrix A; K is a scalar, show that $K+\lambda$ is a characteristic root of the matrix $A+KI$.

Solution: Let λ be a characteristic root of a matrix A, and v be the corresponding characteristic vector.

Then $Av = \lambda v$ and $(A+KI)v = Av + K(Iv)$

$$= \lambda v + Kv$$

$$= (\lambda + K)v$$

Since $v \neq 0$ $\lambda + K$ is a characteristic root of the matrix $A+KI$ and v is a corresponding of characteristic vector.

12.6.5 If $\lambda_1, \lambda_2 \ldots \lambda_n$ are characteristic roots of a n x n matrix A; k is a scalar then the charecteristic roots of $A-kI$ are $\lambda_1 - k, \lambda_2 - k, \ldots, \lambda_n - k$.

Solution: Given $\lambda_1, \lambda_2 \ldots\ldots \lambda_n$ are the charecteristic roots of A.

So the characteristic polynomial of A is $|A-\lambda I|$

$$= (\lambda_1 - \lambda)(\lambda_2 - \lambda)\ldots(\lambda_n - \lambda)$$

The characteristic polynomial of $A+kI$ is $|A-kI-\lambda I| = |A-(k+\lambda)I|$

$$= [\lambda_1 - (k+\lambda)][\lambda_2 - (k+\lambda)]\ldots[\lambda_n - (k+\lambda)]$$

$$= [(\lambda_1 - k) - \lambda)][(\lambda_2 - k) - \lambda)]\ldots[(\lambda_n - k) - \lambda)]$$

Hence the characteristic roots of $A-kI$ are $\lambda_1 - k, \lambda_2 - k; \lambda_3 - k, \ldots, \lambda_n - k$.

12.6.6 If A is non singular prove that the eigen values of A^{-1} are the reciprocals of the eigen values of A.

Solution: Let λ be an eigen value of A and v be the corresponding eigen vector then $Av = \lambda v$.

$$\Rightarrow v = A^{-1}(\lambda v) = \lambda(A^{-1}v)$$

$\Rightarrow \frac{1}{\lambda} v = A^{-1} v$ (Since A is non singular $\lambda \neq 0$)

$\Rightarrow A^{-1} v = \frac{1}{\lambda} v$

So $\frac{1}{\lambda}$ is an eigen value of A^{-1}, and v is the corresponding eigen vector.

Conversely suppose that λ is an eigen value of A^{-1}. Since A is non singular A^{-1} is also non singular and $(A^{-1})^{-1} = A$. So it follows from the first part of this question $\frac{1}{\lambda}$ is an eigen value of A.

Thus each eigen value of A^{-1} is equal to the reciprocal of some eigen value of A.

Hence the eigen values of A^{-1} are nothing put the reciprocals of the eigen values of A.

12.6.7 Corollary: If $\lambda_1, \lambda_2 ... \lambda_n$ are the eigen values of a non singular matrix A, then $\lambda_1^{-1}, \lambda_2^{-1} ... \lambda_n^{-1}$ are the eigen values of A^{-1}.

Solution: The solution follows from the above proof.

12.6.8 Theorem: If $\lambda_1, \lambda_2, ... \lambda_n$ are the eigen values of A, then $K\lambda_1, K\lambda_2, ... K\lambda_n$ are the eigen values of KA.

Proof: If $K = 0$ then $KA = O$ and each eigen value of 0 is 0. Thus $0\lambda_1, 0\lambda_2, ... 0\lambda_n$ are the eigen values of KA when $\lambda_1, \lambda_2, ... \lambda_n$ are eigen values of A.

So let us suppose that $K \neq 0$

We have $|KA - \lambda KI| = |K(A - \lambda)I|$

$$= K^n |A - \lambda I| \text{ Since } |KB| = K^n |B|$$

If $K \neq 0$, then $|KA - \lambda KI| = 0$ if and only if $|A - \lambda I| = 0$.

i.e. $K\lambda$ is an eigen value of KA; if and only if λ is an eigen value of A.

Thus if $\lambda_1, \lambda_2 ..., \lambda_n$ are eigen values of A, that $K\lambda_1, K\lambda_2 ..., K\lambda_n$ are eigen values of KA.

12.6.9 Corollary: Let $0 \neq \lambda$ be an eigen value of an invertible operator T. Show that λ^{-1} is an eigen value of T^{-1}.

Solution: As λ is an eigen value of T, there exists non zero $v \in V$ such that $T(v) = \lambda v$.

$$\Rightarrow v = T^{-1}(\lambda v) = \lambda T^{-1}(v)$$

So $\lambda^{-1} v = \lambda^{-1} \lambda T^{-1}(v)$

$$\Rightarrow \lambda^{-1} v = T^{-1}(v)$$

Hence λ^{-1} is an eigen value of T^{-1}.

12.6.10 If λ is a characteristic root of a non singular matrix A, Show that λ^r is the characteristic root of $A^r; r$ being an integer.

Solution: λ is a characteristic root of a non singular matrix. So $\lambda \neq 0$.

Case i) Let $r > 0$. Let v be a characteristic vector corresponding to λ then $Av = \lambda v$.

So $A^r v = A^{r-1}(Av) = A^{r-1}(\lambda v) = \lambda (A^{r-1} v)$

$$= \lambda A^{r-2}(Av)$$

$$= \lambda A^{r-2}(\lambda v)$$

$$= \lambda^2 A^{r-2} v$$

Proceeding like this we get

$$A^r v = \lambda^r v$$

Thus λ^r is a characteristic root of A^r and v is also a characteristic vector of A^r.

Case ii) Let $r = 0$ then $A^0 = I$ and characteristic roots of I are all unity i.e. λ^0.

Case iii) Let $r = -1$ then $Av = \lambda v \Rightarrow A^{-1}(Av) = A^{-1}(\lambda v)$

So $Iv = \lambda (A^{-1} v)$

$$\Rightarrow \lambda^{-1}(Iv) = \lambda^{-1} \lambda (A^{-1} v)$$

$\Rightarrow \lambda^{-1} v = A^{-1} v$ since $\lambda^{-1} \lambda = 1$

$\Rightarrow \lambda^{-1}$ is a characteristic root of A^{-1}.

i.e. when $r = -1$, λ^r is a characteristic root of A^r.

Case iv) Let r be negative say $r = -s$ where s is a positive integer then $A^r = A^{-s} = (A^{-1})^s$.

By Case iii) λ^{-1} is a characteristic root of A^{-1}.

By case i) $(\lambda^{-1})^s$ is a characteristic root of $(A^{-1})^s$.

$\therefore \lambda^r$ is a characteristic root of A^r.

Hence the theorem.

12.6.11 Corollary: If $\lambda_1, \lambda_2, ...\lambda_n$ are the characteristic roots of A, then the characteristic roots of A^2 are $\lambda_1^2, \lambda_2^2, \ldots\ldots, \lambda_n^2$.

Solution: The solution follows from the above theorem.

12.6.12 Let T be a linear operator on a vector space V. let v be an eigen vector of T corresponding to the eigen value λ. For any positive integer n prove that v is an eigen vector of T^n corresponding to the eigen value λ^n.

Solution: Given v is an eigen vector of the linear operator T. So $Tv = \lambda v$ where $v \neq O$. We have to prove $T^n(v) = \lambda^n v$, n being a positive integer. We prove the result by induction. For $n = 1$, the result is true because of (1).

Let the result be true for a positive integer m

i.e. $T^m(v) = \lambda^m v$ .... (1)

Thus $T^{m+1}(v) = T^m(Tv) = (T^m T)v = T^m(Tv)$

$= T^m(\lambda v)$

$= \lambda(T^m v)$ Since T is linear

$= \lambda(\lambda^m v)$ by (3)

So $T^{m+1}(v) = (\lambda\lambda^m)(v) = \lambda^{m+1}(v)$ so the statement is true for $m+1$

The statement is true for $n = 1$, when it is assumed to be true for m; it is proved to be true for $m+1$. Hence by mathematical induction the statement is true for all positive integral values of n.

12.6.13 Let T be a linear operator an a vector space V; over a field F and let $g(x)$ be a polynomial with coefficients from F. Prove that if v is an eigen vector of T with corresponding eigen value λ, then $g(T)(v) = g(\lambda)v$ i.e. v is an eigen value of $g(T)$ with the corresponding eigen values of $g(\lambda)$.

Proof:

Let $g(x) = a_0 + a_1 x + a_2 x^2 + ... + a_m \lambda^m$ where $a_i \in F$

Let $T(v) = \lambda v$.

$\Rightarrow T^2(v) = \lambda^2 v$. In general $T^r(v) = \lambda^r(v)$

Now $g(T) = a_0 I + a_1 T + a_2 T^2 + ... + a_m T^m$

$$(g(T))(v) = a_0 v + a_1 T(v) + a_2 T^2(v) + ... + a_m T^m(v)$$

$$= a_0 v + a_1 \lambda(v) + a_2 \lambda^2(v) + ... + a_m \lambda^m(v)$$

$$= (a_0 + a_1 \lambda + a_2 \lambda^2 + ... + a_m \lambda^m)v$$

$$= g(\lambda)v$$

Hence the theorem.

12.6.14 Theorem: Let A be n x n triangular matrix over F. Prove that the eigen values of A are diagonal elements.

Let $A = \begin{bmatrix} a_{11} & a_{12} & a_{13}....... & a_{1n} \\ 0 & a_{22} & a_{23}........... & a_{24} \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0......... & a_{nn} \end{bmatrix}$

i.e.. $A = \left[a_{ij}\right]_{n \times n}$ where $a_{ij} = 0$ for $i > j$

Characteristic equation of A is $|A - \lambda I| = 0$

$$\Rightarrow \begin{vmatrix} a_{11} - \lambda & a_{12} & a_{13}........ & a_{1n} \\ 0 & a_{22} - \lambda & a_{23}........ & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & a_{nn} - \lambda \end{vmatrix} = 0$$

$$\Rightarrow (a_{11} - \lambda)(a_{22} - \lambda)...(a_{nn} - \lambda) = 0$$

So the eigen values of A are $a_{11}, a_{22}, ..., a_{nn}$

So in a triangular matrix, the eigen values are the diagonal elements of the matrix.

12.6.5 Corollary: Show that the characteristic values of a diagonal matrix D are the elements in the diagonal.

Proof: Let $D = \begin{bmatrix} a_{11} & 0 & 0\ldots\ldots\ldots & 0 \\ 0 & a_{22} & 0\ldots\ldots\ldots & 0 \\ 0 & 0 & 0\ldots\ldots\ldots & a_{nn} \end{bmatrix}_{n\times n}$

Characteristic equation of D is $|D-\lambda I| = 0$.

$$\Rightarrow \begin{vmatrix} a_{11}-\lambda & 0 & 0\ldots\ldots\ldots & 0 \\ 0 & a_{22}-\lambda & 0\ldots\ldots\ldots & 0 \\ \vdots & \vdots & & \\ 0 & 0 & 0\ldots\ldots\ldots & a_{nn}-\lambda \end{vmatrix} = 0$$

$$\Rightarrow (a_{11}-\lambda)(a_{22}-\lambda)\ldots(a_{nn}-\lambda) = 0$$

Which shows the characteristic values are $a_{11}, a_{22}, \ldots\ldots, a_{nn}$. Which are nothing but the elements in the diagonal.

12.7 Similarity:

12.7.1 Definition: i) Two $n \times n$ matrices A and B are said to be similar if there exists a non singular matrix P such that $AP = PB$ or $A = PBP^{-1}$

Definition II : Two linear operators T_1 and T_2 on V are said to be similar if there exists a nonsingular linear operator T on V such that $T_1T = TT_2$ or $T_1 = TT_2T^{-1}$

12.7.2 Show that similar matrices have the same characteristic polynomial and hence the same eigen value.

Proof: Let A and B are any two similar matrices then for a invertible matrix P, we have $B = P^{-1}AP$.

Let $\det(B-\lambda I) = \det(P^{-1}AP - \lambda I)$

$$= \det\left(P^{-1}(A-\lambda I)P\right)$$

$$= \det(P^{-1})\det(A-\lambda I)\det(P) = \det(P^{-1})\det(P)\det(A-\lambda I)$$

$$= \det(P^{-1}P)\det(A-\lambda I)$$

$$= \det(I)\det(A-\lambda I)$$

$$= 1.\det(A-\lambda I)$$

This shows that the matrices A and B have the same characteristic polynomial. Hence A and B have the same characteristic roots.

12.7.3 Corollary: A square matrix B is similar to a diagonal matrix D. Show that the characteristic roots of B are diagonal elements of D.

Proof: Let the D be the diagonal matrix of order n. B is similar to D.

The characteristic roots of D are the elements along the principal diagonal of D.

By the above theorem, B and D have the same charecteristic equation and hence the same characteristic roots.

So the characteristic roots of B are the diagonal elements of D.

Hence the Theorem.

12.8 Similarity of matrices using Trace:

12.8.1 Theorem: Let A and B are two square matrices of order n. Thun show that trace of $(AB) =$ trace of (BA)

Proof: Let $A=\left[a_{ij}\right]_{n\times n}\ B=\left[b_{ij}\right]_{n\times n}$

$AB=\left[C_{ij}\right]_{n\times n}$ where $C_{ij}=\sum_{k=1}^{n} a_{ik}b_{kj}$

$BA=\left[d_{ij}\right]_{n\times n}$ where $d_{ij}=\sum_{k=1}^{n} b_{ik}a_{kj}$

trace of $(AB)=\sum_{i=1}^{n} C_{ii}=\sum_{c=1}^{n}\left(\sum_{k=1}^{n} a_{ik}b_{ki}\right)$

$$=\sum_{k=1}^{n}\ \sum_{k=1}^{n} a_{ik}b_{ki}$$

Interchanging the order of summation in the last sum.

$$=\sum_{k=1}^{n}\ \left(\sum_{i=1}^{n} b_{ki}a_{ik}\right)$$

$$=\sum_{k=1}^{n} d_{kk} = d_{11} + d_{22} + ... + d_{nn}$$

So trace of (AB) = trace of (BA)

Hence the theorem.

12.8.2 Theorem: Prove that similar matrices have the same trace.

Proof: Let A and B are n x n matrices over F such that A is similar to B. We have to show that trace of A = trace of B.

Let $A=\left[a_{ij}\right]_{n\times n}$ $B=\left[b_{ij}\right]_{n\times n}$. Let A be similar to B; Then there exists a non singular matrix P. such that $A=P^{-1}BP$.

trace of A = trace of $(P^{-1}BP)$

= trace of $(PP^{-1}B)$ since trace of AB is equal to trace of BA

= trace of (IB)

= trace of B.

Hence similar matrices have the same trace.

12.9 Trace of Matrix : The Sum of the elements of a square matrix A lying along the principal diagonal is called the trace of the matrix. If $A=\left[a_{ij}\right]_{n\times n}$ then trace of $A=a_{11}+a_{22}+...+a_{nn}$.

Definition: Trace of a linear operator :

Let T be a linear operator from $V \to V$. Then the trace of T written as tr T is the trace of $M(T)$ where $M(T)$ is the matrix of T in some basis of V.

12.9.1 To show that the definition of trace of a linear operator is well defined:

To show that the trace of linear operator is independent of the basis of V.

Solution: Let $M_1(T); M_2(T)$ are the matrices of T in two different bases of V. We know if T is a linear operator from a n dimensional vector space V to V; over a field F, and T has the matrix $M_1(T)$ in the basis $\{v_1, v_2, ...v_n\}$ and the matrix $M_2(T)$ in the basis $\{w_1, w_2, ...w_n\}$ then there exists a non singular matrix P of order n such that $M_2(T)=P^{-1}M_1(T)P$.

Hence by this theorem, there exists a non singular matrix P such that $M_2(T)=P^{-1}M_1(T)P$.

i.e. $M_1(T)$ and $M_2(T)$ are similar matrices. But similar matrices have the same trace. Hence trace of T does not depend upon any particular basis of V.

Hence the above definition is meaningful. So trace depends only an T and not any particular basis.

W.E. 5: Example: Let $T: R_2 \to R_2$ be the linear transformation defined by $T(x, y) = (2y, 3x - y)$.

Then in the basis $\{(1,0)(0,1)\}$ the matrix of T is $\begin{bmatrix} 0 & 2 \\ 3 & -1 \end{bmatrix}$. So trace of $T = 0 - 1 = -1$.

Also in the basis $\{(1,3)(2,5)\}$ matrix of T is $\begin{bmatrix} -30 & -48 \\ 18 & 29 \end{bmatrix}$

trace of $T = -30 + 29 = -1$

From which we observe trace of T is not dependent on the basis of V.

12.10 Determinant of a linear operator T on V(F):

12.10.1 Definition: Let T be a linear operator on a vector space $V(F)$ and $[T]$ or $[T]_B$ be the matrix of this linear operator relative to a basis B; then $\det T = \det[T]$.

12.10.2 Theorem: Prove that the determinant of a linear operator T on a vector space is unique.

Or

Prove that the determinant of a linear operator is independent of the choice of an ordered basis for V.

Proof:

Let $[a_{ij}]$ and $[b_{ij}]$ are the matrices of the linear operator T with respect to the basis B_1 and B_2 of V.

i.e. $[T : B_1] = [a_{ij}]$; and $[T : B_2] = [b_{ij}]$

Then there exists an invertible matrix $[c_{ij}]$ such that $[b_{ij}] = [c_{ij}]^{-1}[a_{ij}][c_{ij}]$

$$\det[b_{ij}] = \det\left[[c_{ij}]^{-1}[a_{ij}][c_{ij}]\right]$$

$$= \det[b_{ij}] = \det\left[[c_{ij}]^{-1}[a_{ij}][c_{ij}]\right]$$

$$= \det\left[c_{ij}\right]^{-1}\det\left[c_{ij}\right]\det\left[a_{ij}\right]$$

i.e. $$\det\left[b_{ij}\right] = \det\left[\left[c_{ij}\right]\left[c_{ij}\right]^{-1}\right]\det\left[a_{ij}\right]$$

$$= \det(I)\det\left[a_{ij}\right]$$

$$= 1.\det\left[a_{ij}\right]$$

i.e. $$\det\left[b_{ij}\right] = \det\left[a_{ij}\right]$$

Hence the determinant of the operator T is unique even though the matrices of T are different with respect to the bases B_1 and B_2.

12.10.3 Theorem: If T_1 and T_2 are linear operators on a finite dimensional vector space $V(F)$, then prove that $\det(T_1T_2) = (\det T_1)(\det T_2)$.

Proof: T_1 and T_2 are two linear operators on a finite dimensional vector space V(F). Choose B to be an ordered basis of v. Then the matrix of the operator T_1T_2 w.r.t the basis B can be put in the form $[T_1T_2]_B = [T_1]_B.[T_2]_B$.

So det $[T_1T_2]_B = \det\left([T_1]_B.[T_2]_B\right)$

$$= \det[T_1]_B.\det[T_2]_B \quad \text{............ (1)}$$

As we know the det. of the product of two matrices is equal to the product of their determinants.

Now by the definition $\det T = \det[T]_B$ and hence by (1) we have $\det T_1T_2 = \left(\det T_1\right)\left(\det T_2\right)$

12.10.4 T is a linear operator an a finite dimensional vector space V. Prove that T is invertible if and only if det $(T) \neq 0$.

Proof: T is a linear operator on a finite dimensional vector space V. Let B be a basis of the vector space V.

If T is invertible then $TT^{-1} = T^{-1}T = I$ and $\det(T^{-1}T) = \det[I]_B$ where $[I]_B$ is matrix of the identity operator.

$\Rightarrow \det(T).\det(T^{-1}) = 1$ since det. of the unit matrix is 1.

Now $(\det T)$ and $(\det T^{-1})$ are the elements of the field F and a field F is without zero divisors. i.e.

$ab = 0 \Rightarrow a = 0$ or $b = 0$ or both zero.

In other words if $a = 0$ then $ab = 0$ in a field F.

Hence since $(\det T).(\det T^{-1}) = 1$ hence $\det T \neq 0$.

Conversely let $\det T \neq 0$ then by def. of $\det T$ we have $\det [T]_B \neq 0$

$\rightarrow [T]_B$ is invertible.

Hence the operator T is invertible.

12.10.5 Theorem: T is a linear operator on a finite dimensional vector space V. Prove that T is invertible then $\det(T^{-1}) = (\det T)^{-1}$.

Solution: T is invertible then $TT^{-1} = I = T^{-1}T$

$$\det(TT^{-1}) = \det(I) = \det(T^{-1}T)$$

$$\Rightarrow (\det T).(\det T^{-1}) = 1 = \left(\det(T^{-1})\right)(\det T)$$

So $(\det T^{-1}) = \left[\det(T)\right]^{-1}$

12.10.6 Let T be a linear operator on a finite dimensional vector space V. Then show that O is the characteristic value of T iff T is not invertible.

Solution: Case i) Let O be the eigen value of T. Then we have to prove T is singular.

As O is a eigen value of T, there exists a nonzero v in V such that $T(v) = 0v$.

$\Rightarrow T(v) = O$ (zero vector)

$\Rightarrow T$ is singular so T is not invertible.

Case ii) Converse:

Suppose T is not invertible we have to show O is the eigen value of T.

As T is a linear operator on a finite dimensional vector space V,and T is not invertible means T is singular So there exists a nonzero vector v in V such that $Tv = O = 0v$.

$\therefore O$ is the characteristic vector.

Hence the theorem.

12.10.7 Corollary:

Prove that a linear operator on a finite vector space is invertible if and only zero is not an eigen value of T.

Exercise 12.11:

T is linear operator on R^3 defined by

$$T\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{bmatrix} 3a+2b-2c \\ -4a-3b+2c \\ -c \end{bmatrix} \text{ and } B = \left\{ \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} \right\}$$ is an ordered basis of R^3. Compute $[T]_B$ and determine whether B is a basis of consisting of eigen vectors of T.

Ans) $[T]_B = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$; yes

2). T is a linear operator on R^2 defined by $T(a,b) = (-2a+3b, -10a+9b)$. Find the eigen values of T and an ordered basis B for V such that $[T]_B$ is a diagonal matrix.

Ans) $\lambda = 3,4, \quad B = \{(3,5),(1,2)\}$

3) T is a linear operator on $P_3(R)$ defined by $T[f(x)] = f(x) + f(2)x$. Find the eigen values of T and an ordered basis B for V such that $[T]_B$ is a diagonal matrix.

Ans : $\lambda = 1,3 \quad B = \{-2+x, -4+x^2, -8+x^3, x\}$

4) Find the eigen values of the matrix $\begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & -1 \\ 2 & -1 & 0 \end{bmatrix}$

Ans : 2, $-1+\sqrt{3}, \quad -1-\sqrt{3}$

5) Find the characteristic polynomial of $A = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 3 & 2 \\ 1 & 3 & 9 \end{bmatrix}$

Ans : $\lambda^3 - 13\lambda^2 + 31\lambda - 17$

6) If $A = \begin{bmatrix} 4 & 1 & -1 \\ 2 & 5 & -2 \\ 1 & 1 & 2 \end{bmatrix}$ then find i) All eigen values of A.

ii) A maximum set S of linearly independent vectors of A.

Ans: i) $\lambda = 3,3,5$ ii) $\{(1,-1,0),(1,0,1),(1,2,1)\}$ is a maximal set of linearly independent vectors.

7) If $A = \begin{bmatrix} 1 & 1 \\ 4 & 1 \end{bmatrix}$, find the eigen values and eigen vectors A. Prove that A is diagonalizable.

Obtain a basis for R^2 containing eigen vectors of A.

Ans: $\lambda = 3,-1$ and eigen vectors are $\begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ -2 \end{bmatrix}$ $\left\{\begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ -2 \end{bmatrix}\right\}$ is a basis of R^2.

8) Find the eigen values and eigen vectors of the following matrices.

i) $\begin{bmatrix} 6 & -2 & 2 \\ -2 & 3 & -1 \\ 2 & -1 & 3 \end{bmatrix}$ ii) $\begin{bmatrix} 8 & -6 & 2 \\ -6 & 7 & -4 \\ 2 & -4 & 3 \end{bmatrix}$

Ans: i) $2,2,8; a\begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix} + b\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ where a, b are any nonzero scalars

and $c\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$ where c is any non zero scalar.

ii) $0,3,15$ $\begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix}, \begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix}$ and theire nonzero scalar multiples.

9) Show that the matrices $\begin{bmatrix} -10 & 6 & 3 \\ -26 & 16 & 8 \\ 16 & -10 & -5 \end{bmatrix}$ and $\begin{bmatrix} 0 & -6 & -16 \\ 0 & 17 & 45 \\ 0 & -6 & -16 \end{bmatrix}$ are similar.

10) Prove that the matrix. $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \in M_{2\times 2}(R)$ is diagonalizable.

11) Find all the eigen values and a basis for each eigen space of the linear operator $T : R^3 \to R^3$ defined by $T(a,b,c) = (2a+b, b-c, 2b+4c)$

Ans: eigen values 2, 3

eigen space of 2 is spanned by $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$

eigen space of 3 is spanned by $\begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}$

12. Find the eigen values of $\begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}$

Ans : 2, 2, 2

13. Show that $A = \begin{bmatrix} o & h & g \\ h & o & f \\ g & f & o \end{bmatrix}; B = \begin{bmatrix} o & f & h \\ f & o & g \\ h & o & g \end{bmatrix}$

have the same charecteristic equation.

12.12 Diagonalizability:

We have see in the preceeding articles that every linear operator or every matrix is not diagonalizable. We need a simple test to determine whether an operator or a matrix can be diagonalized as well as a method for actually finaling a basis of eigen vectors.

12.12.1 Theorem: Let T be linear operator on a vector space V; Let $\lambda_1, \lambda_2 ... \lambda_k$ be distinct eigen values of T. If $v_1, v_2 ... v_k$ are eigen vectors of T such that λ_i corresponds to V_i $(1 \le i \le k)$ then $\{v_1, v_2 ... v_k\}$ are linearly independent.

Proof: We prove the theorem by mathematical induction on k. If $k = 1,$ then $v_1 \ne O$ since v_i is an eigen vector and so $\{v_1\}$ is linearly independent. We assume th theorem is true for $(k-1)$ distinct eigen values where $(k-1) \ge 1$

Let there be k eigen vectors $v_1, v_2, ..., v_k$ corresponding to the distinct eigen values $\lambda_1, \lambda_2, ..., \lambda_k$. We will show that $\{v_1, v_2, ..., v_k\}$ is linearly independent.

Suppose $a_1, a_2, ..., a_k$ are scalars.

Such that $a_1 v_1 + a_2 v_2 + ... + a_k v_k = O$ ...... (1)

Applying $(T - \lambda_k I)$ on both sides of (1) we get

$$a_1(\lambda_1 - \lambda_k)v_1 + a_2(\lambda_2 - \lambda_k)v_2 + ... + a_{k-1}(\lambda_{k-1} - \lambda_k)v_{k-1} = 0$$

By induction hypothesis $\{v_1, v_2, ... v_{k-1}\}$ is linearly independent. So we must have $a_1(\lambda_1 - \lambda_k) = a_2(\lambda_2 - \lambda_k) = ... = a_{k-1}(\lambda_{k-1} - \lambda_k) = 0$.

As $\lambda_1, \lambda_2, ... \lambda_k$ are distinct it follows

that $\lambda_i - \lambda_k \ne 0$ for $1 \le i \le k-1$

So $a_1 = a_2 = ... = a_{k-1} = 0$

Substituting these values in (1) we get $a_k v_k = 0$.

As $v_k \ne O$ so $a_k = 0$ consequently $a_1 = a_2 = ... = a_3 = ... = a_{k-1} = a_k = 0$

Thus a linear combination of vectors $v_1, v_2 ... v_k$ is equal to a zero vector implies each of the scalar coefficient is zero.

So $\{v_1, v_2, ... v_k\}$ is linearly independent.

Hence the theorem.

12.12.2 Corollary: Let T be a linear operator on an n dimensional vector space V. If T has n distinct eigen values then T is diagonalizable.

Proof: Let the n distinct eigen values of T be $\lambda_1, \lambda_2, ...\lambda_k$. Foe each i, choose an eigen vector v_i corresponding to λ_i . By the above theorem $\{v_1, v_2, ...v_n\}$ is linearly independent and as dim $V = n$; the set $\{v_1, v_2, ...v_n\}$ is a basis of V. Hence T is diagonalizable.

Converse: The converse of the above theorem need not be true. i.e. If T is diagonalizable, then it has n distinct eigen values need not be true. For example the identity operator is diagonalizable even though it has only one eigen value namely 1.

W.E. 6: Worked Out Examples:

Show that $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \in M_{2\times 2}(R)$ is diagonalizable.

Solution: The characteristic polynomial of A is $|A - \lambda I| = \begin{vmatrix} 1-\lambda & 1 \\ 1 & 1-\lambda \end{vmatrix}$

$$= (1-\lambda)^2 - 1^2$$

$$= (1-\lambda+1)(1-\lambda-1)$$

$$= (2-\lambda)(-\lambda)$$

The characteristic equation is $\lambda(\lambda - 2) = 0$,

Thus $\lambda = 0, 2$

Hence the characteristic values of A are 0,2. Hence the characteristic values of L_A are 0,2. Which are distinct.

L_A is a linear operator on vector space R^2 , whose dimension is 2.

$\therefore L_A$ and hence A is diagonaliziable.

W.E. 7: Show that $A = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$ is not diagonalizable.

Solution:

The characteristic polynomial of A is $|A - \lambda I|$

$$= \begin{vmatrix} 1-\lambda & 2 \\ 0 & 1-\lambda \end{vmatrix} = (1-\lambda)^2$$

The characteristic equation of A is $|A-\lambda I|=0$ i.e. $(1-\lambda)^2=0$

$$\rightarrow \lambda = 1,1$$

As A and hence L_A has one distinct eigen value and dim R^2 is 2, it follows that A is not diagonalizable.

12.14 Polynomial Splitting:

12.14.1 Definition: A polynomial $f(t)$ in $P(F)$ splits over F if there are scalars $a_1, a_2, ...a_n$ not necessarily distinct in F such that $f(t)=c(t-a_1)(t-a_2)...(t-a_n)$.

Example:

i) $t^2-1=(t+1)(t-1)$ splits over R.

but ii) $(t^2+1)(t-2)$ does not split over R.

bu it splits over C because it factors into the product $(t+i)(t-i)(t-2)$.

Note: If $f(t)$ is the characteristic polynomial of a linear operator or a matrix over a field F, then the statement that $f(t)$ splits is to understood to mean that it splits over F.

12.14.2 Theorem: The characteristic polynomial of any diagonalizable linear operator splits.

Proof: Let V be a n dimensional vector space. Let T be a diagonalizable linear operator over V. Let B be an ordered basis for V such that $[T]_B = D$ is a diagonal matrix. Suppose that

$$D = \begin{bmatrix} \lambda_1 & 0 & 0........ & 0 \\ 0 & \lambda_2 & 0........ & 0 \\ \vdots & \vdots & \vdots & \\ 0 & 0 & 0 & \lambda_n \end{bmatrix}$$

Let $f(t)$ be the characteristic polynomial of T.

$$f(t)=\det(D-tI)$$

$$= \det \begin{bmatrix} \lambda_1 - t & 0 & .... & 0 \\ 0 & \lambda_2 - t & .... & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & & \lambda_n - t \end{bmatrix}$$

$= (\lambda_1 - t)(\lambda_2 - t)....(\lambda_n - t)$

$= (-1)^n (t - \lambda_1)(t - \lambda_2)....(t - \lambda_n)$

$f(t)$ is factored into a product of linear factors.

$\therefore f(t)$ splits. Hence the theorem.

Note: From the above theorem it is clear that if T is diagonalizable linear operator on an n - dimensional vector space that fails to have distinct eigen values, then the characteristic polynomial of T must have repeated zeros.

ii) The converse of the above theorem need not be true. That is, the characteristics polynomial may split but T need not be diagonalizable. For example consider the examples W.E. 5. in which we observe even though T may split, T need not be diagonalizable.

12.14.3 Algebraic Multiplicity:

Definition: Let λ be an eigen value of a linear operator or matrix with characteristic polynomial $f(t)$. The algebraic multiplicity of λ is the largest positive integer k for which $(t-\lambda)^k$ is a factor of $f(t)$.

W.E. 8 : Example:

i) $\det A = \begin{bmatrix} 3 & 1 & 0 \\ 0 & 3 & 4 \\ 0 & 0 & 4 \end{bmatrix}$

The characteristic polynomial is $|A - tI|$

$\Rightarrow \begin{vmatrix} 3-t & 1 & 0 \\ 0 & 3-t & 4 \\ 0 & 0 & 4-t \end{vmatrix}$ expanding along the first column; the characteristic polynomials is $(3-t)\left[(3-t)(4-t) - 0\right]$

$= (3-t)^2\ (4-t)$

Hence $3, 3, 4$ are the eigen values.

So $\lambda = 3$ is an eigen value of A with multiplicity 2. $\lambda = 4$ is an eigen value of A with multiplicity 1.

ii) For n x n null matrix, zero is the characteristic root of algebraic multiplicity n.

iii) For identity matrix of order n unity is the characteristic root of algebraic multiplicity n.

12.15 Eigen Space:

12.15.1 Definition: Let T be a linear operator on a vector space V, det λ be an eigen value of T. Define $E_\lambda = \{v \in V : T(v) = \lambda v\} = N(T - \lambda I_v)$.

The set E_λ is called the eigen space of T corresponding to the eigen value λ.

Analogously we define the eigen space of a square matrix A to be the eigen space of L_A.

Note: E_λ is a subspace of V consisting of the zero vector and eigen vectors of T. Corresponding to the eigen value λ. So the maximum number of linearly independent eigen vectors of T corresponding to the eigen value λ is the dimension of E_λ.

12.15.2 Theorem: Let T be a linear operator on a finite dimensional vector space V. Let λ be an eigen value of T having multiplicity m. Then $1 \le \dim(E_\lambda) \le m$.

Proof: Choose an ordered basis $\{v_1, v_2, ..., v_p\}$ for E_λ. extend it to an ordered basis $B = \{v_1, v_2, ..., v_p, v_{p+1} ... v_n\}$ for V.

Let $A = [T]_B$. Observe that v_i $(1 \le i \le p)$ is an eigen vector of T corresponding to λ and therefore $A = \begin{bmatrix} \lambda I_p & B \\ 0 & C \end{bmatrix}$

Then the characteristic polynomial of T is $f(t) = \det(A - tI_n) = \det \begin{bmatrix} (\lambda - t)I_p & B \\ 0 & c - tI_{n-p} \end{bmatrix}$

$$= \det\left((\lambda - t)I_p\right)\det\left(c - tI_{n-p}\right)$$

$= (\lambda - t)^p g(t)$ where $g(t)$ is a polynomial.

Thus $(\lambda - t)^p$ is a factor of $f(t)$ and hence the multiplicity of λ is atleast p. But dim $(E_\lambda) = p$. So dim $(E_\lambda) \le m$.

Hence $1 \le \dim(E_\lambda) \le m$.

12.15.3 We state some theorem without proofs:

Lemma: Let T be a linar operator and Let $\lambda_1, \lambda_2 ... \lambda_k$ be distinct eigen values of T. For each $i = 1, 2...k$, let $v_i \in E\lambda_i$, the eigen space corresponding to λ_i. If $v_1 + v_2 + ... + v_k = 0$ then $v_i = 0$ for all i.

12.15.4 Theorem: Let T be a linear operator on a vector space V, and let $\lambda_1, \lambda_2, ... \lambda_k$ be distinct eigen values of T. For each $i = 1, 2...k$; let S_i be a finite linearly independent subset of Eigen space $E\lambda_i$. Then $S = S_1 \cup S_2 \cup S_3 \cup ... \cup S_k$ is a linearly independent subset of V.

12.15.5 Theorem: Let T be a linear operator on a finite dimensional vector space V such that the characteristic polynomial of T splits. Let $\lambda_1, \lambda_2, ... \lambda_k$ be the distinct eigen values of T. Then i) T is diagonalizable if and only if the multiplicity of λ_i is equal to the dim $(E\lambda_i)$ for all i.

ii) If T is diagonalizable and B_i is an ordered basis for $E\lambda_i$ for each i, then $B = B_1 \cup B_2 \cup B_3 \cup ... \cup B_k$ is an ordered basis for V consisting of eigen vectors of T.

12.16 Summary:

Test for Diagonalization :

Let T be a linear operator on an n dimensional vector space V. Then T is diagonalizable if and only if both the following conditions hold.

i) The characteristic polynomial of T splits.

ii) For each eigen value λ of T, the multiplicity of λ equals to n - rank $(T - \lambda I)$

In order to test the diagonalizability of a square matrix, the same conditions can be used since the diagonalizability of A is equivalent to the diagonalizability of the operator L_A.

If T is a diagonalizable operator and $B_1, B_2, ... B_k$ are ordered basis for the eigen space of T, then the union $B = B_1 \cup B_2 \cup ... \cup B_k$ is an ordered basis for V consisting of eigen vectors of T, and hence $[T]_B$ is a diagonal matrix.

When we want to test T for diagonalizability we usually choose a convient basis B for V, and form $A = [T]_B$ if the characteristic polynomial of A splits, then use the condition (ii) above to check if the multiplicity of each of the repeated eigen value of A equal to n - rank $(A - \lambda I)$. If the characteristic polyomial of A is splitting condition. (ii) is automatically satisfied for eigen values with multiplicity 1. If A is diagonlizable then T is also diagonalizable.

If we find T is diagonalizable and want to find a basis B for V, consisting of eigen vectors of T, we addopt the following procedure.

1) We first find a basis for each eigen space of A. the union of these bases is a basis C for F^n consisting of eigen vectors of A. Each vector in C is the coordinate vector relative to B of an eigen vector of T. The set consisting of these n eigen vectors of T is the desired basis B.

Further more, if A is a $n\times n$ diagonalizable matrix, we can find an invertible $n\times n$ matrix Q and a diagonal $n\times n$ matrix D such that $Q^{-1}AQ=D$.

The matrix Q has as its columns the vectors in a basis of eigen vectors of A, and D has as its j th diagonal entry the eigen value of A corresponding to the j th column of Q.

12.17 Workedout Examples:

W.E. 9:

Let T be the linear on $P_2(R)$ defined by $T\left[f(x)\right]=f^{'}(x)$. Test T for diagonalizability.

Solution: T is a linear operator on $P_2\left[R\right]$ defined by $T\left[f(x)\right]=f^{'}(x)$. The standard basis for $P_2(R)$ is $B=\left\{1,x,x^2\right\}$

Given $T\left[f(x)\right]=f^{'}(x)$

$$T(1)=0=0.1+0.x+0.x^2$$

$$T(x)=1=1(1)+0x+0x^2$$

$$T(x^2)=2x=0(1)+2x+0x^2$$

$$A=\left[T\right]_B=\begin{bmatrix}0&1&0\\0&0&2\\0&0&0\end{bmatrix}$$

The characteristic equation is $\left|A-\lambda I\right|=0$

$$\Rightarrow\begin{vmatrix}-\lambda&1&0\\0&-\lambda&2\\0&0&-\lambda\end{vmatrix}=0$$

expanding along the third row we get

$$0-0+(-\lambda)(-\lambda^2-0)=0$$

$$-\lambda^3 - 0$$

Thus T has only one eigen value $\lambda = 0$ with multiplicity 3.

$T\left[f(x)\right] = f^{'}(x) = 0 \Rightarrow f(x)$ is a constant polynomial.

$E_\lambda = N(T - \lambda I) = N(T)$ is the subspace of $P_2(R)$ with constant polynomials.

So $\{1\}$ is a basis of E_λ. So dim $(E_\lambda) = 1$ consequently there is no basis of $P_2(R)$ consisting of eigen vectors of T. And so T is not diagonalizable.

W.E. 10: Test the matrix $A = \begin{bmatrix} 3 & 1 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{bmatrix} \in M_{3\times 3}(R)$ for diagonalizability.

Solution: The characteristic equation is $|A - \lambda I| = 0$.

$$\Rightarrow \begin{bmatrix} 3-\lambda & 1 & 0 \\ 0 & 3-\lambda & 0 \\ 0 & 0 & 4-\lambda \end{bmatrix} = 0$$

expanding along the third row.

$$0 - 0 + (4-\lambda)\left\{(3-\lambda)^2 - 0\right\} = 0$$

$$\Rightarrow (4-\lambda)(3-\lambda)^2 = 0 \text{ so } \lambda = 4, 3, 3$$

Also A has eigen values $\lambda_1 = 4$ and $\lambda_2 = 3$ with multiplicities 1 and 2 respectively.

Since λ_1 has multiplicity 1; condition (ii) is satisfied for λ_1. Thus we need only to test condition (ii) for λ_2.

To find the rank of $(A - \lambda_2 I)$ where $\lambda_2 = 3$

$$(A - \lambda_2 I) = \begin{bmatrix} 3-3 & 1 & 0 \\ 0 & 3-3 & 0 \\ 0 & 0 & 4-3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Which can be put in Echelon form. Here the number of non zero rows 2. So Rank of $(A-\lambda_2 I)=2$.

For $\lambda_2=3$, n - rank $(A-\lambda_2 I)=3-2=1$. Which is not the multiplicity of λ_2. So the condition (ii) fails for λ_2 and so A is not diagonalizable.

W.E. 11:

Test for diagonalizability of the linear operator T on $P_2(R)$ defined as follows:

$$T\left[f(x)\right]=f(1)+f^{'}(0)x+\left(f^{'}(0)+f^{''}(0)\right)x^2$$

Also find an ordered basis for R^3 of eigen vectors of $[T]_B$ where B is the standard basis of $P_2(R)$.

Solution: T is a linear operator on $P_2(R)$ defined by $T\left[f(x)\right]=f(1)+f^{'}(0)x+\left(f^{'}(0)+f^{''}(0)\right)x^2$

$B=\{1,x,x^2\}$ is the standard ordered basis for $P_2(R)$ and $A=[T]_B$

$T(1)=1+0x+0x^2$; $\quad T(x)=1+1.x+(1+0)x^2$

i.e. $T(x)=1+x+x^2$

$T(x^2)=1+0x+(0+2)x^2$

i.e. $T(x^2)=1+0x+2x^2$

Thus $A=[T]_B=\begin{bmatrix}1 & 1 & 1\\ 0 & 1 & 0\\ 0 & 1 & 2\end{bmatrix}$

det $(A-\lambda I)=\begin{vmatrix}1-\lambda & 1 & 1\\ 0 & 1-\lambda & 0\\ 0 & 1 & 2-\lambda\end{vmatrix}$

$=(1-\lambda)\left[(1-\lambda)(2-\lambda)-0\right]$ expanding along the first column.

$=(1-\lambda)^2(2-\lambda)$

The characteristic polynomial of A and hence of T is $(1-\lambda)^2(2-\lambda)$ which splits. Hence the condition (i) is satisfied.

Also $\lambda_1 = 1$ has multiplicity 2.

and $\lambda_2 = 2$ has multiplicity 1 and hence condition (ii) is satisfied.

So we verify condition (ii) for $\lambda_1 = 1$

For this (n - rank $(A-\lambda_1 I)$)

$$= 3 \text{ - rank of } \begin{bmatrix} 1-1 & 1 & 1 \\ 0 & 1-1 & 0 \\ 0 & 1 & 2-1 \end{bmatrix}$$

$$= 3 \text{ - rank } \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

$= 3-(1) = 2$. Since the matrix has only one linearly independent row.

Here n - rank $(A - \lambda_1 I) = 2$ = multiplicity of λ_1.

Hence as the required conditions are satisfied, T is diagonalizable.

We now find the ordered basis C for R^3 of eigen vectors of A. We consider each eigen value separately.

Let $\lambda_1 = 1$; then $(A-\lambda I)v = 0$

$$\Rightarrow \begin{bmatrix} 1-1 & 1 & 1 \\ 0 & 1-1 & 0 \\ 0 & 1 & 1 \end{bmatrix} . \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$\Rightarrow v_2 + v_3 = 0 \rightarrow v_2 = -v_3$ Let $v_1 = s, v_3 = t$

then $\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} s \\ -t \\ t \end{bmatrix} = s\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + t\begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}$

So $C_1 = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} \right\}$ is a basis for the eigen space $E_{\lambda_1} = \left\{ \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \in R^3 \middle/ (A - \lambda I)v = 0 \right\}$

To find the eigen space corresponding to $\lambda_2 = 2$.

$$(A - \lambda_2 I)v = O$$

$$\Rightarrow \begin{bmatrix} 1-2 & 1 & 1 \\ 0 & 1-2 & 0 \\ 0 & 0 & 2-2 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} -1 & 1 & 1 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$\Rightarrow -v_2 = 0, -v_1 + v_2 + v_3 = 0 \rightarrow -v_1 + v_3 = 0$ so $v_3 = v_1$

Put $v_1 = 1$ then $v_2 = 0$, $v_3 = 1$

So $C_2 = \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right\}$ is the basis for the eigen space $E_{\lambda_2} = \left\{ v = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \in R^3 \middle/ (A - \lambda_2 I)v = 0 \right\}$

consider $C = C_1 \cup C_2$ then

$$C = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right\}$$

Thus C is an ordered basis for R^3 consisting of eigen vectors of A.

Finally we observe that the vectors in C are the coordinate vectors relative to B of the vectors in the set.

$A=\{1,-x+x^2,1+x^2\}$ which is an ordered basis for $P_2(R)$ consisting of eigen vectors of T. Thus

$$[T]_A=\begin{bmatrix}1&0&0\\0&1&0\\0&0&2\end{bmatrix}$$ which is the required diagonal matrix.

W.E. 12: Show that the matrix $\begin{bmatrix}0&-2\\1&3\end{bmatrix}$ is diagonalizable and find a 2×2 matrix P such that $P^{-1}AP$ is a diagonalizable matrix.

Solution: The characteristic equation of the given matrix $A=\begin{bmatrix}0&-2\\1&3\end{bmatrix}$ is

$$|A-\lambda I|=0\Rightarrow\begin{vmatrix}-\lambda&-2\\1&3-\lambda\end{vmatrix}=0.$$

$$\Rightarrow -\lambda(3-\lambda)+2=0\rightarrow \lambda^2-3\lambda+2=0$$

$\Rightarrow(\lambda-2)(\lambda-1)=0$ in $\lambda=2,1$

Thus A has two distinct eigen values $\lambda_1=1,\lambda_2=2$. As the diamensionality of the vector space is 2.

We see that A is diagonalizable.

1) To find the eigen space corresponding to $\lambda_1=1$:

We have $(A-\lambda I)v=O$

$$\Rightarrow\begin{vmatrix}-1&-2\\1&3-1\end{vmatrix}\begin{bmatrix}v_1\\v_2\end{bmatrix}=\begin{bmatrix}0\\0\end{bmatrix}\Rightarrow\begin{bmatrix}-1&-2\\1&2\end{bmatrix}\begin{bmatrix}v_1\\v_2\end{bmatrix}=\begin{bmatrix}0\\0\end{bmatrix}$$

$\Rightarrow v_1+2v_2=0\rightarrow v_1=-2v_2$ put $v_2=1$ then $v_1=-2$

So $v=\begin{bmatrix}v_1\\v_2\end{bmatrix}=\begin{bmatrix}-2\\1\end{bmatrix}$

or $C_1 = \left\{ \begin{bmatrix} -2 \\ 1 \end{bmatrix} \right\}$ is a basis of eigen space. $E_{\lambda_1} = \left\{ \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \in R^2 \Big/ (A - \lambda_1 I) = O \right\}$

ii) To find eigen space corresponding to $\lambda_2 = 2$.

$$(A - \lambda_1 I)v = O \Rightarrow \begin{bmatrix} 0-2 & -2 \\ 1 & 3-2 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$\Rightarrow \begin{bmatrix} -2 & -2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = 0 \rightarrow v_1 + v_2 = 0$ so $v_1 = -v_2$

Put $v_2 = 1,$ then $v_1 = -1$ so $v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$

So $C_2 = \left\{ \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right\}$ is the basis of the eigen space $E_{\lambda_2} = \left\{ \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \in R^2 \Big/ (A - \lambda_2 I)v = O \right\}$

$C = C_1 \cup C_2 = \left\{ \begin{bmatrix} -2 \\ 1 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right\}$ is an ordered basis for R^2 consisting of the eigen vectors of A.

Let $P = \begin{bmatrix} -2 & -1 \\ 1 & 1 \end{bmatrix}$ is the matrix whose columns are vectors in C.

$$D = P^{-1}AP = [L_A]_B = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$$

W.E. 13:

Let T be the linear operator on R^3 which is represented in the standard basis by the matrix

$\begin{bmatrix} -9 & 4 & 4 \\ -8 & 3 & 4 \\ -16 & 8 & 7 \end{bmatrix}$. Prove that T is diagonalizable. Find a basis of R^3 consisting of eigen vectors of T.

Solution: The given matrix $A = \begin{bmatrix} -9 & 4 & 4 \\ -8 & 3 & 4 \\ -16 & 8 & 7 \end{bmatrix}$

The characteristic polynomial of A corresponding to the characteristic root λ is

λ^3 (trace of A) $\lambda^2 + (A_{11} + A_{22} + A_{33})\lambda - \det A$

Where A_{ii} is the cofactor of the diagonal element a_{ii}.

trace of $A = a_{11} + a_{22} + a_{33} = -9 + 3 + 7 = 1$

$$A_{11} = (-1)^{1+1} \begin{vmatrix} 3 & 4 \\ 8 & 7 \end{vmatrix} = 21 - 32 = -11$$

$$A_{22} = (-1)^{2+2} \begin{vmatrix} -9 & 4 \\ -16 & 7 \end{vmatrix} = -63 + 64 = 1$$

$$A_{33} = (-1)^{3+3} \begin{vmatrix} -9 & 4 \\ -8 & 3 \end{vmatrix} = -27 + 32 = 5$$

$$A_{11} + A_{22} + A_{33} = -11 + 1 + 5 = -5$$

$$\det A = -9(21-32) - 4(-56+64) + 4(-64+48)$$

$$= 99 - 32 - 64 = 3$$

The characteristic polynomial is

$$\lambda^3 - \lambda^2 - 5\lambda - 3 = f(\lambda) \text{ say}$$

-1	1	-1	-5	-3
		-1	2	3
	1	-2	-3	0

$f(-1) = -1 - 1 + 5 - 3 = 0$

$\lambda + 1$ is a factor of $f(\lambda) = 0$

The other factor is

$$\lambda^2 - 2\lambda - 3$$

$$= (\lambda - 3)(\lambda + 1)$$

So $f(\lambda)=0 \Rightarrow (\lambda+1)^2(\lambda-3)=0$

Hence the characteristic roots are $-1,-1,3$

Thus the eigen value $\lambda_1=-1$ has the multiplicity 2

$\lambda_2=3$ has the multiplicity 1.

1) To find eigen space corresponding to $\lambda_1=-1$

$(A-\lambda I)v=O$

$$\Rightarrow \begin{bmatrix} -9+1 & 4 & 4 \\ -8 & 3+1 & 4 \\ -16 & 8 & 7+1 \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}=O$$

$$\Rightarrow \begin{bmatrix} -8 & 4 & 4 \\ -8 & 4 & 4 \\ -16 & 8 & 8 \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}=O$$

R_2-R_1, R_3-2R_1 gives

$$\begin{bmatrix} -8 & 4 & 4 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}=O$$

$R_1 \times \frac{1}{4}$ gives $\begin{bmatrix} -2 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}=O=\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

$-2v_1+v_2+v_3=0$ put $v_1=t, v_2=s$

then $v_3=2v_1-v_2=2t-s$

$$v=\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}=\begin{bmatrix} t \\ s \\ 2t-s \end{bmatrix}=t\begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}+s\begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$$

So $C_1 = \left\{ \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} \right\}$ is a basis of for the eigen space $E_{\lambda_1} = \left\{ \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \in R^3 \middle/ (A-\lambda_1 I)v = O \right\}$

Dim of $E_{\lambda_1} = 2$ which is equal to the multiplication of $\lambda_1 = -1$.

i.e. multiplicity of $\lambda_1 =$ Dim E_{λ_1}

ii) To find the eigen space corresponding to $\lambda_2 = 3$:

$(A - \lambda_2 I)v = O$

$$\Rightarrow \begin{bmatrix} -9-3 & 4 & 4 \\ -8 & 3-3 & 4 \\ -16 & 8 & 7-3 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = O$$

$$\Rightarrow \begin{bmatrix} -12 & 4 & 4 \\ -8 & 0 & 4 \\ -16 & 8 & 4 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = O$$

$R_1 \times \frac{1}{4}, R_2 \times \frac{1}{4}, R_3 \times \frac{1}{4}$ gives

$$\begin{bmatrix} -3 & 1 & 1 \\ -2 & 0 & 1 \\ -4 & 2 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = O$$

$3R_2 - 2R_1, 3R_3 - 3R_1$ gives

$$\begin{bmatrix} -3 & 1 & 1 \\ 0 & -2 & 1 \\ 0 & 2 & -1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = O$$

$R_3 + R_2$ gives

$$\begin{bmatrix} -3 & 1 & 1 \\ 0 & -2 & 1 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Rightarrow -2v_2 + v_3 = 0$$ so $v_3 = 2v_2$ and

$$-3v_1 + v_2 + v_3 = 0 \to -3v_1 + v_2 + 2v_2 = 0$$

$\Rightarrow 3v_1 = 3v_2$ so $v_2 = v_1$

Put $v_1 = 1,$ so $v_2 = 1,$ $v_3 = 2$ Hence $v = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$

So $C_2 = \left\{ \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} \right\}$ is a basis of eigen space $E_{\lambda_2} = \left\{ \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \in R^3 \Big/ (A - \lambda_2 v) = O \right\}$

Dim $E_{\lambda_2} = 1$ which is equal to the multiplicity of $\lambda_2 = 3$

Multiplicity of $\lambda_2 = \dim E_{\lambda_2}$

If we consider the union of bases of these two subspaces we get $C = \left\{ \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} \right\}$ which is linearly independent. Thus the set C is a basis of R^3 consisting of the eigen vectors of T.

Hence T is diagonalizable.

W.E. 14:

Let $T : R^3 \to R^3$ is defined as follows

$$T\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} 4a_1 + a_3 \\ 2a_1 + 3a_2 + 2a_3 \\ a_1 + 4a_3 \end{bmatrix}$$

Find whether the linear operator T is diagonalizable or not.

Solution: Given $T\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} 4a_1 + a_3 \\ 2a_1 + 3a_2 + 2a_3 \\ a_1 + 4a_3 \end{bmatrix}$

$B = \{e_1, e_2, e_3\}$ where $e_1 = \{1,0,0\}, e_2 = \{0,1,0\}, e_3 = \{0,0,1\}$ is the standard basis of for R^3.

$$T(e_1) = T\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 4+0 \\ 2+0+0 \\ 1+0+0 \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \\ 1 \end{bmatrix}$$

$$= 4(1,0,0) + 2(0,1,0) + 1(0,0,1)$$

$$= 4e_1 + 2e_2 + 1e_3$$

$$T(e_2) = T\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 4(0)+0 \\ 2(0)+3(1)+2(0) \\ 0+4(0) \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \\ 0 \end{bmatrix}$$

$$= 0(1,0,0) + 3(0,1,0) + 0(0,0,1)$$

$$= 0e_1 + 3e_2 + 0e_3$$

$$T(e_3) = T\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 4(0)+1 \\ 2(0)+3(1)+2(1) \\ 0+4(1) \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}$$

$$= 1(1,0,0) + 2(0,1,0) + 4(0,0,1)$$

$$= 1e_1 + 3e_2 + 0e_3$$

Writing $[T]_B$ as A we get

$$A = [T]_B = \begin{bmatrix} 4 & 0 & 1 \\ 2 & 3 & 2 \\ 1 & 0 & 4 \end{bmatrix}$$

Hence the characteristic polynomial is $|A - \lambda I|$.

The characteristic polynomial of A corresponding to the characteristic root λ is

$$\lambda^3 - \text{(trace of A)}\,\lambda^2 + (A_{11} + A_{22} + A_{33})\lambda - \det A .$$

Where A_{ii} is the cofactor or diagonal element a_{ii}. trace of $A = 4+3+4 = 11$

$$A_{11} = (-1)^{1+1}\begin{vmatrix} 3 & 2 \\ 0 & 4 \end{vmatrix} = 12 - 0 = 12$$

$$A_{22} = (-1)^{2+2}\begin{vmatrix} 4 & 1 \\ 1 & 4 \end{vmatrix} = 16 - 1 = 15$$

$$A_{33} = (-1)^{3+3}\begin{vmatrix} 4 & 0 \\ 2 & 3 \end{vmatrix} = 12 - 0 = 12$$

$$A_{11} + A_{22} + A_{33} = 12 + 15 + 12 = 39$$

det $A = 4(12-0) - 0 + 1(0-3) = 48 - 3 = 45$

Hence the characteristic polynomial is $\lambda^3 - 11\lambda^2 + 39\lambda - 45 = f(\lambda)$ say

$f(3) = 27 - 99\lambda - 45 = 0$

$\therefore \lambda - 3$ is a factor of the characteristic equation $f(\lambda) = 0$.

The other factor is

$$\lambda^2 - 8\lambda + 15$$

$$= (\lambda - 3)(\lambda - 5)$$

3	1	–11	39	–45	
		3	–24	45	
	1	–8	15	0	

So $f(\lambda) = (\lambda - 3)^2(\lambda - 5)$

Hence the characteristic roots are 3, 3, 5.

So the eigen values of T are

$\lambda_1 = 5$ with multiplicity 1

$\lambda_2 = 3$ with multiplicity 2

i) To find the eigen space corresponding to $\lambda_1 = 5$:

$$(A - \lambda_1 I)v = O$$

$$\begin{bmatrix} 4-5 & 0 & 1 \\ 2 & 3-5 & 2 \\ 1 & 0 & 4-5 \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} -1 & 0 & 1 \\ 2 & -2 & 2 \\ 1 & 0 & -1 \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$R_2 + 2R_1, R_3 + R_1$ gives

$$\Rightarrow \begin{bmatrix} -1 & 0 & 1 \\ 0 & -2 & 4 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow -2v_2 + 4v_3 = 0 \Rightarrow v_2 = 2v_3$$

$$-v_1 + v_3 = 0 \qquad v_1 = v_3$$

Put $v_3 = 1,$ then $v_2 = 2, v_1 = 1$

So $v = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$

So $C_1 = \left\{ \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \right\}$ is a basis of eigen space $E_{\lambda_1} = \left\{ \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \in R^3 \Big/ (A - \lambda_1 v) = O \right\}$

Dim of $E_{\lambda_1} = 1$ which is equal to the multiplicity of λ_1.

ii) To find the eigen space corresponding to $\lambda_2 = 3$

$$(A - \lambda_2 I)v = O$$

$$\begin{bmatrix} 4-3 & 0 & 1 \\ 2 & 3-3 & 2 \\ 1 & 0 & 4-3 \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = O$$

$$\Rightarrow \begin{bmatrix} 1 & 0 & 1 \\ 2 & 0 & 2 \\ 1 & 0 & 1 \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = O = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$R_2 - 2R_1, R_3 - R_1$ gives

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow 1v_1 + v_3 = 0 \Rightarrow v_3 = -v_1$$

The unknown v_2 does not appear in this system, we assign it a parametric value say $v_2 = s$ and solve the system for v_3 and v_1. If $v_3 = t$, then $v_1 = -t$, introducing another parameter t. The result is the general solution to the system.

$$\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} -t \\ s \\ t \end{bmatrix} = s\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + t\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \text{ for } s,t \in R$$

So $C_2 = \left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\}$ is a basis of the eigen space $E_{\lambda_2} = \left\{ \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \in R^3 \middle/ (A - \lambda_2 v) = O \right\}$

$\dim E_{\lambda_2} = 2$; The multicity of $\lambda_2 = 2$

So $\dim E_{\lambda_2} =$ multicity of $\lambda_2 = 2$

The union of two bases C_1 and C_2.

is $C = C_1 \cup C_2 = \left\{ \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\}$ is linearly independent and hence is a basis of R^3; consisting of eigen vectors of T. Consequenty T is diagonalizable.

W.E.15: Let T be a linear operator an $P_2(R)$ defined by $T[f(x)] = f(x) + (x+1)f'(x)$. Show that T is diagonalizable.

Solution: $B = \{1, x, x^2\}$ is the standard basis of $P_2(R)$ we are given a linear operator on $P_2(R)$ defined by $T[f(x)] = f(x) + (x+1)f'(x)$.

In W.E. 4. We have shown $A = [T]_B = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 2 \\ 0 & 0 & 3 \end{bmatrix}$ and the eigen values of T are 1, 2, 3.

We will now find eigen vectors corresponding to the eigen values 1, 2, 3.

To find the eigen space corresponding to $\lambda_1 = 1$.

We have $(A - \lambda_1 I)v = O$

$$\Rightarrow \begin{bmatrix} 1-1 & 1 & 0 \\ 0 & 2-1 & 2 \\ 0 & 0 & 3-1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = 0 \Rightarrow \begin{bmatrix} 0 & 1 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$\Rightarrow 2v_3 = 0$ so $v_3 = 0$ and $v_2 + 2v_3 = 0 \Rightarrow v_2 = 0$.

Since $v_3 = 0$ and $v_2 = 0$ and v_1 can take any real value. Say $v_1 = 1$

Hence the eigen vector corresponding to $\lambda_1 = 1$ is given by $v = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ and every non zero scalar multiple of it is an eigen vector.

So $C_1 = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right\}$ is a basis for the eigen space.

$$E_{\lambda_2} = \left\{ \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \in R^3 \middle/ (A - \lambda_2 I)v = O \right\}$$

$\dim E_{\lambda_1} = 1$;   Multiplicity of $\lambda_1 = 1$.

$\therefore \dim E_{\lambda_1} =$ Multiplicity of λ_1

To find the eigen space corresponding to $\lambda_2 = 2$:

$$(A - \lambda_2 I)v = O$$

$$\begin{bmatrix} 1-2 & 1 & 0 \\ 0 & 2-2 & 2 \\ 0 & 0 & 3-2 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = O$$

$$\Rightarrow \begin{bmatrix} -1 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = O \Rightarrow v_3 = 0$$

$$-v_1 + v_2 = 0$$

$$v_2 = v_1$$

put $v_1 = 1$, then $v_2 = 1$, $v_3 = 0$

Hence $v = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ is the eigen vector and every scalar multiple of it is also an eigen vector.

So $C_2 = \left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \right\}$ is a basis for the eigen space.

$$E_{\lambda_2} = \left\{ \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \in R^3 \Big/ (A - \lambda_2 I) = O \right\}$$

Dim of $E_{\lambda_2} = 1$; Multiplicity of $\lambda_2 = 1$

$\therefore$ Dim of $E_{\lambda_2} =$ Multiplicity of λ_2

To find the eigen space corresponding to $\lambda_3 = 3$:

$$(A - \lambda_3 I)v = O$$

$$\begin{bmatrix} 1-3 & 1 & 0 \\ 0 & 2-3 & 2 \\ 0 & 0 & 3-3 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = O$$

$$\Rightarrow \begin{bmatrix} -2 & 1 & 0 \\ 0 & -1 & 2 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Rightarrow -v_2 + 2v_3 = 0 \text{ so } v_2 = 2v_3 \quad -2v_1 + v_2 = 0 \quad v_2 = 2v_1$$

put $v_1 = 1,$ so $v_2 = 2, v_3 = 1$

Hence $v = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$ is an eigen vector and every scalar multiple of it is also an eigen vector.

So $C_3 = \left\{ \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \right\}$ is a basis for the eigen space.

$$E_{\lambda_3} = \left\{ \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \in R^3 \Big/ (A - \lambda_3 I)v = O \right\}; \dim E_{\lambda_3} = 1$$

Multiplicity of $\lambda_3 = 1$ so $\dim E_{\lambda_3} =$ Multiplicity of λ_3

Thus we observe that the multiplicity of each eigen value is equal to the dimension of the corresponding eigen space.

Let $C = C_1 \cup C_2 \cup C_3 = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \right\}$ is linearly independent. Thus this set is a basis of R^3 consisting of eigen vectors of T.

So T is diagonalizable.

12.18 Method to find any positive integral power of a diagonalizable matrix A:

A is a diagonalisable matrix of order n. We can find a n x n matrix Q such that Q^{-1}AQ is a diagonal matrix D.

$$D = Q^{-1}AQ = [L_A]_B$$

Q is the matrix whose columns are the eigen vectors.

Preoperating with Q, and post operating with Q^{-1} we get $A = QDQ^{-1}$

$$A^k = (QDQ^{-1})^k$$

So $A^k = (QDQ^{-1})(QDQ^{-1})...(QDQ^{-1})$ (k terms)

$$= QD(Q^{-1}Q)D(Q^{-1}Q)...DQ^{-1}$$

So $A^k = QD^kQ^{-1}$

W.E. 16: Examples:

For $A = \begin{bmatrix} 1 & 4 \\ 2 & 3 \end{bmatrix} \in M_{2\times 2}(R)$ find an expression for A^n where n is a positive integer.

Solution: given $A = \begin{bmatrix} 1 & 4 \\ 2 & 3 \end{bmatrix}$. We show that A is diagonalizable and find a 2×2 matrix Q, such that $Q^{-1}A$Q is a diagonal matrix.

Then we compute A^n for any positive integer no.

The characteristic equation of A is

$$|A-\lambda I|=0 \Rightarrow \begin{vmatrix} 1-\lambda & 4 \\ 2 & 3-\lambda \end{vmatrix}=0$$

$$\Rightarrow (1-\lambda)(3-\lambda)-8=0$$

$$\lambda^2-4\lambda-5=0 \Rightarrow (\lambda-5)(\lambda+1)=0$$

Hence the characteristic values are -1, 5.

Hence the eigen values of L_A are -1, 5.

Which are district, L_A is a linear operator on a two dimensional vector space. So L_A and hence A is diagonalizable.

<u>To find eigen space corresponding to</u> $\lambda_1=-1$:

$$(A-\lambda_1 I)v=0 \Rightarrow \begin{bmatrix} 1-\lambda_1 & 4 \\ 2 & 3-\lambda_1 \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \end{bmatrix}=O$$

$$\Rightarrow \begin{bmatrix} 1+1 & 4 \\ 2 & 3+1 \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \end{bmatrix}=O \Rightarrow \begin{bmatrix} 2 & 4 \\ 2 & 4 \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \end{bmatrix}=O=\begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$\Rightarrow 2v_1+4v_2=0 \Rightarrow v_1=-2v_2$ put $v_2=1$ and so $v_1=-2$. There fore $\begin{bmatrix} v_1 \\ v_2 \end{bmatrix}=\begin{bmatrix} -2 \\ 1 \end{bmatrix}$

Hence $C_1=\left\{\begin{bmatrix} -2 \\ 1 \end{bmatrix}\right\}$ is a basis for the eigen

space $E_{\lambda_1}=\left\{\begin{bmatrix} v_1 \\ v_2 \end{bmatrix}\in R^2 \middle/ (A-\lambda_1 I)v=O\right\}$

<u>To find the eigen space corresponding to</u> $\lambda_2=5$:

$$(A-\lambda_2 I)v=O \quad \Rightarrow \begin{bmatrix} 1-5 & 4 \\ 2 & 3-5 \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \end{bmatrix}=O$$

$$\Rightarrow \begin{bmatrix} -4 & 4 \\ 2 & -2 \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \end{bmatrix}=O=\begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$\Rightarrow -4v_1 + 4v_2 = 0 \Rightarrow -v_1 + v_2 \Longrightarrow v_1 = v_2$ put $v_1 = 1,$ them $v_2 = 1$ so $v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

$C_2 = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$ is a basis for the eigen space $E_{\lambda_1} = \left\{ \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \in R^2 \Big/ (A - \lambda_1 I)v = O \right\}$

So $C = C_1 \cup C_2 = \left\{ \begin{bmatrix} -2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$ is an ordered basis for R^2 consisting of eigen vectors of A.

Let $Q = \begin{bmatrix} -2 & 1 \\ 1 & 1 \end{bmatrix}, Q^{-1} = \frac{-1}{3} \begin{bmatrix} 1 & -1 \\ -1 & -2 \end{bmatrix}$

$$D = Q^{-1}AQ = [L_A] = \begin{bmatrix} -1 & 0 \\ 0 & 5 \end{bmatrix}$$

$$A^n = QD^nD^{-1}$$

$$= Q \begin{bmatrix} (-1)^n & 0 \\ 0 & (5)^n \end{bmatrix} Q^{-1}$$

$$= \begin{bmatrix} -2 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} (-1)^n & 0 \\ 0 & (5)^n \end{bmatrix} \left(\frac{-1}{3} \right) \begin{bmatrix} 1 & -1 \\ -1 & -2 \end{bmatrix}$$

$$= \frac{1}{3} \begin{bmatrix} -2 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} (-1)^n & 0 \\ 0 & 5^n \end{bmatrix} \begin{bmatrix} -1 & 1 \\ 1 & 2 \end{bmatrix}$$

$$= \frac{1}{3} \begin{bmatrix} (-2)(-1)^n & 5^n \\ (1)^n & 5^n \end{bmatrix} \begin{bmatrix} -1 & 1 \\ 1 & 2 \end{bmatrix}$$

$$= \frac{1}{3} \begin{bmatrix} 2(-1)^n + 5^n & (-2)(-1)^n + 2.(5)^n \\ -(1)^n + 5^n & (-1)^n + 2.5^n \end{bmatrix}$$

W.E. 17: If $A = \begin{bmatrix} 1 & 1 \\ 4 & 1 \end{bmatrix}$. Find the eigen values and eigen vectors of A. Prove that A is diagonalizable. Find a basis of R^2 containing eigen vectors of A.

Solution: The characteristic equation is $|A-\lambda I| = O$

$$\Rightarrow \begin{vmatrix} 1-\lambda & 1 \\ 4 & 1-\lambda \end{vmatrix} = O \Rightarrow (1-\lambda)^2 - 4 = 0$$

$$\Rightarrow (1-\lambda+2)(1-\lambda-2) = O \Rightarrow (3-\lambda)(-\lambda-1) = 0$$

$$\Rightarrow \lambda = 3, \lambda = -1$$

Thus A has two eigen values $\lambda_1 = 3, \lambda_2 = -1$

<u>To find the eigen vector corresponding to</u> $\lambda_1 = 3$:

$$(A-\lambda_1 I)v = O \Rightarrow \begin{bmatrix} 1-3 & 1 \\ 4 & 1-3 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = O$$

$$\Rightarrow \begin{bmatrix} -2 & 1 \\ 4 & -2 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = O$$ $R_2 + 2R_1$ gives

$$\begin{bmatrix} -2 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow -2v_1 + v_2 = 0 \to v_2 = 2v_1$$

put $v_1 = t,$ then $v_2 = 2t$ so $v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} t \\ 2t \end{bmatrix} = t\begin{bmatrix} 1 \\ 2 \end{bmatrix}$

Where $t \in R$

Thus $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ is the eigen vector corresponding to $\lambda_1 = 3$

<u>To find the eigen vector corresponding to</u> $\lambda_2 = -1$:

$$(A-\lambda_2 I)v = O \Rightarrow \begin{bmatrix} 1+1 & 1 \\ 4 & 1+1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = O$$

$$\Rightarrow \begin{bmatrix} 2 & 1 \\ 4 & 2 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = O$$ $R_2 - 2R_1$ gives

$$\begin{bmatrix} 2 & 1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow 2v_1 + v_2 = 0 \rightarrow v_2 = 2v_1$$

put $v_1 = s$ them $v_2 = -2s$

Thus $v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} s \\ -2s \end{bmatrix} = s\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ where s $\in R$

So $\begin{bmatrix} 1 \\ -2 \end{bmatrix}$ is the eigen vector corresponding to $\lambda_2 = -1$.

$\left\{\begin{bmatrix} 1 \\ 2 \end{bmatrix}\begin{bmatrix} 1 \\ -2 \end{bmatrix}\right\}$ is a basis of R^2, since there vectors are linearly in dependent and dim $R^2 = 2$.

This is basis of R^2 consisting of eigen vectors of A. So L_A and hence A is diagonalizable

W.E.18: Find the matrix P which diagonalizes the matrix $A = \begin{bmatrix} -2 & 2 & -3 \\ 2 & 1 & -6 \\ -1 & -2 & 0 \end{bmatrix}$ and verify that $P^{-1}AP$ is a diagonal matrix.

Solution: The characteristic equation of A is $|A - \lambda I| = 0$

trace of $A = -2 + 1 + 0 = -1$

$$A_{11} = (-1)^{1+1}\begin{vmatrix} 1 & -6 \\ -2 & 0 \end{vmatrix} = 0 - 12 = -12$$

$$A_{22} = (-1)^{2+2}\begin{vmatrix} -2 & -3 \\ -1 & 0 \end{vmatrix} = 0 - 3 = -3$$

$$A_{33} = (-1)^{3+3}\begin{vmatrix} -2 & 2 \\ 2 & 1 \end{vmatrix} = -2 - 4 = -6$$

$$A_{11} + A_{22} + A_{33} = -12 - 3 - 6 = -21$$

$$\det A = -2(0 - 12) - 2(0 - 6) - 3(-4 + 1)$$

$$= 24 + 12 + 9 = 45$$

The characteristic equation is

$$\lambda^3 - \lambda^2 .\text{ trace of } A + (A_{11} + A_{22} + A_{33})\lambda - \det A = 0$$

i.e. $f(\lambda) = \lambda^3 - (-1)\lambda^2 + (-21)\lambda + 45 = 0$

$$= \lambda^3 + \lambda^2 - 21\lambda + 45 = 0$$

$$f(-3) = -27 + 9 + 63 - 45$$

$$= 0$$

$\lambda + 3$ is a factor.

–3	1	+1	–21	–45
		–3	6	+45
	1	–2	–15	0

The other factor is $\lambda^2 - 2\lambda - 15 = (\lambda - 5)(\lambda + 3)$

So $f(\lambda) = (\lambda + 5)^2 (\lambda - 5) = 0$

The eigen values are $\lambda = -3, -3, 5$

i) To find the eigen space corresponding to $\lambda_1 = 5$.

$$(A - \lambda_1 I)v = O$$

$$\Rightarrow \begin{bmatrix} -2-5 & 2 & -3 \\ 2 & 1-5 & -6 \\ -1 & -2 & -5 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = O$$

$$\Rightarrow \begin{bmatrix} -7 & 2 & -3 \\ 2 & -4 & -6 \\ -1 & -2 & -5 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = O$$

$R_1 - 7R_3, R_2 + 2R_3$ given

$$\begin{bmatrix} 0 & 16 & 32 \\ 0 & -8 & -16 \\ -1 & -2 & -5 \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = O$$

$R_1 + 2R_2$ gives

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & -8 & -16 \\ -1 & -2 & -5 \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Rightarrow -8v_2 - 16v_3 = 0 \Rightarrow v_2 = -2v_3$$

$$-v_1 - 2v_2 - 5v_3 = 0$$

$-v_1 + 4v_3 - 5v_3 = 0$ since $v_2 = -2v_3$

$$\Rightarrow v_3 = -v_1$$

put $v_1 = 1$ then $v_3 = -1, v_2 = 2$

$v = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$ is an eigen vector corresponding to $\lambda_1 = 5$.

$C_1 = \left\{ \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} \right\}$ is a basis of the eigen space $E_{\lambda_1} = \left\{ \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \in R^3 \middle/ (A - \lambda_1 I)v = 0 \right\}$

ii) To find the eigen space corresponding to $\lambda_2 = -3$

$$(A - \lambda_2 I)v = O$$

$$\Rightarrow \begin{bmatrix} -2+3 & 2 & -3 \\ 2 & 1+3 & -6 \\ -1 & -2 & -3 \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = O$$

$$\Rightarrow \begin{bmatrix} 1 & 2 & -3 \\ 2 & 4 & -6 \\ -1 & -2 & +3 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = O$$, $R_2 - 2R_1, R_3 + R_1$ give

$$\begin{bmatrix} 1 & 2 & -3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = O ,$$

So $v_1 + 2v_2 - 3v_3 = 0$

$v_1 = -2v_2 + 3v_3$

put $v_2 = -s; v_3 = t$

Then $v = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 2s+3t \\ -s+0t \\ 0s+1t \end{bmatrix} = s \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix} + t \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix}$

So $C_2 = \left\{ \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix} \right\}$ is a basis corresponding to $E_{\lambda_1} = \left\{ \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \in R^3 \Big/ (A - \lambda_2 I)v = O \right\}$

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

$$= -1 - 4 - 3 = -8$$

$$P^{-1} = \frac{1}{\det P} Adj.A = \frac{-1}{8} \begin{bmatrix} -1 & -2 & 3 \\ -2 & 4 & 6 \\ -1 & -2 & -5 \end{bmatrix}$$

$$P^{-1}A = \frac{-1}{8} \begin{bmatrix} -1 & -2 & 3 \\ -2 & 4 & 6 \\ -1 & -2 & -5 \end{bmatrix} \begin{bmatrix} -2 & 2 & 3 \\ -2 & 4 & 6 \\ -1 & -2 & -5 \end{bmatrix} = \frac{-1}{8} \begin{bmatrix} -5 & -10 & 15 \\ 6 & -12 & -18 \\ 3 & 6 & 15 \end{bmatrix}$$

$$P^{-1}AP = \frac{-1}{8}\begin{bmatrix} -5 & -10 & 15 \\ 6 & -12 & -18 \\ 3 & 6 & 15 \end{bmatrix}\begin{bmatrix} 1 & 2 & 3 \\ 2 & -1 & 0 \\ -1 & 0 & 1 \end{bmatrix} = \frac{-1}{8}\begin{bmatrix} -40 & 0 & 0 \\ 0 & 24 & 0 \\ 0 & 0 & 24 \end{bmatrix}$$

$$= \begin{bmatrix} 5 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & -3 \end{bmatrix} = \text{diag } (5,-3,-3)$$

Hence $P^{-1}AP$ is a diagonal matrix.

W.E.19: Find the matrix P which transforms the matrix $A = \begin{bmatrix} 1 & 0 & -1 \\ 1 & 2 & 1 \\ 2 & 2 & 3 \end{bmatrix}$ to diagonal form and hence calculate A^4.

Solution: Characteristic equation is $|A - \lambda I| = 0$

trace of $A = 1 + 2 + 3 = 6$

$$A_{11} = (-1)^{1+1}\begin{vmatrix} 2 & 1 \\ 2 & 3 \end{vmatrix} = 6 - 2 = 4$$

$$A_{22} = (-1)^{2+2}\begin{vmatrix} 1 & -1 \\ 2 & 3 \end{vmatrix} = 3 + 2 = 5$$

$$A_{33} = (-1)^{3+3}\begin{vmatrix} 1 & 0 \\ 1 & 2 \end{vmatrix} = 2 - 0 = 2$$

$$A_{11} + A_{22+}A_{33} = 4 + 5 + 2 = 11$$

$$\det A = 1(6-2) - 0 + (-1)(2-4) = 4 + 2 = 6$$

The characteristic equation is

$$\lambda^3 - \lambda^2 \text{ trace of } A + \lambda(A_{11} + A_{22} + A_{33}) - \det A = 0$$

$$f(\lambda) = \lambda^3 - 6\lambda^2 + 11\lambda - 6 = 0$$

$f(1) = 1 - 6 + 11 - 6 = 0$ So $\lambda - 1$ is a factor.

$$\begin{array}{r|rrrr} 1 & 1 & -6 & 11 & -6 \\ & & +1 & -5 & -6 \\ \hline & 1 & -5 & 6 & \multicolumn{1}{|r}{0} \end{array}$$

The other factor is $\lambda^2 - 5\lambda + 6$

$$= (\lambda - 5)(\lambda - 2)$$

Hence $f(\lambda) = (\lambda - 1)(\lambda - 2)(\lambda - 3) = 0$

The eigen values of A are 1, 2, 3

i) To find the eigen space corresponding to $\lambda = 1$:

$$(A - \lambda_1 I)v = O \Rightarrow \begin{bmatrix} 1-1 & 0 & -1 \\ 1 & 2-1 & 1 \\ 2 & 2 & 3-1 \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = O$$

$$\Rightarrow \begin{bmatrix} 0 & 0 & -1 \\ 1 & 1 & 1 \\ 2 & 2 & 2 \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = O$$

$R_3 - 2R_2$ gives $\begin{bmatrix} 0 & 0 & -1 \\ 1 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} . \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = O$

$$-v_3 = 0 \Rightarrow v_3 = 0$$

$$v_1 + v_2 + v_3 = 0$$

i.e. $v_1 + v_2 = 0$

$$\Rightarrow v_2 = -v_1$$

Put $v_1 = 1,$ them $v_2 = -1, v_3 = 0$

So $v = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$

So $C_1 = \left\{ \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \right\}$ is a basis of the eigen space $E_{\lambda_1} = \left\{ \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \in R^3 \middle/ (A - \lambda_1 I)v = O \right\}$

ii) To find eigen space corresponding to $\lambda_2 = 2$

$$(A - \lambda_2 I)v = O \Rightarrow \begin{bmatrix} 1-2 & 0 & -1 \\ 1 & 2-2 & 1 \\ 2 & 2 & 3-2 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = O \Rightarrow \begin{bmatrix} -1 & 0 & -1 \\ 1 & 0 & 1 \\ 2 & 2 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = O$$

$R_2 + R_3$ gives

$$\begin{bmatrix} -1 & 0 & -1 \\ 0 & 0 & 0 \\ 2 & 2 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = O \Rightarrow -v_1 - v_3 = 0 \text{ so } v_3 = -v_1$$

$$2v_1 + 2v_2 + v_3 = 0$$

$$\Rightarrow 2v_1 + 2v_2 - v_1 = 0$$

$$\Rightarrow v_1 + 2v_2 = 0 \quad v_1 = -2v_2$$

if $v_2 = 1, v_1 = -2, v_3 = 2$

$v = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} -2 \\ 1 \\ 2 \end{bmatrix}$ so $C_2 = \left\{ \begin{bmatrix} -2 \\ 1 \\ 2 \end{bmatrix} \right\}$ is a basis of eigen space $E_{\lambda_2} = \left\{ \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \in R^3 \middle/ (A - \lambda_2 I)v = O \right\}$

iii) To find the eigen space corresponding to $\lambda_3 = 3$.

$$\begin{bmatrix} 1-3 & 0 & -1 \\ 1 & 2-3 & 1 \\ 2 & 2 & 3-3 \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = O \Rightarrow \begin{bmatrix} -2 & 0 & -1 \\ 1 & -1 & 1 \\ 2 & 2 & 0 \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = O$$

$$\Rightarrow -2v_1 - v_3 = 0 \Rightarrow v_3 = -2v_1$$

$$2v_1 + 2v_2 = 0 \Rightarrow v_2 = -v_1$$

$$v_1 - v_2 + v_3 = 0 \Rightarrow v_1 + v_1 + -2v_1 = 0$$

put $v_1 = -1$, then $v_2 = 1, v_3 = 2$

So $v = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix}$ Hence $C_3 = \left\{\begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix}\right\}$ is a basis of $E_{\lambda_3} = \left\{\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \in R^3 \middle/ (A - \lambda_3 I)v = O\right\}$

Writing the three eigen vectors of the matrix as the three columns, the required transformation matrix $P = \begin{bmatrix} 1 & -2 & -1 \\ -1 & 1 & 1 \\ 0 & 2 & 2 \end{bmatrix}$

det $P = 1(2-2) + 2(-2-0) - 1(-2-0)$

$$= 0 - 4 + 2 = -2$$

$$P^{-1} = \frac{1}{\det P} Adj.P$$

$$= \frac{-1}{2}\begin{bmatrix} 0 & 2 & -1 \\ 2 & 2 & 0 \\ -2 & -2 & -1 \end{bmatrix}$$

Now $P^{-1}AP = \dfrac{-1}{2}\begin{bmatrix} 0 & 2 & -1 \\ 2 & 2 & 0 \\ -2 & -2 & -1 \end{bmatrix}\begin{bmatrix} 1 & 0 & -1 \\ 1 & 2 & 1 \\ 2 & 2 & 3 \end{bmatrix}\begin{bmatrix} 1 & -2 & -1 \\ -1 & 1 & 1 \\ 0 & 2 & 2 \end{bmatrix}$

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

$$= \frac{-1}{2}\begin{bmatrix} -2 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & -6 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} = D \text{ (say)}$$

$$P^{-1}AP = D \Rightarrow A = PDP^{-1}$$

$$A^4 = PD^4P^{-1}$$

$$= \begin{bmatrix} 1 & -2 & -1 \\ -1 & 1 & 1 \\ 0 & 2 & 2 \end{bmatrix}\begin{bmatrix} (1)^4 & 0 & 0 \\ 0 & (2)^4 & 0 \\ 0 & 0 & (3)^4 \end{bmatrix}\begin{bmatrix} 0 & 2 & -1 \\ 2 & 2 & 0 \\ -2 & -2 & -1 \end{bmatrix} \times \left(\frac{-1}{2}\right)$$

$$= \begin{bmatrix} 1 & -2 & -1 \\ -1 & 1 & 1 \\ 0 & 2 & 2 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 16 & 0 \\ 0 & 0 & 81 \end{bmatrix} \times \left(\frac{-1}{2}\right)\begin{bmatrix} 0 & 2 & -1 \\ 2 & 2 & 0 \\ -2 & -2 & -1 \end{bmatrix}$$

$$= \left(\frac{-1}{2}\right)\begin{bmatrix} 1 & -32 & -81 \\ -1 & 16 & 81 \\ 0 & 32 & 162 \end{bmatrix}\begin{bmatrix} 0 & 2 & -1 \\ 2 & 2 & 0 \\ -2 & -2 & -1 \end{bmatrix}$$

$$A^4 = \frac{-1}{2}\begin{bmatrix} 98 & 100 & 80 \\ -130 & -132 & -81 \\ -260 & -260 & -162 \end{bmatrix} = \begin{bmatrix} -49 & -50 & -40 \\ 65 & 66 & 40 \\ 130 & 130 & 81 \end{bmatrix}$$

12.19 Invariant Subspaces:

We observed if v is an eigen vector of a linear operator T, then T maps the span of $\{v\}$ into itself. Subspaces that are mapped into themselves are great importance in the study of linear operator.

12.19.1 T invariant Subspace: Let T be a linear operator on a vector space V. A subspace ω of V is said to be a T - invariant subspace of V if $T(W) \subseteq W$, that is, if $T(v) \in W$ for all $v \in \omega$.

W.E. 20: Example: Suppose that T is a linear operator on a vectorspace V; Then the following subspaces are T - invariant (i) $\{O\}$, (ii) V (iii) R (T) (iv) N (T) (v) E_λ for any eigen value λ of T.

Solution : i) To show $\{0\}$ is T - invariant.

Let $W_1 = \{0\}$. We know that W_1 is a subspace of V.

Also $T(0) = 0 \in W_1$ for $0 \in v$

Thus W_1 is a T invariant subspace of V.

ii) To show that V is T - invariant.

We know V is a subspace of V.

Let $v \in V$ then $T(v) \in V$ for all $v \in V$ which proves that V is a T - invariant subsspace of V.

iii) To show range of T i.e. $R(T)$ is T invariant.

We know that $R(T)$ is a subspace of V.

Let $u \in R(T) \leq V \Rightarrow u \in V$

$\therefore T(u) \in R(T)$ for all $u \in R(T)$ so $R(T)$ is a T - invariant subspace of V.

iv) To show Null space $N(T)$ is T - invarient.

T ' is a linear operator on V.

$N(T) = \{\alpha \in V / T(\alpha) = O(\in V)\}$ and $N(T)$ is a null space of V i.e. a subspace of V.

$T(N(T)) \subseteq N(T)$ i.e $T(u) \in N(T)$ for all $u \in N(T)$.

$\therefore N(T)$ is a T - invariant subspace of V.

W.E. 21: T is the linear operator on R^3 defined by $T(a,b,c) = (a+b, b+c, 0)$

Then xy plane $= \{(x, y, 0)/x, y \in R\}$ and x axis $= \{(x, 0, 0)/x \in R\}$ are T - invariant subspaces of R^3.

Solution: Given $T : R^3 \to R^3$ is defined by $T(a,b,c) = (a+b, b+c, 0)$

We know that $W_1 = xy$ plane = $\{(x, 0, 0)/x \in R\}$ is a subspace of R^3.

But $v=(x,0,0)\in W_1$

$T(v)=T(x,y,0)=(x+y,y+0,0)=(x+y,y,0)\in W_1$ for all $v\in W_1$

Thus W_1 is a T - invariant subspace of R^3.

Similarly Let $W_2 = X-axis=\{(x,0,0/x\in R\}$

We know that W_2 is a subspace of R^3.

Let $v=(x,0,0)\in W_2$

$T(v)=T(x,0,0)=(x+0,0+0,0)=(x,0,0)\in W_2$ for all $v\in W_2$. Thus W_2 is an invariant subspace of R^3.

12.20 T - Cyclic subspace of V generated by a non zero vector:

12.20.1 Definition: Let T be a linear operator on a vector space V and let v be a nonzero vector in V, The subspace W = span $\left(\{v,T(v),T(v^2)...\}\right)$ is called the T - Cyclic subspace of V generated by v.

We can easily prove that W is T - invariant subspace of V. In fact, W is the smallest T - invariant subspace of V containing non zero vector v.

W.E. 22:

T is a linear operator on $P(R)$ defined as $T\left(f(x)\right)=f^{'}(x)$. Find the T - Cyclic subspace generated by x^2.

Solution: We have $T\left[f(x)\right]=f^{'}(x)$ and so

$$T(x^2)=(x^2)^{'}=2x$$

$$T^2(x^2)=T\left[T(x^2)^{'}\right]=T\left[(2x)\right]=(2x)^{'}=2$$

T - Cyclic subspace generated by x^2 = Span $\left[\{x^2,2x,2\}\right]=P_2(x)$

W.E. 23:

Let T be the linear operator on R^3 defined by $T(a,b,c)=\{-b+c,a+c,3c\}$. Find the

T - Cyclic subspace generated by $e_1 = (1,0,0)$.

Solution: Given $T: R^3 \to R^3$ is defined by

$$T(a,b,c) = \{-b+c, a+c, 3c\}$$

$$T(e_1) = T(1,0,0) = e_2$$

$$T^2(e_1) = [T(e_1)] = T(e_2) = T(0,1,0) = (-1,0,0) = -e_1$$

$$T^3(e_1) = [T^2(e_1)] = T(-e_1) = T(-1,0,0) = (0,-1,0) = -e_2$$

Thus T cyclic subspace generated by $e_1 = \{e_1, T(e_1), T^2(e_1), ...\}$

= Span $\{e_1, e_2\} = [(s,t,o)/s,t \in R]$

12.20.2 Theorem: Let T be a linear operator on a finite dimensional vector space V, and let W be a T - invariant subspace of V. Then the characteristic polynomial of T_W devides the characteristic polynomial of T.

Proof: Choose an ordered basis $C = \{v_1, v_2, ...v_k\}$ for W, and it is extended to form an ordered basis $B = \{v_1, v_2, ...v_k, v_{k+1}......v_n\}$ for V. Let $A = [T]_B$ and $B_1 = [T_W]_c$. So A can be written as $\begin{bmatrix} B_1 & B_2 \\ 0 & B_3 \end{bmatrix}$.

Let $f(t)$ be the characteristic polynomial of T and $g(t)$ the characteristic polynomial of T_W.

Then $f(t) = \det(A - tI_n) = \det \begin{bmatrix} B_1 - tI_k & B_2 \\ 0 & B_3 - tI_{n-k} \end{bmatrix}$

$$g(t).\det(B_3 - tI_{n-k})$$

Thus $g(t)$ divides $f(t)$.

W.E. 24: T is a linear operator on R^4 defined by $T(a,b,c,d) = (a+b+2c-d, b+d, 2c-d, c+d)$.

If $W = \{(t,s,o,o)/t,s \in R\}$ is a subspace of R^4. Verify that characteristic polynomial of T_W divides the characteristic polynomial of T.

Solution: Given $T: R^4 \to R^4$ is a linear operator defined by $T(a,b,c,d) = (a+b+2c-d, b+d, 2c-d, c+d)$

$W = \{(t,s,o,o)/t,s \in R\}$ is a subspae of R^4.

Let $(a,b,0,0) \in R^4$ then $T(a,b,0,0) = (a+b,b,0,0) \in W$

consider $C = \{e_1, e_2\}$ which is an ordered basis of W. Extend this to the standard ordered basis B of R^4.

Then $B_1 = [T_W]_c = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ and $A = [T]_B = \begin{bmatrix} 1 & 1 & 2 & -1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 2 & -1 \\ 0 & 0 & 1 & 1 \end{bmatrix}$

Let the characteristic polynomial of the $f(t)$ and $g(t)$ be the characteristic polynomial of T_W.

Then $f(t) = |A - tI| = \begin{vmatrix} 1-t & 1 & 2 & -1 \\ 0 & 1-t & 0 & 1 \\ 0 & 0 & 2-t & -1 \\ 0 & 0 & 1 & 1-t \end{vmatrix}$

$$= (1-t)\begin{vmatrix} 1-t & 0 & 1 \\ 0 & 2-t & -1 \\ 0 & 1 & 1-t \end{vmatrix} = (1-t)(1-t)\begin{vmatrix} 2-t & -1 \\ 1 & 1-t \end{vmatrix}$$

$$= g(t).\left[(2-t)(1-t)+1\right]$$

$$= g(t)(t^2 - 3t + 3)$$

Thus $g(t)$ divides $f(t)$.

Thus the characteristic polynomial of T_W divides the characteristic polynomial of T.

12.20.3 Theorem: Let T be a linear operator on a finite dimensional vector space V, and let w denote the T - Cyclic subspace of V generated by a non zero vector $v \in V$. Let $k = \dim(W)$

Then

i) $\{v, t(v), T^2(v), ...T^{k-1}(v)\}$ is a basis of w.

ii) If $a_0 v + a_1 T(v) + ... + a_{k-1} T^{k-1}(v) + T^k(v) = 0$

then the characteristic polynomial of T_W is $f(t) = (-1)^k (a_0 + a_1 t + a_2 t^2 + ... + a_{k-1} t^{k-1} + t^k)$

Proof: i) Since $v \neq O$ the set $\{v\}$ is linearly independent.

Let j be the largest positive integer for which $B = \{v, T(v), T^2(v).....T^{j-1}(v)\}$ is linearly independent. Such a j must exist because V is finite dimensional. Let Z = Span (B). Then B is a basis for Z. Further more $T^j(v) \in Z$. We use this fact to show that Z is a T - invariant subspace of V. Let $w \in Z$. Since w is a linear combination of the vectors of B, there exists scalars $b_0, b_1, ...b_{j-1}$ such that $w = b_0 v + b_1 T(v) + ... + b_{j-1} T^{j-1}(v)$ and hence $T(w)$ is a linear combination of vectors in Z and hence belongs to Z. So Z is T - invariant. Further more, $v \in Z$, and W is the smallest T - invariant subspace of V, that contains so that . Clearly, , and so we conclude that . It follows that B is a basis for , and therefore . Thus .

Hence is a basis of .

ii) Now view B as an ordered basis for .

Let be scalars such that

Observe that

Which has the characteristic polynomial.

Thus is the characteristic polynomial of , proving (ii).

12.21.1 Cayley Hamilton Theorem for Linear Operators:

Theorem:

Let T be a linear operator on the n dimensional vector space V. (or let A be matrix over the field F) and be the characteristic polynomial for T (or for A) then (zero transformation (or (null matrix)

Or

Every square matrix satisfies its characteristic equation

Or

Every matrix is zero of its characteristic polynomial.

Proof: Consider an n square matrix over a field F for relative to an ordered basis B.

i.e.

The characteristic polynomial of A is given by

i.e. .... (1) for s

The characteristic equation is .

i.e.

The elements of the matrix are polynomials at most of the first degree in t with the result that the elements of the matrix Adj are ordinary polynomials in t of degree or less. As we know that the elements of the matrix Adj are the cofactors of the elements of the matrix . It implies that the matrix adj can be written as

.... (2)

Where is a square matrix of order n, over F with elements independent of t. Now by the property of adjoints we have

Or

Equating the coefficients of corresponding powers of t we get

Multiplying the above matrix equations by respectively we get

Thus

Also we have

or

Thus because

Hence the theorem.

From this theorem, we conclude that if T is a linear transformation on the n dimensional vector space ; then there is a polynomial f of degree n, such that .

Corollary: To find an expression for the inverse of a nonsingular matrix A:

Solution: We have

and i.e.

Thus from the cayley - Hamilton theorem we have

So

or

correspondingly

giving an expression for the inverse of a non singular matrix A.

Aliter : To show every linear operator satisfies its characteristic equation.

Let T be a linear operator defined on V and be its characteristic polynomial we will show that for all . If , as is linear we have Let Let be the T - cyclic subspace generated by . Let . Then is a basis for .

Hence there exists scalars

Such that .... (1)

Hence .... (2) is the characteristic polynomial of

From (1) and (2) we get

But we have divides . Hence there exists a polynomial such that

So

Hence T satisfies its characteristic equation.

12.21.2 Theorem: Cayley - Hamilton theorem for matrices:

Every square matrix satisfies its characteristic equation. i.e. If for a square matrix A of order n;

then the matrix equation

is satisfied by A. i.e. .

Proof: As the elements of are at most of the first degree in , the elements of can be written as a matrix polynomial in given by

where are matrices of the type whose elements are functions of s.

Now

Since A. Adj A = det A.I.

So

comparing coefficients of like powers of on both sides we get

=

Pre multiplying there successively by and adding we get

Thus = **O** .... (1)

So every square matrix satisfies its characteristic equation.

Corollary 1: Let A be a non singular matrix i.e. also and therefore .

Pre multiplying (1) by we get

or

Corollary 2: If m be a positive integer such that , then multiplying the result (1) by we get

Showing that any positive integral power of of A is linearly expressible in terms of those lower order.

W.E. 25: Worked Out Examples:

Let T be a linear operator on defined by and is the standard basis of . Show that T satisfies its characteristic equation and satisfies its characteristic equation.

Solution: is a linear operator defined by

is the standard basis of where and

The characteristic polynomial of T is

So

=

Hence T satisfies its characteristic equation more over

so

Thus satisfies its characteristic equation.

W.E. 26 :

is a linear operator defined by . Find the characteristic polynomial. Verify cayley - Hamilton theorem.

Solution: is alinear transformation defined by

is the standard ordered basis of T where

The characteristic equation is .

i.e.

We have to show where O is the zero operator from .

=

and

So

So

So T satisfies its characteristic equation.

W.E. 27: Find the characteristic equation of the matrix

and verify that it is satisfied by A and hence find .

Solution: The characteristic equation of the matrix is

or

We have to show that where O is the null matrix.

So

So

Hence A satisfies its characteristic equation preoperating with we get

(null matrix)

So

W.E. 28:

If then express as a linear polynomial in A by using Cayley. Hamilton theorem.

Solution: The characteristic equation is .

.

By Cayley - Hamilton theorem every matrix satisfies its own characteristic equation.

So = O (null matrix)

.... (1)

The given linear polynomial is using (1) we get

since

Thus is expressed as a linear polynomial .

Aliter:

By (1) (Zero matrix)

So .... (2)

Multiply (2) by

.... (3)

.... (4)

..............(5)

Now

using (5)

using (4)

using (3)

using (2)

which is a linear polynomial.

W.E. 29 :

Using Cayley - Hamilton theorem find the inverse and of the matrix

trace of

det

Characteristic equation is

trace of

.... (1)

By Cayley - Hamilton theorem every matri satisfies its characteristic equation

So .... (2)

Pre operating with we get

=

By (2)

multiplying by A

So

W.E. 30:

Find the inverse of the matrix by using cayley hamilton theorem.

Solution: We know if A is an n x n square matrix, its characteristic polynomial is

Where is the sum of the principal minors of order K.

By this result, the characteristic polynomial of the given matrix

given by

trance of A and is det. of A.

To find (i.e. the sum of principal minors of order 2)

The rows and columns deleted.

1)

2)

3)

4)

5)

6)

To find the sum of principal minors of order 3 i.e. :

Deleted row and column

1)

2)

3)

4)

gives

expanding along the first column.

Hence the characteristic equation is

i.e.

By Cayley Hamilton theorem every matrix satisfies its characteric equation.

Hence

Pre operating with .

So .... (1)

By (1) .

So

12.22 Summary:

In this lesson we discussed about

i) Eigen vectors and eigen values of linear operator and of a square matrix - Properties of eigen values. Diagonalizability of linear operators and matrices. Test for diagonalization and numerical problems - Invariant subspaces. Cayley-Hamilton theorem for linear operators and matrices.

12.23 Technical Terms:

In this chapter we come across the following technical terms.

Eigen vectors

Eigen values

Characteristic equation

Diagonalization

Trace of a linear operator

Eigen space.

12.24 Model Questions:

1. Prove that if the characteristic roots of a matrix A are then the characteristic roots of are .

2. If is an eigen value of a nonsingular matrix A, then show that is an eigen value of .

3. Find the eigen values and the corresponding eigen vectors of the matrix

i)

Ans : i) 2, 2, 8 Characteristic vector corresponding 2 is where a, b are scalars and coresponding to 8 is where k is scalar.

ii)

Ans : ii) 0, 3, 15, k, l, m are scalars.

4) Using Cayley - Hamilton theorem show that where

5) State and prove Cayley - Hamilton Theorem.

6) If then express as a linear polynomial in A by using Cayley Hamilton theorem.

Ans:

7) Find the characteristic equation of the matrix A and verify that it is satisfied by A and hence find .

i)

Ans:

ii)

Ans:

8) Find the inverse of the matrix by Cayley Hamilton theorem

Ans :

12.25 Exercises:

1. For each of the following matrices test A for diagonalizability and if A is diagonalizable, find an invertible matrix Q and a diagonal matrix D such that

(i) Ans: diagonalizable

ii) Ans: Not diagonalizable

iii) Ans : Diagonalizable

iv) Ans : Diagonalizable

2. For each of the following linear operators T on a vector space V, test T for diagonalizability and if T is diagonalizable find a basis B for V such that is a diagonal matrix.

i) and T is defined by respectively.

Ans: T is not diagonalizable

ii) and T is defined by

Ans: T is not diagonalizable.

iii) and T is defined by

Ans: T is diagonalizable

iv) and T is defined

Ans : T is diagonalizable

3. Prove that if T is diagonalizable, then is diagonalizable, when T is a linear operator on a finite dimensional vector space V.

4. Show that is not diagonalizable.

5. Prove that the matrix is not diagonalizable over the field C.

6. For the matrix over the field C, Find the diagonal form and a diagonalizing matrix Q.

Ans :

7. Let T be a linear operator on which is represented in the standard ordered basis by the matrix . find the characteristic values of A and prove that T is diagonalizable.

Ans: 1, 2, 2

8) Let T be a linear operator on a finite dimensional vector space V; and let be an eigen value of T. Show that the eigen space of i.e. is invariant under T.

9) For each of the following linear operators T on the vector space V, determine whether the given subspace is a T - invariant subspace of V.

(i) and

Ans: T - invariant

(ii) and

Ans: T - invariant

(iii) and

Ans: not T - invariant

10) Find the characteristic equation of the matrix and verify that it is satisfied by A and hence find .

Ans:

11) Verify that the matrix satisfies its characteristic equation and conpute .

Ans :

12) Find the characteristic equation of the matrix and show that it is satisfied by A. Hence obtain n the inverse of the given matrix.

Ans :

13) Find the characteristic roots of the matrix and verify cayley - Hamilton theorem for this matrix. Find the inverse of the matrix A and also express as a linear polynomial.

Ans :

The given matric polynomial = .

14) If express as a linear polynomial in A.

Ans: - .

15) Calculate by using Cayley - Hamilton theorem given

Ans :

16) If show that and hence find .

Ans : $B^5 = \begin{bmatrix} -61 & 93 \\ -62 & 94 \end{bmatrix}$

12.26 Reference Books:

1) Linear Algebra 4th edition : Stephen H. Friedberg, Arnold J. Insel, Lawrence E. Spence.

2) Schaum's out lines : Beginning Linear Algebra Seymour Lipschutz

3) A course in abstract Algebra : Vijay K . khanna. S.K. Bhambari

4) Linear Algebra : Gupta and Sharma

5) Fundamentals of Linear Algebra : M.L. Aggarwal, Romesh Kumar

- A. Mallikharjana Sarma

LESSON - 13

INNER PRODUCT SPACES

13.1 Objective of the Lesson:

In this previous lessons the properties of vector spaces discussed are based on addition and scalar multiplication of vectors. In this lesson we introduce the concept of length of vectors by means of an additional structure on the vector space known as inner product.

13.2. Structure of the Lesson:

In this lesson the following concpts are discussed:

13.3 Introduction and properties of complex numbers

13.4 Inner product and inner product space - definitions - examples - basic theorems

13.5 Norm of a vector definition, theorems in inner product spaces

13.6 Norm of a vector, normed vector spaces - definitions and theorems

13.7 Worked out examples

13.8 Summary

13.9 Technical terms

13.10 Model Questions

13.11 Exercises

13.12 Reference Books

13.3.1 Introduction:

In general a vector space is defined over an arbitrary field F. In this lesson we restrict the field F to be the field of real numbers or complex numbers. In th first case the vector space is called a real vector space and in the second case it is called a complex vector space. We study real vector space in analytical geometry and vector analysis. There the concept of length and orthogonality is disscussed.

In this lesson we introduce the concept of length and orthogonality of vectors by means of an additional structure on the vector space known as an inner product.

We also have dot or scalar product of two vectors whose properties are discussed in vector algebra. An inner product on a vector space is a generatisation of dot product in R^3.

Before defining inner product and inner product spaces we shall state some important properties of complex numbers.

13.3.2 Some Properties of Complex Numbers:

Let $Z = x + iy$ for some $x; y \in R$ and $i = \sqrt{-1}$ be the given complex number. Here x is called the real part of the complex number Z. y is called the imaginary part of Z and we write $x = \operatorname{Re} z$ and $y = I_m Z$.

The modules of the complex number $Z = x + iy$ is the non negative real number $\sqrt{x^2 + y^2}$ and is denoted by $|z|$.

Also if $z = x + iy$ is a complex number then $\overline{z} = x - iy$ is called the conjugate complex number of z.

If $z = \overline{z}$, then $x + iy = x - iy$ so $y = 0$.

i.e. $z = \overline{z} \Rightarrow z$ is a real number. Obviously we have

i) $z + \overline{z} = 2x = 2\operatorname{Re} z$

ii) $z - \overline{z} = 2iy = 2\operatorname{Im} z$

iii) $z\overline{z} = x^2 + y^2 = |z|^2$

iv) $|z| = 0 \Rightarrow x = 0, y = 0$ i.e. $|z| = 0 \Rightarrow z = 0$

v) $\overline{(\overline{z})} = z$

vi) $|\overline{z}| = |z|$

vii) $|z| = \sqrt{x^2 + y^2} \geq 0$ i.e. $|z| \geq \operatorname{Re} z$

viii) If z_1, z_2 are two complex numbers then

i) $|z_1 + z_2| \leq |z_1| + |z_2|$

ii) $\overline{z_1 + z_2} = \overline{z}_1 + \overline{z}_2$

iii) $\overline{z_1 z_2} = \overline{z}_1 . \overline{z}_2$

iv) $\overline{z_1 - z_2} = \overline{z}_1 - \overline{z}_2$

v) $\overline{\left[\frac{z_1}{z_2}\right]} = \frac{\overline{z}_1}{\overline{z}_2}$ provided $z_2 \neq 0$

13.4 Inner Product and Inner Product Space:

13.4.1 Definition: Let V be a vector space over F. An inner product on V is a function that assigns to every order pair of vectors u and v in V, a scalar in field F denoted by $< u, v >$ such that for all u, v in V, and for all $a \in F$ the following holds good.

i) $< u + v, w > = < u, w > + < v, w >$ (Linearity)

ii) $< au, v >= a < u, v >$ Linearity

There two conditions can be clubbed as $< au + bv, w >= a < u, w > +b < v, w >$

iii) $\overline{< u, v >} =< v, u >$ where bar denotes the complex conjagate (conjugate symmetry)

iv) $< u, v > \geq 0$ and $< u, u >= 0 \Leftrightarrow u = 0$ where O denotes the zero vector in V (non negativity)

13.4.2. Definition II: A vector space together with an inner product defined on it is called an inner product space.

Thus inner product space is a vector space over the field of real or complex numbers with an inner product function.

Note: a) Conditions (i) and (ii) simply requires that the inner product is linear in the first component.

(b) If $F = R$, the condition (iii) reduces to $< u, v >=< v, u >$. In this case i.e. when $F = R$, the inner product space $V(F)$ is called Eucledian space.

(c) If $F = C$. Then the inner product space $V(F)$ is called a unitary space or complex inner product space.

(d) It can be easily shown that if $a_1, a_2, ... a_n \in F; u_1, u_2, ... u_n$ and $v \in V$, then $< \sum_{i=1}^{n} a_i u_i, v >= \sum_{i=1}^{n} a_i < u_i, v >$

W.E. I:

Worked Examples:

If $u = (a_1, a_2, ... a_n), v = (b_1, b_2, ... b_n) \in V_n(C)$ then show that $< u, v >= a_1\overline{b} + a_2\overline{b_2} + ... + a_n\overline{b_n}$ defines an inner product on $V_n(C)$.

Solution: We will now show that all the postulates of an inner product holds for $< u, v >= a_1\overline{b} + a_2\overline{b_2} + ... + a_n\overline{b_n}$ .... (1)

1) Linearity: Let $w = (c_1, c_2, ... c_n) \in V_n(C)$

Let $a, b, c \in C$ we have

$$au + bv = a(a_1, a_2, ... a_n) + b(b_1, b_2, ... b_n)$$

$$= (aa_1 + bb_1, aa_2 + bb_2, ..., aa_n + bb_n)$$

So $< au+bv; w >= (aa_1+bb_1)\overline{c}_1 + (aa_2+bb_2)\overline{c}_2 + ... + (aa_n+bb_n)\overline{c}_n$

$$= (aa_1\overline{c}_1 + aa_2\overline{c}_2 + ... + aa_n\overline{c}_n) + (bb_1\overline{c}_1 + bb_2\overline{c}_2 + ... + bb_n\overline{c}_n)$$

$$= a(a_1\overline{c}_1 + a_2\overline{c}_2 + ... + a_n\overline{c}_n) + b(b_1\overline{c}_1 + b_2\overline{c}_2 + ... + b_n\overline{c}_n)$$

$$= a < u, w > + b < v, w >$$

Thus $< au+bv, w >= a < u, w > + b < v, w >$

ii) Congugate symmetry : From the definition of the product given in (1)

$$< v, u >= b_1\overline{a}_1 + b_2\overline{a}_2 + ... + b_n\overline{a}_n$$

So $\overline{< v, u >} = \overline{(b_1\overline{a}_1 + b_2\overline{a}_2 + ... + b_n\overline{a}_n)}$

$$= \overline{(b_1\overline{a}_1)} + \overline{(b_2\overline{a}_2)} + ... + \overline{(b_n\overline{a}_n)}$$

$$= \overline{b}_1\overline{(\overline{a}_1)} + \overline{b}_2\overline{(\overline{a}_2)} + ... + \overline{b}_n\overline{(\overline{a}_n)}$$

$$= \overline{b}_1 a_1 + \overline{b}_2 a_2 + ... + \overline{b}_n a_n$$

$$= a_1\overline{b}_1 + a_2\overline{b}_2 + ... + a_n\overline{b}_n$$

$=< u, v >$ by (i) (Since multiplication is commulative)

So $\overline{< v, u >} =< u, v >$

iii) Nonnegativity:

$$< u, u >= a_1\overline{a}_2 + a_2\overline{a}_2 + ... + a_n\overline{a}_n$$

$$= |a_1|^2 + |a_2|^2 + ... + |a_n|^2 \text{ .... (2)}$$

as a_i is a complex number so $|a_i|^2 \geq 0$

So (2) is a sum of n non-negative real numbers and so ≥ 0. Thus $< u, u > \geq 0$ and also $< u, u >= 0 \Rightarrow |a_1|^2 + |a_2|^2 + ... + |a_n|^2 = 0$

$\Rightarrow$ each $|a_i|^2 = 0$ so each $a_i = 0$

So $u = (a_1, a_2, ...a_n) = (0, 0, ...0) = O$

Hence the product defined in (1) is an Inner product on $V_n(C)$ and with respect to this inner product $V_n(C)$ is an inner product space.

13.4.3 Note: i) Standard inner product:

Definition: The inner product $<u,v>$ on $V_n(C)$ defined as $<u,v>= a_1\overline{b_1} + a_2\overline{b_2} + ...a_n\overline{b_n}$ where $u = (a_1, a_2, ...a_n)$ and $v = (b_1, b_2, ...b_n)$ is called the standard inner product on $V_n(C)$.

ii) If u, v are two vectors in $V_n(R)$, then the standard inner product of u, v is given by

$$<u,v>= a_1\overline{b_1} + a_2\overline{b_2} + ... + a_n\overline{b_n}$$

$$= a_1b_1 + a_2b_2 + ... + a_nb_n$$

$\because$ in the field of real nos $\overline{b_1} = b_1$

= $u.v$ which is the dot product of u and v and the inner product $<u,v>$ is denoted by $u.v$.

iii) $<u,v>= a_1(a_1) + a_2(a_2) + ...a_n(a_n)$

$$= a_1^2 + a_2^2 + ...a_n^2 = u.u$$

W.E. 2:

Let V be the vector space over C of all continuous complex valued functions defined on $[0,1]$. If $f, g \in V$; then $< f, g >= \int_0^1 f(t)\overline{g(t)}\,\mathrm{dt}$ defines an inner product.

Solution: Let $f, g, h \in V$ and $a, b \in C$, then

i) Linearity : $\int_0^1 [af(t) + bg(t)]\overline{h(t)}\,\mathrm{dt}$

$$= a\int_0^1 f(t)h(t)\,\mathrm{dt} + b\int_0^1 g(t)\overline{h(t)}\,\mathrm{dt}$$

$$= a < f, h > + b < g, h >$$

ii) Conjugate symmetry:

$$<\overline{g, f}>=\overline{\int_0^1 g(t)\overline{f(t)}\,dt}$$

$$=\int_0^1 \overline{g(t)\overline{f(t)}}\,dt$$

$$=\int_0^1 f(t)\overline{g(t)}\,dt=< f, g >$$

Thus $<\overline{g, f}>=< f, g>$

iii) Non negativity:

$$< f, f >=\int_0^1 f(t)\overline{f(t)}\,dt$$

$$=\int_0^1 |f(t)|^2\,dt \geq 0$$

Also $< f, f >= 0 \Leftrightarrow \int_0^1 |f(t)|^2\,dt = 0$

$$\Leftrightarrow f(t) = 0\{\forall t \in [0,1]\}$$

$$\Leftrightarrow f = 0$$

As the required conditions are satisfied V is an inner product space.

W.E. 3:

For $f(x), g(x) \in P(R)$, the vector space of polynomials over the field R, defined on $[0,1]$

if $< f(x), g(x) >=\int_0^1 f'(t).g(t)\,dt$ then prove that it is not an inner product.

Solution: Take $f(x) = x; g(x) = x^2$ on $[0,1]$

then $f'(x) = 1$ $\qquad g'(x) = 2x$

So $< f,g >= \int_0^1 1(t^2)\,\mathrm{dt} = \left[\frac{t^3}{3}\right]^1 = \frac{1}{3}$

$< g,f >= \int_0^1 2t(t)\,\mathrm{dt} = 2\left[\frac{t^3}{3}\right]_0^1 = 2\left(\frac{1}{3}\right) = \frac{2}{3}$

Hence $< f,g > \neq < g,f >$

So as the conjugate symmentry is not satisfied. Hence it is not an inner product.

13.4.4. Some Key points in Matrices:

1) Definition: i) Let $A \in M_{m\times n}(F)$. We define the conjugate transpose or Adjoint of A to be the n x m matrix denoted by A * and is defined as if $A = \left[a_{ij}\right]_{m\times n}$ then $A^* = \left[b_{ij}\right]_{n\times m}$ where $b_{ij} = \overline{a_{ji}}$ i.e. $\overline{(A^T)} = (\overline{A})^T$

For example Let $A = \begin{bmatrix} i & 1+2i \\ 2 & 3+4i \end{bmatrix}$ then

$A^* = \begin{bmatrix} -i & 2 \\ 1-2i & 3-4i \end{bmatrix}$

Note: If A has real entries then A * is the transpose of A. i.e. $A^* = A^T$.

ii) Trace of a matrix : Let A be a square matrix of order n. The sum of all the elements of A lying along the principal diagonal is the called the trace of A. We write trace of A as tr (A).

Thus trace of $A = \sum_{i=1}^{n} a_{ii}$

Ex : If $A = \begin{bmatrix} 1 & 3 & 4 \\ 2 & -1 & 3 \\ 2 & 1 & 2 \end{bmatrix}$ then tr $A = (1) + (-1) + 2 = 2$

ii) $tr(A+B) = trA + trB$

iii) $tr(\lambda A) = \lambda trA$ where $\lambda \in C$

iv) $tr(AB) = tr(BA)$

v) $trA = trA^T$

W.E. 4:

Let $V = M_{m\times n}(F)$. Define $< A,B >= tr(B*A)$ for all $A,B \in V$, then show that V is an inner product space.

Solution: We are given $V = M_{m\times n}(F)$ is a vector space and $< A,B >= tr(B*A)$ .... (1)

with this definition we will show that $V(F)$ is a inner product space.

i) Linearity: Let $A,B \in V$. Let $a,b \in F$ then by

1) $< aA+bB, C >= tr\left(C*(aA+bB)\right)$

$$= tr\left(C*(aA)+C*(bB)\right)$$

$$= tr\left(a(C*A)+b(C*B)\right)$$

$$= tr\left(a(C*A)\right)+tr\left(b(C*B)\right)$$

$$= tr(C*A)+b. \ \left(tr(C*B)\right)$$

$$= a < A,C > +b < B,C >$$

Hence the condition of linearity holds good.

ii) Non-negativity :

if $A = \left[a_{ij}\right]_{n\times n}$ then $A* = \left[b_{ij}\right]_{n\times n}$ where $b_{ij} = \overline{a_{ji}}$

$A*A = \left[c_{ij}\right]_{n\times n}$ where $c_{ij} = \sum_{k=1}^{n} b_{ik}a_{kj} = \sum_{k=1}^{n} a_{ki}a_{kj}$

$$< A,A >= tr(A*A) = c_{11}+c_{22}+...+c_{nn} = \sum_{i=1}^{n} c_{ii} = \sum_{i=1}^{n}\sum_{i=1}^{n} b_{ik}a_{ki}$$

$$= \sum_{i=1}^{n}\sum_{i=1}^{n} \overline{a_{ki}}a_{ki} = \sum_{i=1}^{n}\sum_{k=1}^{n} \left|a_{ki}\right|^2 \qquad \text{........... (1)}$$

Now if $A \neq O$ then $a_{ki} \neq 0$ for some K and i So $< A,A >$ > 0

If $A = O$ (null matrix) then $a_{ij} = 0$ for all i, j.

So $a_{ki} = 0$ for all k, i

So $$\sum_{i=1}^{n} |a_{ki}|^2 = 0$$

$\Rightarrow$ trace of $(A^* A) = 0$

$\Rightarrow < A, A >= 0$ .... (2)

Thus if $A = O$ (null matrix) $< A, A >= 0$.

Thus from (1) and (2) $< A, A > \geq 0$

Hence the condition of non negativity is satisfied.

Conjugate Symmetry:

Let $A, B \in V$

Then $< A, B >= traceof (B^* A)$ by given data.

$$< \overline{A, B} >= \overline{tr(B^* A)} = tr(\overline{B^* A})$$

$$= tr\left(\overline{B^*}\overline{(A^T)^T}\right) \because (A^T)^T = A$$

$$= tr\left(\overline{\left(\overline{B}\right)^T}\, \overline{(A^T)}^T\right)$$

$$= tr\left(B^T\left(\left(\overline{A}\right)^T\right)^T\right)$$

$$= tr\left(\left(\left(\overline{A}\right)^T B\right)^T\right) \because tr(A) = tr(A^T)$$

$$= tr(A^* B)$$

$$=< B, A >$$

So $=< \overline{A, B} >=< B, A >$

As all the three required conditions are satisfied,

$V = M_{n\times n}(F)$ is an inner product space.

13.4.5 Frobenius Inner Product:

Definition: The inner product on $V = M_{n\times n}(F)$ defined by $< A, B >= tr(B^* A)$ for all $A, B \in V$ is called the Frobenius inner product and then $V = M_{n\times n}(F)$ is an inner product space with the inner product defined above.

W.E. 5 :

Provide the reason why the product $<(a,b),(c,d)>= ac - bd$ on R^2 is not an inner product on the given vector space.

Solution: Let $(a,b) = (3,4) \in R^2$

then by given data

$$<(a,b),(a,b)>=<(3,4),(3,4)>$$

$$= (3)(3) - (4)(4) = 9 - 16 = -7 \ngeq 0$$

Hence the condition of non negativity is not satisfied. Hence $<(a,b),(c,d)>= ac - bd$ is not an inner product.

W.E. 6:

Provide the reason why the product $< f(x), g(x) >= \int_0^1 f'(t) g(t)\,\mathrm{dt}$ on $P(R)$ where ' denotes differentiation, is not an inner product on the given vector space.

Solution: Take $f(x) = x \quad g(x) = x^2$ on $[0,1]$. Then

$$< f, g >= \int_0^1 1.t^2\,\mathrm{dt} = \left[\frac{t^3}{3}\right]_0^1 = \left(\frac{1}{3} - 0\right) = \frac{1}{3}$$

$$< g, f >= \int_0^1 2t.t\,\mathrm{dt} = 2\int_0^1 t^2\,\mathrm{dt} = 2.\left[\frac{t^3}{3}\right]_1^0 = \frac{2}{3}$$

Hence $< f, g > \neq < g, f >$

Hence the conjugate symmatry does not hold. So the given product is not an inner product.

13.4.6 Theorem:

Let V be an inner product space. Then for $u, v, w \in V$ and $a, b, c \in F$ the following statements are true.

i) $< u, v+w >=< u, w > + < u, w >$

ii) $< u, cv >= \bar{c} < u, v >$

iii) $< u, O >=< O, u >= 0$

iv) $< au - bv, w >= a < u, w > -b < v, w >$

v) $< u, av + bw >= \bar{a} < u, v > +\bar{b} < v, w >$

vi) If $< u, v >=< u, w >$ for all $u \in V$ then $v = w$.

Proof:

i) To show $< u, v+w >=< u, v > + < u, w >$

By definition $< u, v+w >= \overline{< v+w, u >}$

$$= \overline{< v, u > + < w, u >}$$

$$= \overline{< u, v >} + \overline{< w, u >}$$

$$= < u, v > + < u, w >$$

Thus $< u, v+w >=< u, v > + < u, w >$

ii) To show $< u, cv >= \bar{c} < u, v >$

By Definition : $< u, cv >= \overline{< cv, u >}$ (by conjugate symmetry)

$= \overline{c < v, u >}$ by linearity

$$= \bar{c} < \overline{v, u} >$$

$$= \bar{c} < u, v >$$

Thus $< u, cv >= \bar{c} < u, v >$

iii) To show $< u, O >= \overline{< O, u, >} = 0$

Now $< u, O >=< u, 0(O) >$

$= 0 < u, O >= 0$ .... (1)

$\because 0$ is real number its conjugate is itself.

So $<u,O>=0$

IIIly $<0,u>=<0(O),u>$

$=0<O,u>$

$=0$ .... (2)

From (1) and (2)

$<u,O>=<O,u>=0$

iv) To show that $<au=bv,w>=a<u,w>-b<v,w>$

Here $<au-bv,w>=<au+(-b)v,w>$

$=a<u,w>+(-b)<v,w>$

$=a<u,w>-b<v,w>$ by linearity.

Thus $<au-bv,w>=a<u,w>-b<v,w>$

v) To show that $<u,av+bw>=\bar{a}<u,v>+\bar{b}<v,w>$

Solution: $<u,av+bw>=\overline{<av+bw,u>}$ by conjugate symmetry.

$=\overline{a<v,u>+b<w,u>}$

$=\bar{a}<\overline{v,u}>+\bar{b}<\overline{w,u}>$

$=\bar{a}<u,v>+\bar{b}<u,w>$ by conjugate sym metry.

Thus $<u,av+bw>=\bar{a}<u,v>+\bar{b}<u,w>$

Corollary: If a, b are real numbers then

i) $<u,av+bw>=a<u,v>+b<u,w>$

ii) $<u,av-bw>=a<u,v>-b<u,w>$ since if x is real number then $\bar{x}=x$.

vi) If $<u,v>=<u,w>$ for all $u\in V$, then to show $v=w$.

Solution: given $<u,v>=<u,w>$ for all $u\in V$

$\Rightarrow<u,v>-<u,w>=0$ for all $u\in V$

$\Rightarrow < u, v-w > = 0$ for all $u \in V$

$\Rightarrow < v-w, v-w > = 0$ choosing $u = v-w$

$\Rightarrow v-w = 0$

$\Rightarrow v = w$

Thus if $< u,v > = < u,w >$ for all $u \in V$, then $v = w$

Remark: From (ii) and (v) of the above theorem, the reader should observe that the inner product is conjugate linear in the second part.

Note: i) $< -u,-v > = (-1) < u,-v > = (-1)\overline{(-1)} < u,v >$

$= (-1)(-1) < u,v > = < u,v >$

ii) $< u, av-bw > = \overline{a} < u,v > - \overline{b} < u,w >$

iii) $< u+v, u+v > = < u,u > + < u,v > + < v,u > + < v,v >$

iv) If $u_1, u_2, ...u_n, v \in V$ and $a_1, a_2, ...a_n \in F$ then $< a_1u_1 + a_2u_2 + ... + a_nu_n, v >$

$= a_1 < u_1, v > + a_2 < u_2, v > + ... + a_n < u_n, v >$

i.e. $< \sum_{i=1}^{n} a_i u_i, v > = \sum_{i=1}^{n} a_i < u_i, v >$

Also $< v, \sum_{i=1}^{n} a_i u_i > = \sum_{i=1}^{n} \overline{a_i} < v, u_i >$

v) $< au, bv > = a\overline{b} < u,v >$ for all $u, v \in V$ and $a, b \in F$.

13.5 Norm or length of a vector in an inner product space:

Consider the vector space $V_3(R)$ with standard inner product defined on it.

If $u = (a_1, a_2, a_3) \in V_3(R)$ then $< u,u > = a_1^2 + a_2^2 + a_3^2$,

Now we know that in the three dimensional Euclidean space $\sqrt{a_1^2 + a_2^2 + a_3^2}$ is the length of the vector $u = (a_1, a_2, a_3)$. Motivated by this fact, we make the following definition.

13.5.1 Definition: Let V be an inner product space if $u \in V$, then the norm or length of the vector u written as $\|u\|$ is defined as the positive square root of $< u,u >$ i.e. $\|u\| = \sqrt{< u,u >}$.

Note: i) If u is a vector in an inner product space $V(F)$ then $< u,u >$ is always non negative and hence $\|u\| = \sqrt{< u,u >}$ is meaning ful and it is non negative.

ii) In the inner product space $V_2(R) = R^2(R)$. If $u = (a,b) \in V_2$ then $\|u\| = \|(a,b)\| = \sqrt{a^2 + b^2} = \sqrt{< u,u >}$

iii) In the inner product space $V_n(C) = C_n$

if $u = (a_1, a_2, ..., a_n)$ then

$$\|u\| = \|(a_1, a_2, ..., a_n)\| = \sqrt{|a_1|^2 + |a_2|^2 + ... + |a_n|^2}$$

$$= \sqrt{\sum_{i=1}^{n} |a_i|^2} = \sqrt{< u,u >}$$

13.5.2 Unit Vector:

Definition: Let V be an inner product space. If $u \in V$ is such that $\|u\| = 1$, then u is called a unit vector. thus in an inner product space a vector is called a unit vector, if its length is one unit.

13.5.3 Theorem: Let $V(F)$ be an inner product space. u is a non zero vector in V. Then show that $\frac{1}{\|u\|}u$ is a unit vector.

Proof: $< \frac{1}{\|u\|}u, \frac{1}{\|u\|}u >= \frac{1}{\|u\|} < u, \frac{1}{\|u\|}u >= \frac{1}{\|u\|} . \frac{1}{\|u\|} < u,u >$

$$= \frac{1}{\|u\|^2} \|u\|^2 = 1$$

$\Rightarrow \left\| \frac{1}{\|u\|}u \right\|^2 = 1 \Rightarrow \left\| \frac{u}{\|u\|} \right\| = 1 \Rightarrow \frac{u}{\|u\|}$ is a unit vector.

Note: If u is a non zero vector in an inner product space $V(F)$, then the unit vector $\frac{u}{\|u\|}$ is called the unit vector corresponding to u. This process of getting a unit vector along u is called normalizing u.

13.5.4 Definition: Normalising:

The process of multiplying a non zero vector in an inner product space by the reciprocal of its length is called normalizing.

Note ii) In the inner product space $R^2, i=(1,0) \quad j=(0,1)$ are unit vectors since the length of each is 1.

iii) In the inner product space $R^3; i=(1,0,0) \quad j=(0,1,0); k=(0,0,1)$ are unit vectors since the length of each is 1.

iv) In the inner product space R^3; if $u=(a_1,a_2,a_3)$; then $\|u\|=\sqrt{a_1^2+a_2^2+a_3^2}$ and the unit vector corresponding to u is

$$\frac{1}{\|u\|}u=\frac{1}{\sqrt{a_1^2+a_2^2+a_3^2}}(a_1,a_2,a_3)$$

$$=\left(\frac{a_1}{\sqrt{a_1^2+a_2^2+a_3^2}},\frac{a_2}{\sqrt{a_1^2+a_2^2+a_3^2}},\frac{a_3}{\sqrt{a_1^2+a_2^2+a_3^2}}\right)$$

W.E.7:

If $u=(1-i,2+3i)$, $v=(2-5i,3-i)$ are two vectors in a complex inner product space then find $<u,v>, \|u\|, \|v\|$.

Solution: $<u,v>=<(1-i,2+3i),(2-5i,3-i)>$

$$=(1-i)\overline{(2-5i)}+(2+3i)\overline{(3-i)}$$

$$=(1-i)(2+5i)+(2+3i)(3+i)$$

$$=(2+3i+5)+(6+11i-3)$$

$$<u,v>=(7+3i)+(3-11i)=10-8i$$

$$\|u\|^2=<u,u>=<(1-i,2+3i),(1-i,2+3i)>$$

$$=(1-i)\overline{(1-i)}(2+3i)\overline{(2+3i)}$$

$$=(1-i)(1+i)+(2+3i)(2-3i)$$

$$\|u\|^2=(1+1)+(4+9)=15$$

So $\|u\| = \sqrt{15}$

iii) $\|v\|^2 = < v, v > = < (2-5i, 3-i), (2-5i, 3-i) >$

$= (2-5i)\overline{(2-5i)} + (3-i)\overline{(3-i)}$

$= (2-5i)(2+5i) + (3-i)(3+i)$

So $\|v\|^2 = < v, v > = (4+25) + (9+1) = 39$

So $\|v\| = \sqrt{39}$

W.E.8:

Find the unit vector corresponding to $(2-i, 3+2i, 2+\sqrt{3}i)$ of $V_3(C)$ with respect to the standard inner product.

Solution: Let $u = (2-i, 3+2i, 2+\sqrt{3}i)$

$\|u\|^2 = < u, u > = < (2-i, 3+2i, 2+\sqrt{3}i), (2-i, 3+2i, 2+\sqrt{3}i) >$

$= (2-i)\overline{(2-i)} + (3+2i)\overline{(3+2i)} + (2+\sqrt{3i})\overline{(2+\sqrt{3}i)}$

$= (2-i)(2+i) + (3+2i)(3-2i) + (2+\sqrt{3i})(2-\sqrt{3}i)$

$= (4+1) + (9+4) + (4+3) = 25$

So $\|u\|^2 = 25$ Hence $\|u\| = \sqrt{25} = 5$

Hence unit vector corresponding to u is $\frac{1}{\|u\|}u$

$= \frac{1}{5}(2-i, 3+2i, 2+\sqrt{3}i)$

W.E. 9:

If $u = (0,3,4); v = \left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right)$ are two vectors in a real inner product space, then find $< u, v >$.

Solution: $<u,v>=<(0,3,4)\,,\left(\frac{1}{\sqrt{2}},0,\frac{1}{\sqrt{2}}\right)>$

$$=<(0,3,4),\frac{1}{\sqrt{2}}(1,0,1)>$$

$$=\frac{1}{\sqrt{2}}<(0,3,4),(1,0,1)>$$

$$=\frac{1}{\sqrt{2}}\left\{0(1)+3(0)+4(1)=\frac{4}{\sqrt{2}}=2\sqrt{2}\right\}$$

13.5.5 Theorem:

Let V be an inner product space over a field F then show that

$$\|cu\|=|c|\|u\| \;\forall\, u\in V,\forall\, c\in F$$

Proof: $\|cu\|^2=<cu,cu>$

$$=c\bar{c}<u,u>$$

$$=|c|^2\|u\|^2$$

Hence $\|cu\|=|c|\|u\|$

13.5.6 Theorem:

Let V be an inner product space over a field F. Then show that $\|u\|=0$ if and only if $u=0$ and in any case $\|u\|\geq 0$.

Proof: $\|u\|=\sqrt{<u,u>}$ by definition of norm.

So $\|u\|^2=<u,u>\Rightarrow\|u\|^2\geq 0$ since by definition $<u,u>\geq 0$

$$\Rightarrow\|u\|\geq 0$$

Also we know by definition of inner product

$<u,u>=0$ if and only if $u=0$

i.e. $\|u\|^2 = 0$ if and only if $u = 0$

i.e. $\|u\| = 0$ if and only if $u = 0$

Thus in an inner product space $\|u\| > 0$

if and only if $u \neq O$.

13.5.7 Theorem:

CAUCHY - SCHWARZ'S INEQUALITY : If u, v are any two vectors in an inner product space $V(F)$ then $|<u,v>| \leq \|u\|\|v\|$.

Proof: Case (i) if $v = O$ then $<u,v> = <u,O> = <u,0O>$

$$O<u,O> = 0$$

So $|<u,v>| = |0| = 0$ .... (1)

and $\|u\|\|v\| = \|u\|\|O\| = \|u\|(0) = 0$ .... (2)

From (1) and (2) $|<u,v>| = \|u\| \; \|v\|$

Hence $|<u,v>| \leq \|u\|\|v\|$ holds good.

Case (ii) Let $v \neq O$ for any C in F $\|u - cv\| \geq 0$

So $0 \leq \|u - cv\|$ .... (3)

Now $\|u - cv\|^2 = <u - cv, u - cv>$

$= <u, u - cv> - c<v, u - cv>$

$= <u,u> - \bar{c}<u,v> - c\{<v,u> - \bar{c}<v,u>\}$

$= <u,u> - \bar{c}<u,v> - c<v,u> + c\bar{c}<v,u>$ .... (4)

In particular set $c = \dfrac{<u,v>}{<v,u>}$

Using the value of c in (4) we get

$$\|u-cv\|^2 = <u,u> - \frac{\overline{<u,v>}}{<v,v>}<u,v> - \frac{<u,v>}{<v,v>}\overline{<u,v>} + \frac{<u,v>}{<v,v>}\frac{\overline{<u,v>}}{<v,v>}<v,v>$$

So $\|u-cv\|^2 = <u,u> - \dfrac{\overline{<u,v>}<u,v>}{<v,v>}$ .... (5) Since $<v,v>$ is real number. $\overline{<v,v>} = <v,v>$.

Thus from (3) we have $0 \le \|u-cv\|$

i.e. $0 \le \|u-cv\|^2$

We get $0 \le <u,u> - \dfrac{\overline{<u,v>}<u,v>}{<v,v>}$ using (5)

$$\Rightarrow 0 \le \|u\|^2 - \frac{|<u,v>|^2}{\|v\|^2}$$

$$\Rightarrow 0 \le \|u\|^2\|v\|^2 - |<u,v>|^2$$

$$\Rightarrow |<u,v>|^2 \le \|u\|^2\|v\|^2$$

$$\Rightarrow |<u,v>| \le \|u\|\|v\|$$

Hence the theorem.

13.5.8 Special Case of Cauchy - Schwaz's inequality :

Cauchy's inequality Theorem:

In the vector space $V_n(C)$ with standard inner product defined on it,

$$\left|\sum_{i=1}^{n} a_i\overline{b_i}\right| \le \left[\sum_{i=1}^{n}|a_i|^2\right]^{1/2}\left[\sum_{i=1}^{n}|b_i|^2\right]^{1/2}$$

Or

If $a_1, a_2, a_3 ... a_n$ and $b_1, b_2, ... b_n$ are complex numbers, then

$$\left|a_1\overline{b_1} + a_2\overline{b_2} + ... + a_n\overline{b_n}\right| \le \left(\sqrt{|a_1|^2 + |a_2|^2 + ... + |a_n|^2}\right) \times \left(\sqrt{|b_1|^2 + |b_2|^2 + ... + |b_n|^2}\right)$$

Proof: Let $u=(a_1,a_2,...a_n)$ and $v=(b_1,b_2,...b_n)$ are any two vectors in the vector space $V_n(C)$ with standard inner product defined on it.

Then $a_1,a_2,...a_n$ and $b_1,b_2,...b_n$ are all complex numbers.

We have $<u,v>=a_1\overline{b_1}+a_2\overline{b_2}+...+a_n\overline{b_n}$

$$|<u,v>|^2=|a_1\overline{b_1}+a_2\overline{b_2}+...+a_n\overline{b_n}|^2$$

Also $\|u\|^2<u,v>=a_1\overline{a_1}+a_2\overline{a_2}+...+a_n\overline{a_n}$

$$=|a_1|^2+|a_2|^2+...+|a_n|^2$$

Similarly $\|v\|^2=<v,v>=|b_1|^2+|b_2|^2+...+|b_n|^2$

By cauchy - Schwarz's inequality

$$|<u,v>|\le\|u\|\|v\|$$

i.e. $|a_1\overline{b_1}+a_2\overline{b_2}+...+a_n\overline{b_n}|\le\left(\sqrt{|a_1|^2+|a_2|^2+...+|a_n|^2}\right)\times\left(\sqrt{|b_1|^2+|b_2|^2+...+|b_n|^2}\right)$

Note : If $u,v\in V_n(R)$ then $\overline{a_i}=a_i;\overline{b_i}=b_i$ so i.e. when $a_1,a_2,...a_n$, $b_1,b_2,...b_n$ are real numbers.

then $|a_1b_1+a_2b_2+...+a_nb_n|\le\left(\sqrt{a_1^2+a_2^2+...+a_n^2}\right)\times\left(\sqrt{b_1^2+b_2^2+...+b_n^2}\right)$

Or

$$|a_1b_1+a_2b_2+...+a_nb_n|\le\left(\sqrt{a_1^2+a_2^2+...+a_n^2}\right)\left(\sqrt{b_1^2+b_2^2+...+b_n^2}\right)$$

W.E. 10: Using Cauchy - Schwarz, inequality, prove that the also vector value of the cosine of an angle can not be greater than 1.

Solution: Let F be the field of real numbers R and $V=F^{(3)}$

Consider the standard inner product on V.

Let $u=(a_1,a_2,a_3);v=(b_1,b_2,b_3)$ be any two non zero vectors in V.

$$O=(0,0,0)$$

Let θ be the angle between the vectors u and v then

$$\cos\theta = \frac{a_1b_1 + a_2b_2 + a_3b_3}{\sqrt{a_1^2 + a_2^2 + a_3^2}.\sqrt{b_1^2 + b_2^2 + b_3^2}}$$

$$= \frac{<u,v>}{\|u\|\|v\|}$$

$$|\cos\theta| = \frac{<u,v>}{\|u\|\|v\|} \le \frac{\|u\|\|v\|}{\|u\|\|v\|}$$

Since by Cauchy Schwarz inequality $|<u,v>| \le \|u\|\|v\|$

$\therefore |\cos\theta| \le 1$

Hence the absolute value of the cosine of an angle can not be greater than 1.

13.5.9 Triangular inequality:

in an inner product space $V(F)$, prove that $\|u+v\| \le \|u\| + \|v\|$.

Proof: By definition of norm

$$\|u+v\|^2 = <u+v, u+v>$$

$$= <u, u+v> + <v, u+v>$$

$$= <u,u> + <u,v> + <v,u> + <v,v>$$

$$= \|u\|^2 + <u,v> + \overline{<u,v>} + \|v\|^2 \text{ since } <v,u> = \overline{<u,v>}$$

$$= \|u\|^2 + 2\,\mathrm{Re} <u,v> + \|v\|^2$$

$$\le \|u\|^2 + 2|<u,v>| + \|v\|^2 \text{ Since } \mathrm{Re}\, z \le |z|.$$

Hence $\|u+v\|^2 \le \|u\|^2 + 2\|u\|\|v\| + \|v\|^2$ since By Cauchy - Schwarz inequality $|<u,v>| \le \|u\|\|v\|$

$$\Rightarrow \|u+v\|^2 \le \left(\|u\| + \|v\|\right)^2$$

$$\Rightarrow \|u+v\| \le \|u\| + \|v\|$$

Geometrical interpretation of triangular inequality:

Consider u, v to be the vectors in an Inner product space $V_3(R)$ with standard inner product defined on it. Let the vectors u, v be represented by the sides AB, BC of triangle ABC. Then evedently we have $u+v = AC$ and $\|u\| = AB, \|v\| = BC$ and $\|u+v\| = AC$.

Then Cauchy - Schwarz's inequality implies that $AC \le AB + BC$.

W.E. 11:

Prove that if V is an inner product space then $|<u,v>| = \|u\|\|v\|$ if and only if one of the vectors u or v is a multiple of the other (or u, v are linearly dependent).

Solution:

Case i) Let $|<u,v>| = \|u\|\|v\|$

if $v = O$ (zero vector) then clearly $|<u,v>| = 0$ and $\|u\|\|v\| = \|u\|\|0\| = 0$

so $|<u,v>| = \|u\| \; \|v\|$ and we can write $v = 0u$ i.e. v is a scalar multiple of u. (i.e. u, v are linearly dependent).

Similarly if $u = O$ then $u = 0v$ i.e. u is a scalar multiple of v (i.e. u, v are linearly dependent)

Let $v \ne O$ and let $c = \dfrac{<u,v>}{\|v\|^2}$

Let $w = u - cv$

$<w,w> = <u-cv, u-cv>$

$= <u,u> - \bar{c} <u,v> - c <v,u> + c\bar{c} <v,v>$ by (4) Cauchy - Schwarz's inequality.

$$= \|u\|^2 - \frac{\overline{<u,v>}}{\|v\|^2} <u,v> - \frac{<u,v>}{\|v\|^2} <v,u> + \frac{<u,v>}{\|v\|^2} \frac{\overline{<u,v>}}{\|v\|^2} \|v\|^2$$

$$= \|u\|^2 - \frac{<u,v><v,u>}{\|v\|^2}$$

$$= \|u\|^2 - \frac{<u,v>\overline{<u,v>}}{\|v\|^2}$$

$$= \|u\|^2 - \frac{|<u,v>|^2}{\|v\|^2}$$

$$= \|u\|^2 - \frac{\|u\|^2\ \|v\|^2}{\|v\|^2}$$ since $= |<u,v>| = \|u\|\ \|v\|$

$$= 0$$

Hence $<w,w>= 0 \Rightarrow w = O$

$$\Rightarrow u - cv = O \Rightarrow u = cv$$ i.e.

u is a scalar multiple of v. Similarlity when $u \neq O$ we can prove v is a scalar multiple of u.

i.e. u, v are linearly dependent.

Converse: If one of the vectors u, v are zero vectors then they are linearly dependent i.e. one can be expressed as a scalar multiple of other and

$|<u,v>| = 0$ and $\|u\|\ \|v\|$=0

So $|<u,v>| = \|u\|\|v\|$

Let us suppose that both u, v are non zero vectors and they are linearly dependent.

So one is a scalar multiple of the other.

Let $u = cv$ for some $c \in F$

$$<u,v>=<cv,v>= c<v,v>= c\|v\|^2$$

$\therefore |<u,v>| = |c|\|v\|^2$ .... (1)

Also $\|u\| = \|cv\| = |c|\|v\|$

Hence $\|u\|\|v\| = |c|\|v\|^2$ .... (2)

From (1) and (2) it follows

$$|<u,v>| = \|u\|\|v\|$$

Hence from the above two cases the theorem follows.

13.5.9 Theorem:

If u, v are vectors of an inner product space $V(F)$, then $\left|\|u\|-\|v\|\right| \le \|u-v\|$

Proof: u, v are vectors in an inner product space $V(F)$.

$\|u\| = \|u-v+v\| \le \|u-v\| + \|v\|$ by triangular inequality.

So $\|u\| - \|v\| \le \|u-v\|$ .... (1)

Again $\|v\| = \|v-u+u\| \le \|v-u\| + \|u\|$ by triangular inequality

$\Rightarrow \|v\| - \|u\| \le \|v-u\|$ .... (2)

From (1) and (2)

$\left|\|u\| - \|v\|\right| \le \|u-v\|$

13.5.10 Parallelogram law on an inner product space :

If u, v are any two vectors in an inner product space $V(F)$ then show that

$$\|u+v\|^2 + \|u-v\|^2 = 2\|u\|^2 + 2\|v\|^2$$

Proof: u, v are any two vectors in an inner product space $V(F)$.

So $\|u+v\|^2 = <u+v, u+v>$ by definition of norm.

$$= <u, u+v><v, u+v>$$

$$= <u,u> + <u,v> + <v,u> + <v,v>$$

i.e. $\|u+v\|^2 = \|u\|^2 + <u,v> + <v,u> + \|v\|^2$ .... (1)

Also $\|u+v\|^2 = <u-v, u-v>$

$$= <u, u-v> - <v, u-v>$$

$$= <u,u> - <u,v> - \{<v,u> - <v,v>\}$$

So $\|u-v\|^2 = \|u\|^2 - <u,v> - <v,u> + \|v\|^2$ .... (2)

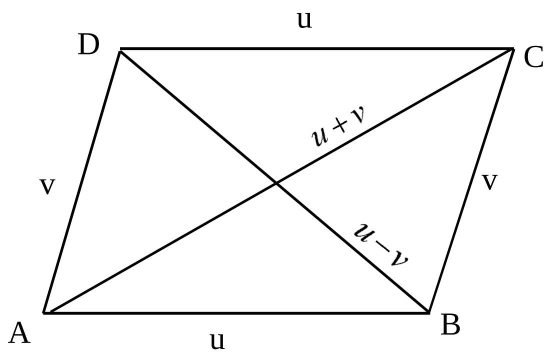
Adding (1) and (2) we get

$$\|u+v\|^2 + \|u-v\|^2 = 2\|u\|^2 + \|v\|^2$$

Hence the theorem.

Geometrical inter pretation of parallelogram Law:

Let u and v be two vectors in the vector space $V_3(R)$ with standard inner product defined on it. Suppose the vector u is represented by the side AB, and the vector v is represented by the side BC of a parallelogram ABCD then th vectors $u+v, u-v$, represents the diagonals AC and DB of the parallelogram.



From the theorem of parallelogram law $\|u+v\|^2 + \|u-v\|^2 = 2\left(\|u\|^2 + \|v\|^2\right)$

$$\Rightarrow AC^2 + DB^2 = 2(AB^2 + BC^2)$$

$$= AB^2 + BC^2 + CD^2 + DA^2$$

$\Rightarrow$ The sum of the squares on the diagonals of a parallelogram is equal to the sum of the squares on the four sides.

13.5.11 Theorem:

If u, and v are two vectors in an innter product space $V(F)$ such that $\|u+v\| = \|u\| + \|v\|$ then the vectors are linearly dependent.

Proof: u, v are two vectors in an inner product space $V(F)$ such that $\|u+v\| = \|u\| + \|v\|$

$$\Rightarrow \|u+v\|^2 = \left(\|u\| + \|v\|\right)^2$$

$\Rightarrow < u+v, u+v >= \|u\|^2 + \|v\|^2 + 2\|u\|\|v\|$

$\Rightarrow < u,u > + < u,v > + < v,u > + < v,v >= \|u\|^2 + \|v\|^2 + 2\|u\|\|v\|$

$\Rightarrow \|u\|^2 + < u,v > + \overline{< u,v >} + \|v\|^2 = \|u\|^2 + \|v\|^2 + 2\|u\|\|v\|$

$\Rightarrow 2\,\mathrm{Re} < u,v >= 2\|u\|\ \|v\|$

i.e. $R\mathrm{e} < u,v >= \|u\|\|v\|$ .... (1)

We know $\mathrm{Re}\, z \leq |z|$

So $\mathrm{Re}(u,v) \leq |< u,v >|$

$\Rightarrow \|u\|\|v\| \leq |< u,v >| \leq \|u\|\|v\|$ by (1) and Cauchy - Schawarz's inequality.

$\Rightarrow |< u,v >| = \|u\|\|v\|$

Hence u,v are linearly independent.

Note: The converse of the above theorem need not be true.

For example consider the inner product space $V_3(R)$ with standard inner product defined on it.

Let $u = (-1,0,1), v = (3,0,-3) \in V_3(R)$

then $v = -3u$. Hence u and v are linearly dependent.

$$\|u\| = \sqrt{(-1)^2 + 0^2 + (1)^2} = \sqrt{2}, \|v\| = \sqrt{(3)^2 + 0^2 + (-3)^2}$$

$$= \sqrt{18} = 3\sqrt{2}$$

$\|u\| + \|v\| = \sqrt{2} + 3\sqrt{2} = 4\sqrt{2}$ .... (1)

$(u+v) = (-1,0,1) + (3,0,-3) = (2,0,-2)$

$\|u+v\| = \sqrt{4+0+4} = \sqrt{8} = 2\sqrt{2}$ .... (2)

From (1) and (2) $\|u+v\| \neq \|u\| + \|v\|$.

13.5.12 If u, v are vectors in an inner product space $V(F)$ then show that

$\text{Re} < u, v > = \frac{1}{4}\|u+v\|^2 \frac{-1}{4}\|u-v\|^2$ and if $F = \mathbb{R}$, then show that $< u, v > = \frac{1}{4}\|u+v\|^2 - \frac{1}{4}\|u-v\|^2$

Proof: u, v are two vectors in an inner product space $V(F)$,

We have $\|u+v\|^2 = < u+v, u+v >$

$= < u, u+v > + < u, u+v >$

$= < u, u > + < u, v > + < v, u > + < v, v >$

$= \|u\|^2 + < u, v > + \overline{< u, v >} + + \|v\|^2$

$= \|u\|^2 + 2\,\text{Re} < u, v > + \|v\|^2$ .... (1)

Also $\|u-v\|^2 = < u-v, u-v >$

$= < u, u-v > - < v, u-v >$

$= < u, u > - < u, v > - \{< v, u > - < v, v >\}$

So $\|u-v\|^2 = \|u\|^2 - < u, v > - \overline{< u, v >} + \|v\|^2$

$= \|u\|^2 - 2\,\text{Re} < u, v > + \|v\|^2$ .... (2)

Subtracting (2) from (1) we get

$\|u+v\|^2 - \|u-v\|^2 = 4\,\text{Re} < u, v >$

So $\text{Re} < u, v > = \frac{1}{4}\|u+v\|^2 - \frac{1}{4}\|u-v\|^2$ .... (3)

If $F = \mathbb{R}$, then $\text{Re} < u, v > = < u, v >$

So (3) becomes

$< u, v > = \frac{1}{4}\|u+v\|^2 - \frac{1}{4}\|u-v\|^2$

13.5.13 Theorem:

If u and v are vectors in a unitary space then

$$4 < u, v > = \|u+v\|^2 - \|u-v\|^2 + i\|u+iv\|^2 - i\|u-iv\|^2$$

Proof: u and v are vectors in a unitary space $V(F)$

$$\|u+v\|^2 = < u, v >, < u, v >$$

$$= < u, u > + < u, v > + < v, u > + < v, v >$$

i.e. $\|u+v\|^2 = \|u\|^2 + < u, v > + < v, u > + \|v\|^2$

$$\|u-v\|^2 = \|u\|^2 - < u, v > - < v, u > + \|v\|^2$$

So $\|u+v\|^2 - \|u-v\|^2 = 2 < u, v > + 2 < v, u >$ .... (1)

$$\|u+iv\|^2 = < u+iv, u+iv >$$

$$= < u, u+iv > + i < v, u+iv >$$

$$= < u, u > + \bar{i} < u, v > + i\{< v, u > + \bar{i} < v, v >\}$$

$$= < u, u > - i < u, v > + i\{< v, u > - i < v, v >\}$$

Since $\bar{i} = -i$

$$= \|u\|^2 - i < u, v > + i < v, u + \|v\|^2$$

So $i\|u+iv\|^2 = i\|u\|^2 + < u, v > - < v, u > + i\|v\|^2$ .... (2)

Also $\|u-iv\|^2 = < u-iv, u-iv >$

$$= < u, u-iv > - i < v, u-iv >$$

$$= < u, u > - \bar{i} < u, v > - i\{< v, u > - \bar{i} < u, v >\}$$

So $\|u-iv\|^2 = \|u\|^2 + i < u, v > - i < v, u > + \|v\|^2$ since $\bar{i} = -i$

So $-i\|u-iv\|^2 = -i\|u\|^2 + < u, v > - < v, u > - i\|v\|^2$ .... (3)

Adding (2) and (3) we get

$$i\|u+iv\|^2 - i\|u-iv\|^2 = 2<u,v> -2<v,u> \text{ .... (4)}$$

Adding (1) and (4) we get

$$4<u,v> = \|u+v\|^2 - \|u-v\|^2 + i\|u+iv\|^2 - i\|u-iv\|^2$$

13.5.14 Theorem:

If u and v are vectors in a unitary space then prove that $<u,v> = \operatorname{Re}<u,v> + i\operatorname{Re}<u,iv>$.

Proof:

If $z = x+iy$ then $y = \operatorname{Im} z$

$$= \operatorname{Re}\{-i(x+iy)\}$$

i.e. $y = \operatorname{Re}(-iz)$

So by this $\operatorname{Im}<u,v> \operatorname{Re}\{-i<u,v>\}$

$= \operatorname{Re}\{<u,iv>\}$ Since $<u,iv> = \bar{i}<u,v>$

$= -i<u,v>$

Hence $<u,v> = \operatorname{Re}<u,v> + i\operatorname{Re}<u,iv>$. Since Z = real $Z + i\operatorname{Im} Z$

13.6 Norm of a Vector in a vector space:

Definition: Let V be a vector space over F, where F is either R or C. Regardless of whether V is or is not an inner product space, we define a norm $\|.\|$ as a real valued function on V satisfying the following three conditions for $u, v \in V$ and $a \in F$.

i) $\|u\| \geq 0$ and $\|u\| = 0$ if and only if $u = 0$

ii) $\|au\| = |a|\|u\|$

iii) $\|u+v\| \leq \|u\| + \|v\|$

A vector space $V(F)$ in which the above three conditions are satisfied is called a normed vector space.

13.6.2 Definition: Normed Vector Space:

Let $V(F)$ be an inner product space in which norms of a vector $u \in V$ is defined as $\|u\| = \sqrt{<u,u>}$. The inner product space with this definition of norm is called a normed vector space if the following three conditions are satisfied.

i) $\|u\| \geq 0$ and $u = 0$ if and only if $u = O$

ii) $\|au\| = |a|\|u\|$ and

iii) $\|u+v\| \leq \|u\| + \|v\| \quad \forall u, v \in V$ and $a \in F$

13.6.3 Theorem: Every inner product space is a normed vector space.

Proof: As the three conditions required for a normed vector space are true in every inner product space, it follows that every inner product space is a normed vector space.

13.6.4 Distance in an inner product space:

Definition: Let u and v be two vectors in an inner product space $V(F)$. The distance between the vectors u and v is denoted by d (u,v) and is defined as $d(u,v) = \|u-v\| = \sqrt{<u-v, u-v>}$

Note: $d(u,v)$ is a non negative real number.

Ex: Let $u = (a_1, a_2, a_3) v = (b_1, b_2, b_3)$ be two vectors in the inner product space R^3. Then

$$u - v = (a_1 - b_1, a_2 - b_2, a_3 - b_3)$$

$$d(u,v) = \|u-v\| = \sqrt{(a_1-b_1)^2 + (a_2-b_2)^2 + (a_3-b_3)^2}$$

13.6.5 Theorem: If u, v , w are any three vectors in an inner product space $V(F)$ then prove that

i) $d(u,v) \geq 0$ and $d(u,v) = 0$ iff $u = v$

ii) $d(u,v) = d(v,u)$

iii) $d(u,v) \leq d(u,w) + d(w,v)$

iv) $d(u,v) = d(u+w, v+w)$

Proof: i) To show that $d(u,v) \geq 0$ and $d(u,v) = 0$ iff $u = v$.

By definition $d(u,v) = \|u-v\| \geq 0$ since the norm of a vector is a non negative real number.

and $d(u,v)=0 \Leftrightarrow \|u-v\|=0 \Leftrightarrow \|u-v\|^2=0$

$$\Leftrightarrow < u-v, u-v >= 0$$

$$\Leftrightarrow u-v=0 \Leftrightarrow u=v$$

Thus $d(u,v)=0$ if and only if $u=v$.

ii) To show $d(u,v)=d(v,u)$

Proof: We have by definition

$$d(u,v)=\|u-v\|$$

$$=\|(-1)(v-u)\|$$

$$=|(-1)|\|u-v\|$$

$$=1\|v-u\|$$

$$=\|v-u\|=d(v,u)$$

So $d(u,v)=d(v,u)$

iii) To show that $d(u,v)\le d(u,w)+d(w,v)$

Proof: $d(u,v)=\|u-v\|$

$$=\|(u-w)+(w-v)\|$$

$\le \|(u-w)\|+\|w-v\|$ by triangle inequality.

$$\le d(u,w)+d(v,w)$$

So $d(u,v)\le d(u,w)+d(v,w)$

iv) To show $d(u,v)=\|u-v\|$

Proof: $d(u,v)=\|u-v\|$

$$=\|(u+w)-(v+w)\|$$

$= d(u+w, v+w)$

Thus $d(u,v) = d(u+w, v+w)$.

13.7 Worked Out Examples:

W.E.12 :

Show that $V_3(R)$ is an inner product space under $<(x_1,x_2,x_3),(y_1,y_2,y_3)>= x_1y_1 + x_2y_2 + x_3y_3$.

Solution : As the given field is the field of real numbers the conjugate symmetry is nothing but symmetry

i) Symmetry:

Let $u = (x_1,x_2,x_3), v = (y_1,y_2,y_3)$

given $<u,v>=<(x_1,x_2,x_3),(y_1,y_2,y_3)>= x_1y_1 + x_2y_2 + x_3y_3$

So $<v,u>=<(y_1,y_2,y_3),(x_1,x_2,x_3) = y_1x_1 + y_2x_2 + y_3x_3$

$= x_1y_1 + x_2y_2 + x_3y_3$

$=<u,v>$

Thus $<v,u>=<u,v>$

ii) Linearity: Let $a,b \in R$ then

$au + bv = a(x_1,x_2,x_3) + b(y_1,y_2,y_3)$

$= (ax_1 + by_1, ax_2 + by_2, ax_3 + by_3)$

Let $w = (z_1,z_2,z_3)$ be any vector in $V_3(R)$.

Then $<au+bv, w>=<(ax_1+by_1, ax_2+by_2, ax_3+by_3),(z_1,z_2,z_3)>$

$= (ax_1+by_1)z_1 + (ax_2+by_2)z_2 + (ax_3+by_3)z_3$

$= (ax_1z_1 + by_1z_1) + (ax_2z_2 + by_2z_2) + (ax_3z_3 + by_3z_3)$

$= a(x_1z_1 + x_2z_2 + x_3z_3) + b(y_1z_1 + y_2z_2 + y_3z_3)$

$= a<u,w> + b<v,w>$

Thus $< au+bv, w >= a < u, w > +b < v, w >$

iii) Non negativity : $< u, u >=< (x_1, x_2, x_3), (x_1, x_2, x_3) >$

$$= x_1(x_1) + x_2(x_2) + x_3(x_3)$$

$$< u, u >= x_1^2 + x_2^2 + x_3^2 \geq 0$$

$$< u, u >= 0 \Leftrightarrow x_1^2 + x_2^2 + x_3^2 = 0 \Leftrightarrow x_1 = 0, x_2 = 0, x_3 = 0$$

So $u = (x_1, x_2, x_3) = (0,0,0) = O$

Hence $< u, u >= 0 \Leftrightarrow u = O$.

As all the three required conditions are satisifed, the given product is an inner product.

So $V_3(R)$ is an inner product space.

W.E. 13:

Which of the following define inner product in $V_2(R)$. Give reasons

i) $< u, v >= x_1y_1 + 2x_1y_2 + 2x_2y_1 + 5x_2y_2$

ii) $< u, v >= 2x_1y_1 + 5x_2y_2$ where $u = (x_1, x_2)$ $v = (y_1, y_2)$

Solution: $u = (x_1, x_2), v = (y_1, y_2)$ are any two vectors in $V_2(R)$.

i) To verify $< u, v >= x_1y_1 + 2x_1y_2 + 2x_2y_1 + 5x_2y_2$ is an inner product or not.

i) Symmetry :

$$< v, u >=< (y_1, y_2), (x_1, x_2) >$$

$$= y_1x_1 + 2y_1x_2 + 2y_2x_1 + 5y_2x_2$$

$$= x_1y_1 + 2x_1y_2 + 2x_2y_1 + 5x_2y_2$$

$$=< u, v >$$

Hence $< v, u >=< u, v >$

ii) Linearity :

Let $a, b \in R$; then

$au+bv=a(x_1,x_2)+b(y_1,y_2)$

$=(ax_1+by_1,ax_2+by_2)$

Let $w=(z_1,z_2)$ be any vector in $V_2(R)$.

Then $<au+bv,w>=<(ax_1+by_1,ax_2+by_2),(z_1,z_2)>$

$=(ax_1+by_1)z_1+2(ax_1+by_1)z_2(ax_2+by_2)z_1+5(ax_2+by_2)z_2$

Thus

$<au+bv,w>=ax_2z_1+by_1z_1+2ax_1z_2+2by_1z_2+2ax_2z_1+2by_2z_1+5ax_2z_2+5by_2z_2$

$=a(x_1z_1+bx_1z_2+2x_2z_1+5x_2z_2)+b(y_1z_1+2y_1z_2+2y_2z_1+5y_2z_2)$

$=a<u,w>+b<v,w>$

Thus $<au+bv,w>=a<u,w>+b<v,w>$

Non Negativity:

$<v,v>=<(x_1,x_2),(x_1,x_2)>$

$=x_1x_1+2x_1x_2+2x_2x_1+5x_2x_2$

$=(x_1+2x_2)^2+x_2^2\geq 0$

and $<u,u>=0 \Leftrightarrow (x_1+2x_2)^2+x_2^2=0$

$\Leftrightarrow x_1+2x_2=0; x_2=0 \Leftrightarrow x_1=0, x_2=0$

$\Leftrightarrow u=(0,0)=O$

Hence $<u,u>=0 \Leftrightarrow u=O$.

As the three required conditions are satisfied $<u,v>=x_1y_1+2x_1y_2+2x_2y_1+5x_2x_2$ is an inner product.

2. To verify $<u,v>=2x_1y_1+5x_2y_2$ is an inner product or not.

given $u=(x_1,x_2)v=(y_1,y_2).<u,v>=2x_1y_1+5x_2y_2$

Symmetry: $< v,u >=< (y_1, y_2)(x_1, x_2) >$

$= 2y_1x_1 + 5y_2x_2$

$= 2x_1y_1 + 5x_2y_2$

$=< u,v >$

Thus $< v,u >=< u,v >$

ii) Linearity :

Let $w = (z_1, z_2)$ be any vector in $V_2(R)$

Let $a,b \in R$

Then $au + bv = a(x_1, x_2) + b(y_1, y_2)$

$= \left((ax_1 + by_1), (ax_2 + by_2)\right)$

Now $< au + bv, w >=< \left((ax_1 + by_1), (ax_2 + by_2)\right), (z_1, z_2) >$

$= 2(ax_1 + by_1)z_1 + 5(ax_2 + by_2)z_2$ by given definition

$= 2ax_1z_1 + 2by_1z_1 + 5ax_2z_2 + 5by_2z_2$

$= a(2x_1z_1 + 5x_2z_2) + b(2y_1z_1 + 5y_2z_2)$

$= a < u,w > +b < v,w >$

Thus $< au + bv, w >= a < u,w > +b < v,w >$

iii) Non Negativity : $< u,u >=< (x_1, x_2), (x_1, x_2) >$

$= 2x_1x_1 + 5x_2x_2$

$= 2x_1^2 + 5x_2^2 \geq 0$

More over $< u,u >= 0 \Leftrightarrow 2x_1^2 + 5x_2^2 = 0$

$\Leftrightarrow x_1 = 0, x_2 = 0 \Leftrightarrow u = (x_1, x_2) = (0,0)$

$\Leftrightarrow u = O$

Thus $<u,u>=0 \Leftrightarrow u=O$

As all the three required conditions are satisfied $<u,v>=2x_1y_1+5x_2y_2$ where $u=(x_1,x_2), v=(y_1,y_2)$ is an inner product.

W.E. 14:

If u,v are vectors in an inner product space $V(F)$ and $a,b \in F$, then prove that $\|au+bv\|^2 = |a|^2\|u\|^2 + a\bar{b}<u,v> + \bar{a}b<v,u> + |b|^2\|v\|^2$

Solution: u,v are vectors in the inner product space $V(F)$.

$$\|au+bv\|^2 = <au+bv, au+bv>$$

$$= a<u, au+bv> + b<v, au+bv>$$

$$= a\left\{\bar{a}<u,u> + \bar{b}<u,v>\right\} + b\left\{\bar{a}<v,u> + \bar{b}<v,v>\right\}$$

$$= a\bar{a}<v,u> + a\bar{b}<u,v> + b\bar{a}<v,u> + b\bar{b}<v,v>$$

$$= |a|^2\|u\|^2 + a\bar{b}<u,v> + \bar{a}b<v,u> + |b|^2\|v\|^2$$

Hence the problem.

W.E. 15 :

Use the Frobenius inner product, compute $\|A\|, \|B\|$ and $<A,B>$ for $A=\begin{bmatrix}1 & 2+i\\ 3 & i\end{bmatrix}$ and $B=\begin{bmatrix}1+i & 0\\ i & -i\end{bmatrix}$ also compute the angle between A and B on $M_{2\times 2}(F)$.

Solution: $B=\begin{bmatrix}1+i & 0\\ i & -i\end{bmatrix}$ Hence $B^*=\begin{bmatrix}1-i & -i\\ 0 & i\end{bmatrix}$

$$B^*A=\begin{bmatrix}1-i & -i\\ 0 & i\end{bmatrix}\begin{bmatrix}1 & 2+i\\ 3 & i\end{bmatrix}=\begin{bmatrix}1(1-i)+3(-i) & (1-i)(2+i)+(-i)i\\ 0(1)+3i & 0(2+i)+i(i)\end{bmatrix}$$

$$=\begin{bmatrix}1-4i & 2-i+1+1\\ 3i & -1\end{bmatrix}$$

So $B^*A=\begin{bmatrix}1-4i & 4-i\\ 3i & -1\end{bmatrix}$

trace of $(B^*A)=1-4i-1=-4i$

As $A=\begin{bmatrix}1 & 2+i\\ 3 & i\end{bmatrix}$; $A^*=\begin{bmatrix}1 & 3\\ 2-i & -i\end{bmatrix}$

$$A^*A=\begin{bmatrix}1 & 3\\ 2-i & -i\end{bmatrix}\begin{bmatrix}1 & 2+i\\ 3 & i\end{bmatrix}=\begin{bmatrix}10 & 2+i+3i\\ 2-i-3i & 4-i^2-1^2\end{bmatrix}$$

$$=\begin{bmatrix}10 & 2+4i\\ 2-4i & 6\end{bmatrix}$$

trace of $A^*A=10+6=16$

$\|A\|^2=<A,A>=$ trace of $A^*A=16$

$\|A\|=\sqrt{<A,A>}=4$

$$B^*B=\begin{bmatrix}1-i & -i\\ 0 & i\end{bmatrix}\begin{bmatrix}1+i & 0\\ i & -i\end{bmatrix}$$

$$=\begin{bmatrix}1-i^2-i^2 & +i^2\\ i^2 & -i^2\end{bmatrix}=\begin{bmatrix}3 & -1\\ -1 & 1\end{bmatrix}$$

trace of $B^*B=3+1=4$

$\|B\|^2=<B,B>=$ trace of $B^*B=4$

$\|B\|=\sqrt{<B,B>}=2$

$<A,B>=$ trace of $B^*A=$ trace of $\begin{bmatrix}1-4i & 4-i\\ 3i & -1\end{bmatrix}$

$$=1-4i-1=-4i$$

If θ is the angle between A and B.

then $\cos\theta = \dfrac{|< A, B >|}{\|A\|\|B\|} = \dfrac{|-4i|}{4(2)} = \dfrac{4}{8} = \dfrac{1}{2}$

$$\cos\theta = \frac{1}{2} \Rightarrow \theta = \frac{\pi}{3}$$

Thus the angle between A and B is $\dfrac{\pi}{3}$

W.E. 16:

Let T be a linear operator an inner product space suppose that $\|T(u)\| = \|u\| \forall u \in V$. Prove that T is one one:

Solution: $V(F)$ is an inner product space.

T is a linear operator on V.

So $T(au+bv) = aT(u)+bT(v)$ for all $u, v \in V$ for all $a, b \in F$.

Show that T is one one:

Let $u, v \in V$ such that

$$T(u) = T(v)$$

$\Rightarrow T(u) - T(v) = O$ (Zero vector in V)

$\Rightarrow T(u-v) = O$ (Since T is a linear operator)

$\Rightarrow \|T(u-v)\| = |O|$

$\Rightarrow \|(u-v)\| = 0$ Since $\|T(u)\| = \|u\|$ for all $u \in V$

$\Rightarrow \|u-v\|^2 = 0$

$\Rightarrow < u-v, u-v >= 0$

$\Rightarrow u - v = O$

$\Rightarrow u = v$

Thus $T(u) = T(v) \Rightarrow u = v \forall u, v \in V$

So T is one one.

W.E. 17:

Let $u=(2,1+i,i)$ and $v=(2-i,2,1+2i)$ be vectors in C^3. Compute $<u,v>, \|u\|, \|v\|$ and $\|u+v\|$. Then verify both the Cauchy - Schwarz inequality and the triangle in equality.

Solution : $<u,v>=<(2,1+i,i),(2-i,2,1+2i)>$

$$=2\overline{(2-i)}+(1+i)2+i\overline{(1+2i)}$$

$$=2(2+i)+2(1+i)+i(1-2i)$$

$$=4+2i+2i+2+i-2i^2$$

$$=4+5i+2+2=8+5i$$

So $<u,v>=8+5i$

$$\|u\|^2=<u,u>=<(2,1+i,i),(2,1+i,i)>$$

$$=2(2)+(1+i)\overline{(1+i)}+i\overline{(i)}$$

$$=2(2)+(1+i)(1-i)+i(-i)$$

$$\|u\|^2=4+1-i^2-i^2=7$$

So $\|u\|=\sqrt{7}$

$$\|v\|^2=<v,v>=<(2-i,2,1+2i),(2-i,2,1+2i)>$$

$$=(2-i)\overline{(2-i)}+2(2)+(1+2i)\overline{(1+2i)}$$

$$=(2-i)(2+i)+4+(1+2i)(1-2i)$$

$$=4-i^2+4+1-4i^2$$

$$\|v\|^2=4+1+4+1+4=14$$

So $\|v\|=\sqrt{14}$

$$u+v=(2,1+i,i)+(2-i,2,1+2i)$$

$$=(2+2-i,1+i+2,i+1+2i)=(4-i,3+i,1+3i)$$

$\|u+v\|^2 = < u+v, u+v > = < (4-i, 3+i, 1+3i), (4-i, 3+i, 1+3i) >$

$= (4-i)\overline{(4-i)} + (3+i)\overline{(3+i)}(1+3i)\overline{(1+3i)}$

$= (4-i)(4+i) + (3+i)(3-i) + (1+3i)(1-3i)$

$= 16 - i^2 + 9 - i^2 + 1 - 9i^2 = 26 + 1 + 1 + 9 = 37$

Hence $\|u+v\| = \sqrt{37}$

To verify Cauchy - Schwarz in equality $|< u, v >| \le \|u\|\|v\|$.

We have shown $< u, v > = 8 + 5i$

$|< u, v >| = \sqrt{64+25} = \sqrt{89}$ ...... (1)

$\|u\| = \sqrt{7}, \|v\| = \sqrt{14}$

$\|u\|\|v\| = \sqrt{7}\sqrt{14} = \sqrt{7 \times 14} = \sqrt{98}$ ..... (2)

but $\sqrt{89} < \sqrt{98}$

So $|< u, v >| < \|u\|\|v\|$

Hence Cauchy - Schwarz's in equality is verified.

To verify triangle inequality $\|u+v\| \le \|u\| + \|v\|$

we have $\|u\| + \|v\| = \sqrt{7} + \sqrt{14}$

But $\sqrt{7} + \sqrt{14} > \sqrt{37}$

So $\|u\| + \|v\| > \|u+v\|$

Hence triangle inequality is verified.

W.E. 18 :

In $C([0,1])$ if $f(t) = t; \quad g(t) = e^t$. Compute $< f, g >, \|f\|, \|g\|,$ and $\|f+g\|$. Then verify both the Cauchy - Schwarz in equality and the triangle in equality. If $< f, g > = \int_0^1 f(t)g(t)dt$.

Solution: Let $V = C([0,1])$ be the inner product space of real valued continuous functions on $[0,1]$ with the inner product f and g defined by

$$< f,g >= \int_0^1 f(t)g(t)\,\mathrm{dt}$$

$$< f,g >= \int_0^1 e^t.t\,\mathrm{dt}$$

$$= \left[t.e^t\right]_0^1 - \int_0^1 1.e^t\,\mathrm{dt}$$

$$1.e^1 - \left[e^t\right]_0^1 = e^1 - \left(e^1 - e^0\right) = e - e + 1 = 1$$

$$< f,g >= 1$$

$$\|f\|^2 =< f,f >=< t,t >= \int_0^1 t.t\,\mathrm{dt} = \int_0^1 t^2\,\mathrm{det} = \left[\frac{t^3}{3}\right]_0^1$$

i.e. $\|f\|^2 = \frac{1}{3} - 0 = \frac{1}{3}$

$\|f\|^2 = \frac{1}{3}$ So $\|f\| = \frac{1}{\sqrt{3}}$

$$\|g\|^2 =< g,g >= \int_0^1 e^t.e^t\,\mathrm{dt} = \int_0^1 e^{2t}\,\mathrm{dt} = \left[\frac{e^{2t}}{2}\right]_0^1$$

i.e. $\|g\|^2 = \frac{1}{2}\left(e^2 - e^0\right) = \frac{1}{2}\left(e^2 - 1\right)$

$$\|g\| = \sqrt{\frac{1}{2}\left(e^2 - 1\right)}$$

$$f + g = t + e^t$$

$$\|(f+g)\|^2 =< f+g, f+g >=< (t+e^t),(t+e^t) >$$

$$= \int_0^1 (t + e^t)^2 dt = \int_0^1 (t^2 + 2te^t + e^{2t})\, \mathrm{dt}$$

$$= \left[\frac{t^3}{3} + \frac{e^{2t}}{2} + 2e^t.(t-1) \right]_0^1$$

$$\|f+g\|^2 = \frac{1}{3} + \frac{1}{2}(e^2 - 1) + 2 = \frac{1}{6}(3e^2 + 11)$$

i.e. $\|f+g\| = \sqrt{\frac{1}{6}(3e^2+11)}$

To verify Cauchy - Schwarz inequality $|<u,v>| \le \|u\|.\|v\|$

We have $|<f,g>| = |1| = 1$ .... (1)

$\|f\|\|g\| = \frac{1}{\sqrt{3}}.\sqrt{\frac{1}{2}(e^2-1)}$ .... (2)

From (1) and (2)

$|<f,g>| < \|f\|\|g\|$

Hence Cauchy - Schwarz in equality is verified.

ii) To verify triangular in equality $\|u + v\| \le \|u\| + \|v\|$

We have $\|f+g\| = \sqrt{\frac{1}{6}(3e^2+11)}$

$$\|f\| + \|g\| = \frac{1}{\sqrt{3}} + \sqrt{\frac{1}{2}(e^2-1)}$$

Hence $\|f+g\| < \|f\| + \|g\|$

So triangle inequality is verified.

W.E. 19 :

In the vector space $V_n(F)$ with standard inner product defined on it show that

$$\left[\sum_{i=1}^{n}|a_i+b_i|^2\right]^{1/2} \le \left[\sum_{i=1}^{n}|a_i|^2\right]^{1/2} + \left[\sum_{i=1}^{n}|b_i|^2\right]^{1/2}$$

Solution: In the vector space $V_n(F)$ with the standard inner product defined on it, let $u=(a_1,a_2,...a_n)$, $v=(b_1,b_2,...b_n)$ then $u+v=(a_1,a_2,...a_n)+(b_1,b_2,...b_n)$

So $u+v=(a_1+b_1,a_2+b_2,...a_n+b_n)$

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

$$\|u\|=\sqrt{|a_1|^2+|a_2|^2+...+|a_n|^2}=\sqrt{\sum_{i=1}^{n}|a_i|^2}$$

Similarly $\|v\|=\sqrt{\sum_{i=1}^{n}|b_i|^2}$

By triangle in equality $\|u+v\|\le\|u\|+\|v\|$

So $\sqrt{\sum_{i=1}^{n}|a_i+b_i|^2} \le \sqrt{\sum_{i=1}^{n}|a_i|^2} + \sqrt{\sum_{i=1}^{n}|b_i|^2}$

13.8 Summary:

In this lesson we discussed about inner products - inner product spaces, norm or length of a vector in an inner produc space - normalising the vectors - Cauchy - Schwarz's in equality triangle in equality - Parallellogram Law - Norm of a vector in a vector space - distance between two vectors.

13.9 Technical Terms:

Inner Product, Inner Product Space, Norm of a vector, Normed Vector space, Distance between vectors.

13.10 Model Questions:

1. Find unit vector corresponding to $(2-i,3+2i,2+\sqrt{3}i)$ of $V_3(C)$ with respect to the standard inner product.

Ans : $\frac{1}{5}(2-3i, 3+2i, 2+\sqrt{3}i)$

2. Which of the following define inner product in $V_2(R)$. Give reasons

i. $<u,v>= x_1y_1 + 2x_1y_2 + 2x_2y_1 + 5x_2y_2$

ii. $<u,v>= 2x_1y_1 + 5x_2y_2$ where $u=(x_1,x_2)$ $v=(y_1,y_2)$

Ans : i) inner product

ii) inner product

3. Show that $V_3(R)$ is an inner product space under the inner product $<(x_1,x_2,x_3),(y_1,y_2,y_3)>= x_1y_1 + x_2y_2 + x_3y_3$.

13.11 Exercises:

1. If $u=(a_1,a_2), v=(b_1,b_2)\in V_2(R)$ then define $<u,v>= a_1b_2 - a_1b_2 + 4a_2b_2$. Show that it is an inner product on $V_2(R)$.

2. Show that $V_3(R)$ is an inner product space under the product defined by $<(x_1,x_2,x_3),(y_1,y_2,y_3)> \; = x_1y_1 + x_2y_2 + x_3y_3$

3. If $u=(a_1,a_2), v=(b_1,b_2)$ then show that $<u,v>= 2a_1\overline{b_1} + a_1\overline{b_2} + a_2\overline{b_1} + a_2\overline{b_2}$ is an inner product on $V_2(R)$.

4. Prove that $<u,v>= a_1 + a_2 + b_1 + b_2$ does not define an inner product in $V_2(R)$.

5. Let $u=(1,3,-4,2), v=(4,-2,2,1), w=(5,-1,-2,6)$ in R^4 then find $<u,w>, <v,w>$ and verify that $<u+v,w>=<u,w>+<v,w>$ and compute $\|u\|, \|v\|, \|w\|$.

Ans: $<u,w>= 22, <v,w>= 24$

$\|u\| = \sqrt{30}, \|v\| = 5, \|w\| = \sqrt{66}$

6. $u=(1,2), v=(-1,1)$ are two vectors in the vector space R^2, with standard inner product. If w is a vector such that $<u,w>= -1, <v,w>= 3$ then find w.

Ans : $w = \left(\frac{-7}{3}, \frac{2}{3}\right)$

7. In the vector space $P(t)$ of all polynomials with inner product $< f, g >= \int_0^1 f(t)g(t)dt$

If $f(t) = t+2, \quad g(t) = 3t-2, \quad h(t) = t^2 - 2t - 3$ then find

i) $< f, g >,$ and $< f, h >,$

ii) $\| f \|$ and $\| g \|$

iii) Normalize f and g.

Ans: $< f, h >= -1, < f, h >= \frac{-37}{4}$

$\| f \| = \frac{\sqrt{57}}{3} \| g \| = 1$,

$\hat{f} =$ unit vector along $f = \frac{3}{\sqrt{57}}(t+2); \hat{g} = g = 3t - 2$

8. Let $M = M_{2\times3}$ with inner product $< A, B >= tr(B^T A)$ and Let $A = \begin{bmatrix} 9 & 8 & 7 \\ 6 & 5 & 4 \end{bmatrix} B = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} C = \begin{bmatrix} 3 & -5 & 2 \\ 1 & 0 & -4 \end{bmatrix}$ then find

i) $< A, B >, < A, C >, < B, C >$ ii) $< 2A + 3B, 4C >$ iii) $\| A \|$ and $\| B \|$

Ans: $< A, B >= 119; < A, C >= -9, < B, C >= -21$

$< 2A + 3B, 4C >= -324$

$\| A \| = \sqrt{271}, \quad \| B \| = \sqrt{91}$

9. Find cos θ where θ is the angle between

i) $u = (1, -3, 2)$ and $v = (2, 1, 5)$ in R^3

ii) $u = (1, 3, -5, 4)$ and $v = (2, -3, 4, 1)$ in R^4

iii) $f(t) = 2t - 1, \quad g(t) = t^2$ where $< f, g >= \int_0^1 f(t)g(t)\,\mathrm{dt}$

iv) $A=\begin{bmatrix} 2 & 1 \\ 3 & -1 \end{bmatrix}; B=\begin{bmatrix} 0 & -1 \\ 2 & 3 \end{bmatrix}$ where $< AB >= tr(B^T A)$

i) $\frac{9}{\sqrt{105}}$ ii) $\frac{-23}{3\sqrt{130}}$ iii) $\frac{\sqrt{15}}{6}$ iv) $\frac{2}{\sqrt{210}}$

10. If $u=(1,-5,3)$ and $v=(4,2,-3)$ are two vectors in R^3 then find the distance between u and v.

Ans : $d(u,v)=\sqrt{94}$

11. In the inner product space C^2 for $u,v \in C^2$ and $A=\begin{bmatrix} 1 & i \\ -i & 2 \end{bmatrix}$ if the inner product $< u,v >= uAv*$ compute $< u,v >$ for $u=[1-i, 2+3i]$, and $v=[2+i, 3-2i]$

Ans : $6+2i \in C$

12. If $u=(1-i, 2+3i)$, $v=(2-5i, 3-i)$ are two vectors in $V(C)=C^2(C)$ with standard inner product find $< u,v >$ and $\|u\|, \|v\|$.

Ans: $10+14i, \sqrt{27}, \sqrt{51}$

13. If u,v are two vectors in a Eucledian space $V(R)$ such that $\|u\| = \|v\|$, then prove that $< u+v, u-v >= 0$.

14. Find the norm of the vector $v=(1,-2,5)$ and also normalise this vector.

Ans : $\|v\| = \sqrt{30}, \quad \hat{v} = \left(\frac{1}{\sqrt{30}}, \frac{-2}{\sqrt{30}}, \frac{5}{\sqrt{30}} \right)$

15. Let $V(R)$ be the vector space of polynomials with inner product determined by $< f,g >= \int_0^1 f(t)g(t)\,dt$ for $f,g \in V$. If $f(x)=x^2+x-4, g(x)=x-1$ for all $x \in [0,1]$ then find $< f,g >, \|f\|, \|g\|$.

Ans : $\frac{7}{4}, \sqrt{\frac{311}{30}}, \frac{1}{\sqrt{3}}$

13.12 Reference Books :

1. Linear Algebra - 4th edition

Stephen. H - Friedberg, Arnold J. insel, Lawrence E - Spence.

2. Topics in Algebra - I.N. Herstein

3. Modern Algebra Vol. II K.S. Narayanan, T.K. Mani Cavachagom pillay

4. A course in Abstract Algebra - Vijay K. Khanna, S.K. Bhambhari

- A. Mallikharjuna Sarma

LESSON - 14

ORTHOGONALIZATION

14.1 Objective of the Lesson:

In geometry perpendicularity is an useful concept. We introduce now a similar concept in inner product spaces. In previous chapters, we have seen the special role of the standard ordered bases for C^n and R^n. The special properties of these bases stem from the fact that the basis vectors form an orthonormal set. Just as bases are the building blocks of vector spaces, bases that are also orthonormal sets are the building blocks of Inner product spaces.

14.2. Structure of the Lesson:

This lesson contains the following items.

14.3 Introduction:

Let us consider the case of vectors in R^2 and we see that how the perpendicularity is considered here. The two vectors u and $v \in R^2$ are perpendicular if and only if the pythogorean relation $\|u\|^2 + \|v\|^2 = \|u+v\|^2$ ......(1) holds. In real inner product spaces this pythagorean relation can be written in a very simple form by using the condition that the angle between the vectors is 90^0 or cosine of the angle between the vectors u and v is zero. Here the condition (1) is equivalent to a very simple condition $< u, v >= 0$. We extend this idea to the vectors of Inner product spaces.

14.4 Orthogonality:

14.4.1 Orthogonality of Two Vectors:

Definition: Two vectors u and v in an inner product space V are said to be orthogonal or perpendicular if $< u, v >= 0$.

14.4.2 Orthogonal Set :

Definition: A subset S of an inner product space is said to be orthogonal if any two distinct vectors in S are orthogonal.

14.4.3 A Vector Orthogonal to a Subset S of V:

Definition: A vector u is said to be orthogonal to a subset S of Inner product space V; if it is orthogonal to each vector in S.

14.4.4 Orthogonal Subspaces :

Two subspaces of an inner product space are called orthogonal if every vector in each is orthogonal to every vector in the other.

Two subspaces W_1 and W_2 of an inner product space $V(F)$ are said to be orthogonal if $< u, v >= 0 \ \forall u \in W_1$ and $\forall v \in W_2$

14.4.5 A Theorem:

Show that orthogonality in an inner product space is symmetric.

Proof: Let $V(F)$ the given inner product space. u, v are two vectors in V such that u is orthogonal to v.

So $< u, v >= \Rightarrow \overline{< u, v >} =$ conjugate of 0.

So $<v,u>=0$. Hence v is orthogonal to u.

Hence orthogonality in an inner product is symmetric.

14.4.5B Theorem:

If u is orthogonal to v, then every scalar multiple of u is orthogonal to v. Where u,v are vectors in an inner product space.

Proof: u, v are vectors in an inner product space $V(F)$

As u, is orthogonal to v; $v;<u,v>=0$ .... (1)

Let k be a scalar belonging to F.

then $<ku,v>=k<u,v>=k(0)=0$ using (1)

$\therefore k\ u$ is orthogonal to v.

Hence every scalar multiple of u is orthogonal to v.

14.4.6 Theorem:

Zero vector in V is orthogonal to every vector in an Inner product space.

Proof: **O** is the zero vector in the inner product space $V(F)$.

Let u be any vector in V. There $<0,u>=<0u,u>=0<u,u>=0$

As u is arbitrary, zero vector is orthogonal to every vector.

14.4.7 Theorem: Show that zero vector is the only vector which is orthogonal to itself in an Inner Product Space.

Proof: Let u be any vector in the given inner product space $V(F)$, which is orthogonal to itself.

So $<u,u>=0 \Rightarrow u=O$ (zero vector) by definition of inner product.

Hence zero vector is the only vector which is orthogonal to itself.

14.4.8 Theorem:

The vectors u, v of a real inner product space $V(F)$ are orthogonal if and only if $\|u+v\|^2=\|u\|^2+\|v\|^2$.

Proof: u, v are two vectors in an inner product space $V(F)$.

Now $\|u+v\|^2=\|u\|^2+\|v\|^2$

$\Leftrightarrow <u+v,u+v>=<u,u>+<v,v>$

$\Leftrightarrow <u,u+v> + <v,u+v> = <u,u> + <v,v>$

$\Leftrightarrow <u,u> + <u,v> + <v,u> + <v,v> = <u,u> + <v,v>$

$\Leftrightarrow <u,v> + <v,u> = 0$

$\Leftrightarrow = <u,v> + <u,v> = 0$ since $<u,v>$ is real $<\overline{u,v}> = <u,v>$

$\Leftrightarrow <u,v> = 0 \Leftrightarrow u,v$ are orthogonal vectors.

14.4.9 Geometrical interpretation:

Let u, v be two vectors in the inner product space $V_3(R)$ with standard inner product defined on it. Let u, v represent the sides $\overline{AB}, \overline{BC}$ of triangle ABC. In the three dimensional Eucledian space. Then $\|u\| = AB.$ $\|v\| = BC$ ..

Also the vector u + v represent the side $\overline{AC}$ of the triangle ABC and $\|u+v\| = AC$. Then from the above theorem $\underline{|ABC} = 90^0$ if and only if $AC^2 = AB^2 + BC^2$ which is pythogorean theorem.

Note The above theorem does not held in complex inner product space.

If $u = (0,i), v = (0,1)$ in $V_2(C)$

There $\|u+v\|^2 = <u+v, u+v>$

$= <u,u> + <u,v> + <v,u> + <v,v>$

$= \|u\|^2 + <u,v> + \overline{<u;v>} + \|v\|^2$

$= \|u\|^2 + \|v\|^2 + 2\operatorname{Re} <u,v>$ .... (1)

since $z + \overline{z} = 2\operatorname{Re} z$

But $<u,v> = <(0,i),(0,1)>$

$= 0(0) + 0(1) + i(0) + c(1)$

$= 0 + i \neq 0$

and $\operatorname{Re} <u,v> = 0$ using this in (1)

We get $\|u+v\|^2 = \|u\|^2 + \|v\|^2$

Which does not imply $<u,v>=0$

14.5 Worked out Examples:

W.E.1 : Find a unit vector orthogonal to $(4,2,3)$ in R^3.

Solution : Let $u=(4,2,3)$ and $v=(a,b,c)$ be orthogonal to u.

Hence $<u,v>=0 \Rightarrow <(4,2,3),(a,b,c)>=0$

$\Rightarrow 4a+2b+3c=0$ .... (1)

Any solution of this equation gives a vector orthogonal to u.

$a=1,b=1,c=-2$ satisfy the equation (1). So $u=(a,b,c)=(1,1,-2)$ is orthogonal to u. $\|v\|=\sqrt{(1)^2+(1)^2+(-2)^2}=\sqrt{6}$

Hence $\hat{v}=\frac{1}{\|v\|}v=\frac{1}{\sqrt{6}}(1,1,-2)=\left(\frac{1}{\sqrt{6}},\frac{1}{\sqrt{6}},\frac{-2}{\sqrt{6}}\right)$ is the unit vector orthogonal to the given vector $u=(4,2,3)$.

W.E. 2 : Find a non-zero vector w which orthogonal to $u=(1,2,1)$ and $v=(2,5,4)$ in R^3.

Solution: Let $w=(x,y,z)$ be the vector which is orthogonal to both u and v.

So $<u,w>=0=<(1,2,1),(x,y,z)>=0$

$\Rightarrow x+2y+12=0$ .... (1)

and $<v,w>=0 \Rightarrow <(2,5,4),(x,y,z)>=0$

$\Rightarrow 2x+5y+4z=0$ .... (2)

(1) x 2 : $2x+4y+2z=0$

Subtracting from (2) $y+2z=0$

So $y=-2z$

Put $z=1, y=-2$

using these values in (1)

$x-4+1=0 \Rightarrow x=3$

Thus $w=(x,y,z)=(3,-2,1)$ is the desired non zero vector orthogonal to both u and v.

W.E. 3 : In a real inner product space if u, v are two vectors such that $\|u\|=\|v\|$ then prove that $u-v,$ and $u+v$ are orthogonal. Interpret the result geometrically.

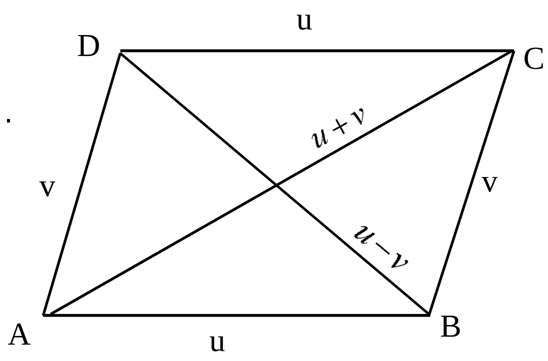
Solution: $<u-v,u+v>=<u,u>+<u,v>-<v,u>-<v,v>$

$$= \|u\|^2+<u-v>-\overline{<u,v>}-\|v\|^2=0$$

Since in the real inner product space $\overline{<u,v>}=<u,v>$ and given $\|u\|=\|v\|$.

So as $<u-v,u+v>=0$, the vectors $u-v,u+v$ are orthogonal.

Geometrical interpretation:



In the 3 dimensional space Let $u=\overline{AB},v=\overline{BC}$.

then $\|u\|=\|v\|\Rightarrow AB=BC$.

In the rhombus $ABCD,\overline{AC}=u+v,\overline{DB}=u-v$

where AC,DB are the diagonals.

$$<u-v,u+v>=0$$

$\Rightarrow u-v,u+v$ are orthogonal

$\Rightarrow$ The diagonals AC and DB are perpendicular.

Hence in a Rhombus the diagonals are perpendicular.

W.E. 4: If u, v are two vectors in a real inner products space and $u+v,u-v$ are orthogonal then $\|u\|=\|v\|$

Solution: $u+v,u-v$ are orthogonal.

$$\Rightarrow <u+v,u-v>=0$$

$$\Rightarrow <u,u>-<u,v>+<v,u>-<v,v>=0$$

$$\Rightarrow \|u\|^2-<u,v>+\overline{<u,v>}-\|v\|^2=0$$

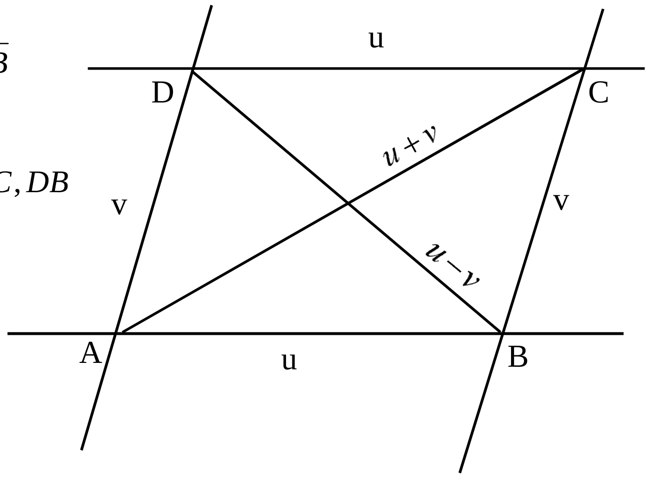
$\Rightarrow \|u\|^2-\|v\|^2=0$ since in a real inner product space $\overline{<u,v>}=<u,v>$

$\Rightarrow \|u\|^2 = \|v\|^2 \Rightarrow \|u\| = \|v\|$

Thus the vectors $u+v, u-v$ are orthogonal $\Rightarrow \|u\| = \|v\|$

Geometrical interpretation:

Let u, v be two vectors in the inner product space $V_3(R)$ with standard inner product defined on it. Let u, v represent the sides AB and BC of a parallelogram $ABCD$. Then $\|u\| = AB, \|v\| = BC$.



$u+v, u-v$ represent the diagonal $\overline{AC}$ and $\overline{DB}$ of the parallelogram.

As $u+v, u+v$ are orthogonal the diagonals AC, DB are perpendicular.

$\|u\| = \|v\| \Rightarrow AB = BC$. Thus if the diagonals of a parallelogram are perpendicular, then the parallelogram is a rhombus.

W.E. 5: Let V be an inner product space and suppose that u and v are orthogonal vectors in an inner product space V,

Prove that $\|u+v\|^2 = \|u\|^2 + \|v\|^2$. Deduce the Pythagorean theorem in R_2.

Solution: As u and v are orthogonal vectors in an inner product space V, $<u,v>=0$ ...... (1)

Now

$$\|u+v\|^2 = <u+v, u+v>$$

$$= <u,u> + <u,v> + <v,u> + <v,v>$$

$$= \|u\|^2 + <u,u> + \overline{<u,v>} + \|v\|^2 \text{ .... (2)}$$

$$= \|u\|^2 + 0 + \text{ conjugate of } 0 + \|v\|^2 \text{ using (1)}$$

So $\|u\|^2 + \|v\|^2 = \|u\|^2 + \|v\|^2$

Deduction of Pythogorem theorem:

If u, v are two orthogonal vectors in a real inner product space, then $\overline{<u,v>} = <u,v>$ .... . (3) using in (2).

$$\|u+v\|^2 = \|u\|^2 + \|v\|^2 + 2<u,v>$$

$= \|u\|^2 + \|v\|^2 + 0$. As u, v are orthogonal vectors $<u,v> = 0$

So $\|u+v\|^2 + \|u\|^2 + \|v\|^2$ .... (4)

In R_2 if $\overline{AB} = u, \overline{BC} = v$ then $\overline{AC} = u+v$ and (4) states $AC^2 = AB^2 + BC^2$ which is pythogorean theorem.

W.E. 6 : If u, v are two orthogonal vectors in an inner product space $V(F)$ and $\|u\| = \|v\| = 1$ then prove that $\|u-v\| = d <u,v> = \sqrt{2}$

Solution: u, v are orthogonal vectors in an inner product space $V(F)$. So $<u,v> = 0$ .... (1)

$$(d(u,v))^2 = \|u-v\|^2$$

$$= <u-v, u-v>$$

$$= <u,u> - <u,v> - <v,u> + <v,v>$$

$= \|u\|^2 - 0 - 0 + \|v\|^2$ since, u, v are orthogonal.

$= (1)^2 + (1)^2$ since given $\|u\| = \|v\| = 1$

$$= 2$$

So $d(u,v) = \|u-v\| = \sqrt{2}$

14.6.1 Theorem:

Show that any orthogonal set of non zero vectors in an inner product space V is linearly independent.

Proof: Let S be an orthogonal set of non zero vectors in an inner product space V.

Let $S_1 = \{u_1, u_2 ..., u_n\}$ be a finite subset of S containing m vectors which are distinct.

Let $\sum_{j=1}^{m} c_j u_j = c_1 u_1 + c_2 u_2 + c_n u_n = O$ .... (1)

We will show that each scalar coefficient is zero. Let u_i be any vector in S_1. i.e. $1 \le i \le m$.

Now consider $< c_1 u_1 + c_2 u_2 + ... + ... + c_{i-1} u_{i-1} + c_i u_i + c_i u_i + c_m u_m, u_i >$

$$= c_1 < u_1, u_i > + c_2 < u_2, u_i > + ... + c_{i-1} < u_{i-1}, u_i > + c_i < u_i, u_i > + c_{i+1} < u_{i+1}, u_i > + ... c_m < u_m, u_i >$$

i.e. $< \sum_{j=1}^{m} c_j u_j, u_i >= c_i < u_i, u_i >$ .... (2)

$\because$ As the vectors are orthogonal $< u_i, u_j >= 0$ for $i \ne j$

But by (1) $\sum_{j=1}^{m} c_j u_j = O$

Using in (2)

$$< O, u_i >= c_i \|u_i\|^2 \Rightarrow c_i^2 \|u_i\|^2 = 0$$

So $c_i \|u_i\| = 0$. But u_i is a non zero vector. $\|u_i\| \ne 0$

Hence $c_i = 0$ for $1 < i \le m$

Thus $c_1 u_1 + c_2 u_2 + ... + c_m u_m = 0$

$$\Rightarrow c_1 = 0, c_2 = 0, ... c_m = 0$$

Hence $S_1 = \{c_1, c_2, ..., c_m\}$ is a linearly independent set.

Thus every finite subset of S is linearly independent. Hence S is linearly independent.

14.6.2 Theorem:

Let $S = \{u_1, u_2, ..., u_m\}$ be an orthogonal set of non zero vectors in an inner product space $V(F)$.

If a vector v in V is in the linear span of S,

Then $v = \sum_{i=1}^{m} \frac{< v, u_i >}{\|u_i\|^2} u_i$

Proof: v is a vector in the inner product space V, which is in the linear span of $S = \{u_1, u_2, ...u_m\}$.

So v can be expressed as a linear combination of the vectors of S. So there exists scalars $c_1, c_2, ...c_m$ in F. Such that $v = c_1u_1 + c_2u_2 + ... + c_mu_m = \sum_{j=1}^{m} c_ju_j$

then for each i where $1 \le i \le m$

$$< v, u_i > = < \sum_{j=1}^{m} c_ju_j, u_i > \text{ .... (1)}$$

$$= \sum_{J=1}^{m} c_j < u_j, u_i >$$ by linear property of inner products.

= $c_i < u_i, u_i >$ on summing up with respect to j, and S is an orthogonal set of non zero vectors. So $< u_j, u_i > = 0$ if $j \ne i$.

So $< v, u_i > = c_i \|u_i\|^2$ .... (2)

As u_i is a non zero vector in S, $\|u_i\| \ne 0$

So by (2) $c_i = \frac{< v, u_i >}{\|u_i\|^2}$

But $v = c_1u_1 + c_2u_2 + ... + c_mu_m$

So $v = \frac{< v, u_1 >}{\|u_1\|^2}u_1 + \frac{< v, u_2 >}{\|u_2\|^2}u_2 + ... + \frac{< v, u_m >}{\|u_m\|^2}u_m$

Hence $v = \sum_{i=1}^{m} \frac{< v, u_i >}{\|u_i\|^2}u_i$

Hence the theorem.

14.7.1 Definition:

Orthonormal set: Let S be a set of unit vectors in an inner product space V, which are mutually orthogonal then S is said to be an orthonormal set.

Or

Let S be a set of vectors in an inner product space V. Then S is said to be orthonormal set if

i) $u \in S \Rightarrow \|u\| = 1$ ie $< u, u >= 1$ and

ii) $u, v \in S$ and $u \neq v \Rightarrow < u, v >= 0$

Or

A finite set $S = \{u_1, u_2, ..., u_m\}$ of an inner product space V is orthonormal if

$< u_i, u_j >= \delta_{ij}$ where δ_{ij} denotes the kronecker delta.

i.e. $\delta_{ij} = \begin{cases} 1 \text{ if } i = j \\ 0 \text{ if } i \neq j \end{cases}$

Note: i) An orthonormal set is an orthogonal set in which every vector is a unit vector i.e. A set consisting of mutually orthogonal unit vectors is called an orthonormal set.

Note ii): An orthonormal set does not contain zero vector.

14.7.2 Worked Out Examples:

W.E. 7: Prove that $S = \left\{ \left(\frac{1}{3}, \frac{-2}{3}, \frac{-2}{3} \right), \left(\frac{2}{3}, \frac{-1}{3}, \frac{2}{3} \right), \left(\frac{2}{3}, \frac{2}{3}, \frac{-1}{3} \right) \right\}$ is an orthonormal set in R^3 with standard inner product.

Solution: Let $u = \left(\frac{1}{3}, \frac{-2}{3}, \frac{-2}{3} \right), v = \left(\frac{2}{3}, \frac{-1}{3}, \frac{2}{3} \right)$, $w = \left(\frac{2}{3}, \frac{2}{3}, \frac{-1}{3} \right)$ are the given vectors of S.

$$\|u\| = \sqrt{\left(\frac{1}{3}\right)^2 + \left(\frac{-2}{3}\right)^2 + \left(\frac{-2}{3}\right)^2} = \sqrt{\frac{1}{9} + \frac{4}{9} + \frac{4}{9}} = 1$$

$$\|v\| = \sqrt{\left(\frac{2}{3}\right)^2 + \left(\frac{-1}{3}\right)^2 + \left(\frac{2}{3}\right)^2} = \sqrt{\frac{4}{9} + \frac{1}{9} + \frac{4}{9}} = \sqrt{1} = 1$$

$$\|w\| = \sqrt{\left(\frac{2}{3}\right)^2 + \left(\frac{2}{3}\right)^2 + \left(\frac{-1}{3}\right)^2} = \sqrt{\frac{4}{9} + \frac{4}{9} + \frac{1}{9}} = 1$$

$$< u, w > = < \left(\frac{1}{3}, \frac{-2}{3}, \frac{-2}{3}\right), \left(\frac{2}{3}, \frac{2}{3}, \frac{-1}{3}\right) > \quad = \frac{1}{3}\left(\frac{2}{3}\right) + \left(\frac{-2}{3}\right)\left(\frac{2}{3}\right) + \left(\frac{-2}{3}\right)\left(\frac{-1}{3}\right)$$

i.e. $< u, w > = \frac{2}{9} - \frac{-4}{9} + \frac{2}{9} = 0 = < w, u >$

$$< u, v > = < \left(\left(\frac{1}{3}, \frac{-2}{3}, \frac{-2}{3}\right), \left(\frac{2}{3}, \frac{-1}{3}, \frac{2}{3}\right)\right) >$$

$$= \frac{1}{3}\left(\frac{2}{3}\right) + \left(\frac{-2}{3}\right)\left(\frac{-1}{3}\right) + \left(\frac{-2}{3}\right)\left(\frac{2}{3}\right) = \frac{2+2-4}{9} = 0$$

i.e. $< u, v > = 0 = < v, u >$

$$< v, w > = < \left(\frac{2}{3}, \frac{-1}{3}, \frac{2}{3}\right), \left(\frac{2}{3}, \frac{2}{3}, \frac{-1}{3}\right) >$$

$$= \left(\frac{2}{3}\right)\left(\frac{2}{3}\right) + \left(\frac{-1}{3}\right)\left(\frac{2}{3}\right) + \left(\frac{2}{3}\right)\left(\frac{-1}{3}\right) = \frac{4-2-2}{9} = 0$$

Thus $< v, w > = 0 = < w, v >$

As the length of each vector in S is unity, the inner product of two different vectors in S is 0, So S is an orthonormal set.

W.E.7b : Consider the usual basis of $E = \{e_1, e_2, e_3\}$ of the Eucledian space R^3 where $e_1 = (1,0,0), e_2 = (0,1,0), e_3 = (0,0,1)$. Show that E is orthonormal.

Solution:

$e_1 = (1,0,0)$ so $\|e_1\| = \sqrt{1^2 + 0^2 + 0^2} = 1$

$e_2 = (0,1,0)$ so $\|e_2\| = \sqrt{0^2 + 1^2 + 0^2} = 1$

$e_3 = (0,0,1)$ so $\|e_3\| = \sqrt{0^2 + 0^2 + 1^2} = 1$

Thus $\|e_1\| = \|e_2\| = \|e_3\| = 1$

More over $< e_1, e_2 > = < (1,0,0), (0,1,0) >$

$= 1(0) + 0(1) + 0(0) = 0$

$\therefore < e_1, e_2 >= 0 =< e_2, e_1 >$

$< e_2, e_3 >=< (0,1,0), (0,0,1) >= 0(0) + 1(0) + 0(1) = 0$

$\therefore < e_2, e_3 >= 0 =< e_3 = e_2 >$

$< e_3, e_1 >=< (0,0,1), (1,0,0) >= 0(1) + 0(0) + 0(0) = 0$

Thus $< e_3, e_1 >= 0 =< e_1, e_3 >$

Thus the length of each vector in E is unity and the inner product of any two different vectors of E is zero. So E is an orthonormal set.

W.E. 8: If $S = \{(1,2,-3,4), (3,4,1,-2), (3,-2,1,1)\}$ is a subset of R^4. Obtain orthonormal set from S. Verify Pythagorean theorem.

Solution: Let $u = (1,2,-3,4), v = (3,4,1,-2); w = (3,-2,1,1)$

Now $< u, v >=< (1,2,-3,4), (3,4,1,-2) >= 1(3) + 2(4) + (-3)(1) + 4(-2)$

$= 3 + 8 - 3 - 8 = 0$

Thus $< u, v >= 0 =< v, u >$

$< v, w >= (3,4,1,-2), (3,-2,1,1) >= 3(3) + 4(-2) + 1(1) + (-2)(1)$

$= 9 - 8 + 1 - 2 = 0$

So $< v, w >= 0 =< w, v >$

$< w, u >=< (3,-2,1,1), (1,2,-3,4) >= 3(1)(-2)2 + 1(-3) + 1(4)$

$= 3 - 4 - 3 + 4 = 0$

Thus $< w, u >= 0 =< u, w >$

Thus the inner product of any two different vectors in S is zero. So S is orthogonal.

We normalise S to obtain an orthonormal set.

$\|u\| = \sqrt{1^2 + (2)^2 + (-3)^2 + 4^2} = \sqrt{1 + 4 + 9 + 16} = \sqrt{30}$

$\|v\| = \sqrt{3^2 + 4^2 + 1^2 + (-2)^2} = \sqrt{9 + 16 + 1 + 4} = \sqrt{30}$

$\|w\| = \sqrt{3^2 + (-2)^2 + 1^2 + 1^2} = \sqrt{9+4+1+1} = \sqrt{15}$

Hence the required orthonormal set of vectors is

$$\left\{\left(\frac{1}{\sqrt{30}}, \frac{2}{\sqrt{30}}, \frac{-3}{\sqrt{30}}, \frac{4}{\sqrt{30}}\right), \left(\frac{3}{\sqrt{30}}, \frac{4}{\sqrt{30}}, \frac{1}{\sqrt{30}}, \frac{-2}{\sqrt{30}}\right), \left(\frac{3}{\sqrt{15}}, \frac{-2}{\sqrt{15}}, \frac{1}{\sqrt{15}}, \frac{1}{\sqrt{15}}\right)\right\}$$

More over $u + v + w = (1,2,-3,4) + (3,4,1,-2) + (3,-2,1,1)$

$$= (7,4,-1,3)$$

$$\|u+v+w\|^2 = (49+16+1+9) = 75$$

$\|u\|^2 + \|v\|^2 + \|w\|^2 = 30 + 30 + 15 = 75$

Hence $\|u+v+w\|^2 = \|u\|^+ + \|v\|^2 + \|w\|^2$

Which verifies the pythogorean theorem for the orthogonal set S.

W.E. 9: Let $V(F)$ be an inner product space. If u is a non zero vector in V then show that $\left\{\frac{u}{\|u\|}\right\}$ is an orthonormal set.

Solution: u is a non zero vector in V. So $\|u\| \neq 0$

Hence $< \frac{u}{\|u\|}, \frac{u}{\|u\|} >= \frac{1}{\|u\|} < u, \frac{u}{\|u\|} >$

$$= \frac{1}{\|u\|} \cdot \frac{1}{\|u\|} < u, u >= \frac{1}{\|u\|^2} \|u\|^2 = 1$$

As $\frac{u}{\|u\|}$ is an unit vector and so $\left\{\frac{u}{\|u\|}\right\}$ is an orthonormal set. it is a subset of V.

Note: Every inner product space has an orthonormal subset.

14.8.1 Theorem: Every orthonormal set of vectors in an inner product space $V(F)$ is linearly independent.

Proof: Let S be an orthonormal set of vectors in an inner product space V. Let $S_1 = \{u_1, u_2, ...u_m\}$ be a finite subset of S containing m vectors.

Let $c_1, c_2, ...c_m$ be scalars belonging to F such that

$$c_1u_1 + c_2u_2 + ... + c_mu_m = O \text{ .... (1) (zero vector)}$$

Now

$$< c_1u_1 + c_2u_2 + ... + c_iu_i + ... + c_mu_m, u_i > = < O, u_i > = 0 \text{ where } 1 \le i \le m.$$

$$\Rightarrow c_1 < u_1, u_i > + c_2 < u_2, u_i > + ... c_{i-1} < u_{i-1}, u_i > + c_i < u_i, u_i > + c_{i+1} < u_{i+1}, u_i > + ... + c_m < u_m, u_i > = 0 \text{ .(2)}$$

As $u_1, u_2 ... u_m$ are orthonormal vectors

$$< u_j, u_i > = \begin{cases} 1 \text{ if } j \neq i \\ 0 \text{ if } j = i \end{cases}$$

using this in the above (2) we get

$$c_i < u_i, u_i > = 0$$

$\Rightarrow c_i = 0$ since $< u_i, u_i > \neq 0$ where $1 \le i \le m$.

So $c_i = 0, c_2 = 0, ...c_m = 0$

Thus

$$c_1u_1 + c_2u_2 + ... c_mu_m = O \Rightarrow c_1 = 0, c_2 = 0, ...c_m = 0$$

Hence $S_1 = \{u_1, u_2, ...u_m\}$ is linearly independent.

As S_1 is arbitarary, every finite sub set of S is linearly independent. Hence S is linearly independent.

Aliter : Let S be an orthonormal set in an inner product space. Let $u \in S$, then $u \neq O$, since $u = O \Rightarrow < u, u > = 0 \neq 1$ is a contradiction.

So S is an orthogonal set of non zero vectors.

As we know every orthogonal set of non-zero vectors in an inner product space V is linearly independent, if follows S is linearly independent.

Hence every orthonormal set of vectors in an inner product space $V(F)$ is linearly independent.

14.8.2 Theorem: Let $S = \{u_1, u_2, ...u_m\}$ be an orthonormal set of vectors in an inner product space $V(F)$. If a vector v is in the linear span of S; then

$$v = \sum_{i=1}^{m} < v, u_i > u_i$$

Proof: v is given to be a vector in the linear span of S. So v can be expressed as a linear combination of the vectors of S. Hence there exists scalars $c_1, c_2, ...c_m$ in F such that

$$v = c_1 u_1 + c_2 u_2 + ... + c_m u_m = \sum_{J=1}^{m} c_j u_j \quad .... (1)$$

We have for each i where $1 \le i \le m$

$$< v, u_i > = < \sum_{J=1}^{m} c_j u_j, u_i >$$

$$= \sum_{J=1}^{m} c_j < u_j, u_i > \text{ by linearty of inner product}$$

$$= c_i < u_i, u_i > \text{ since}$$

$$< u_j, u_i > = \begin{cases} 0 \text{ if } j \neq i \\ 1 \text{ if } j = i \end{cases}$$

$$= c_i(1) = c_i$$

Thus $< v, u_i > = c_i$ where $1 \le i \le m$.

Putting these values of $c_1, c_2, ...c_m$ in (1)

We get $v = \sum_{i=1}^{m} < v, u_i > u_i$

14.8.3 Theorem:

If S is an orthonormal set of vectors of an inner product space $V(F)$ then $v = a_1u_1 + a_2u_2 + ... + a_nu_n$ implies $a_1 =< v, u_i >$ where $u_i \in S, a_i \in F$ for $i = 1, 2, ...n$.

Proof: S is an orthonormal set. So for $u_i, u_j \in S$

$\|u_i\| = 1, \|u_j\| = 1$ and $< u_i, u_j >= 0$ where $i \neq j$

Thus for each $i = 1, 2, ...n$;

$$< v, u_i >=< a_1u_1 + a_2u_2 + ... + a_nu_n, u_i >= a_i < u_i, u_i >$$

$$= a_i \text{ .... (1)}$$

Hence $a_i =< v, u_i >$ and so $v = a_1u_1 + a_2u_2 + ... + a_nu_n = \sum_{c=1}^{n} a_iu_i$

$$v = \sum_{i=1}^{n} < v, u_i > u_i \text{ using (1)}$$

14.8.4 Theorem:

If $S = \{u_1, u_2, ...u_m\}$ is an orthonormal set in an inner product space $V(F)$ and $u \in V$ then $w = v - \sum_{i=1}^{m} < v, u_i > u_i$ is orthogonal to each of $u_1, u_2 ... u_m$.

Proof:

For $j = 1, 2, ..., m$

$$< w, u_j >=< v - \sum_{i=1}^{m} < v, u_i > u_i, u_j >$$

$$=< v, u_j > - \sum_{i=1}^{m} < v, u_i >< u_i, u_j >$$

$$=< v, u_j > - \{< v, u_1 >< u_1, u_j > + < v, u_2 >< u_2, u_j > + ... + < v, u_j >< u_j, u_j > + ... + < v, u_m >< u_m, u_j >\}$$

$$=< v, u_j > \{< v, u_1 > 0 + < v, u_2 > 0 + ... < v, u_j > (1) + < v, u_{j+1} > (0) + ... + < v, u_m > 0\}$$

$= < v, u_j > - < v, u_j > = 0$

$\therefore < w, u_j > = 0 \Rightarrow w$ is orthogonal to each of $u_1, u_2 ... u_m$.

$\Rightarrow v - \sum_{i=1}^{m} < v, u_i > u_i$ is orthogonal to each of $u_1, u_2 ... u_m$.

14.8.5 Corollary: If $S = \{u_1, u_2 ... u_m\}$ is an orthonormal set in an inner product space $V(F)$ and $u \in V$, then $w = v - \sum_{i=1}^{m} < v, u_i > u_i$ is orthogonal to each vector of $L(S)$.

Proof: From the above the orem the vector w is orthogonal to each of $u_1, u_2 ... u_m$.

So $< w, u_i > = 0 = < u_i, w >$ for each i where $1 \le i \le m$ .... (1)

Let $u \in L(S)$. Then there exists scalars $c_1, c_2, ... c_m$ in F such that

$$u = c_1 u_1 + c_2 u_2 + ... + c_m u_m$$

Now $< u, w > = < c_1 u_1 + c_2 u_2 + ... + c_m u_m, w >$

$$= c_1 < u_1, w > + c_2 < u_2, w > + ... + c_m < u_m, w >$$

$= c_1(0) + c_2(0) + ... + c_m(0)$ using (1)

$\therefore u$ is orthogonal to w. So w is orthogonal to u. Hence w is orthogonal to every vector of $L(S)$.

14.9 Worked Out Examples:

W.E.10: If $S = \{u_1, u_2, ... u_n\}$ is an orthogonal set of an inner product space $V(F)$ then prove that $\{a_1 u_1, a_2 u_2, ... a_n u_n\}$ is also an orthogonal set for any choice of non zero scalars $a_1, a_2, ... a_n \in F$.

Solution: $S = \{u_1, u_2, ... u_n\}$ is an orthogonal set

$\Rightarrow < u_i, u_j > = 0$ ..... (1) $\forall u_i, u_j \in S$ and $i \ne J$.

Consider $a_i u_i, a_j u_j \in S_1 = \{a_1 u_1, a_2 u_2, ... a_n u_n\}$

Then $< a_i u_i, a_j u_j > = a_i \overline{a_j} < u_i, u_j >$

$= a_i \overline{a_j}(0)$ using (1)

$\Rightarrow < a_i u_i, a_j u_j >= 0 \forall a_i u_i, a_j u_j \in S_1$ and $i \neq J$

So $S_1 = \{a_1 u_1, a_2 u_2, ... a_n u_n\}$ is also an orthogonal set.

W.E.11: If $S = \left\{ \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0 \right), \left(\frac{1}{\sqrt{3}}, \frac{-1}{\sqrt{3}}, \frac{1}{\sqrt{3}} \right), \left(\frac{-1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}} \right) \right\}$ is an orthonormal subset of the inner product space $R^3(R)$ express the vector $(2,1,3)$ as a linear combination of the basis vectors of S.

Solution: Let $u_1 = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0 \right); u_2 = \left(\frac{1}{\sqrt{3}}, \frac{-1}{\sqrt{3}}, \frac{1}{\sqrt{3}} \right); u_3 = \left(\frac{-1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}} \right)$

Then $S = \{u_1, u_2, u_3\}$ and let $v = (2,1,3)$

Let $v = c_1 u_1 + c_2 u_2 + c_3 u_3$ where $c_1, c_2, c_3 \in R$ by 14.8.3.

$$c_1 =< v, u_1 >= 2\left(\frac{1}{\sqrt{2}} \right) + 1\left(\frac{1}{\sqrt{2}} \right) + 3(0) = \frac{3}{\sqrt{2}}$$

$$c_2 =< v, u_2 >= 2\left(\frac{1}{\sqrt{3}} \right) + 1\left(\frac{-1}{\sqrt{3}} \right) + 3\left(\frac{1}{\sqrt{3}} \right) = \frac{4}{\sqrt{3}}$$

$$c_3 =< v, u_3 >= 2\left(\frac{-1}{\sqrt{6}} \right) + 1\left(\frac{1}{\sqrt{6}} \right) + 3\left(\frac{2}{\sqrt{6}} \right) = \frac{5}{\sqrt{6}}$$

So $v = (2,1,3) = \frac{3}{\sqrt{2}}\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0 \right) + \frac{4}{\sqrt{3}}\left(\frac{1}{\sqrt{3}}, \frac{-1}{\sqrt{3}}, \frac{1}{\sqrt{3}} \right) + \frac{5}{\sqrt{6}}\left(\frac{-1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}} \right)$

W.E.12 : Let $V(C)$ be the inner product space of continuous complex valued functions on $[0, 2\pi]$ with inner product $< f, g >= \frac{1}{2\pi} \int_0^{2\pi} f(t)\overline{g(t)} \, dt$. Prove that $S = \left\{ f_n(t) = c^{\text{int}} / n \in Z \right\}$ is an orthonormal subset of V.

Solution: $< f_n , f_n >= \frac{1}{2\pi}\int_0^{2\pi} e^{\text{int}} \overline{e^{\text{int}}} \, dt$

$$= \frac{1}{2\pi}\int_0^{2\pi} e^{\text{int}} \overline{e}^{\text{int}} \, dt$$

So $< f_n, f_n >= \frac{1}{2\pi}\int_0^{2\pi} 1dt = \frac{1}{2\pi}[t]_0^{2\pi} = \frac{2\pi}{2\pi} = 1$

Let $m \neq n$ then

$$< f_m, f_n >= \frac{1}{2\pi}\int_0^{2\pi} e^{imt}\overline{e^{\text{int}}} \, dt = \frac{1}{2\pi}\int_0^{2\pi} e^{imt}e^{-\text{int}} \, dt$$

$$= \frac{1}{2\pi}\int_0^{2\pi} e^{i(m-n)t} \, dt$$

$$= \frac{1}{2\pi}\int_0^{2\pi} \{\cos(m-n)t + i\sin(m-n)t\}$$

$$= \frac{1}{2\pi}\left[\frac{1}{(m-n)}\sin(m-n)t - \frac{1}{(m-n)}\cos(m-n)t\right]_0^{2\pi}$$

$$= \frac{1}{2\pi}(0) = 0$$

Thus $< f_n, f_n >= 1 \forall n \in Z$ and $< f_m, f_n >= 0$ if $m \neq n, m, n \in Z$

As the required two conditions are satified $S = \left\{ f_n(t) = e^{\text{int}} / n \in Z \right\}$ is an orthonormal subset of V.

W.E.13 : Find k, so that the following pair is orthogonal $f(t) = t + k; g(t) = t^2$ where $< f, g >= \int_0^1 f(t)g(t)\,dt$

Solution: We first find $< f, g >$.

$$< f,g >= \int_0^1 (t+k)t^2 dt = \int_0^1 (t^3 + kt^2)dt$$

$$= \left[\frac{t^4}{4} + \frac{kt^3}{3} \right]_0^1 = \frac{1}{4} + \frac{k}{3}$$

As f, g are orthogonal vectors $< f,g >= 0$.

So $\frac{1}{4} + \frac{k}{3} = 0 \Rightarrow 3 + 4k = 0$

So $k = \frac{-3}{4}$

14.10 Exercise:

1. Find the vector of unit length which is orthogonal to $u = (2,-1,6)$ of $V_3(R)$ with respect to standard inner product.

Ans: $\left(\frac{2}{5}, \frac{-2}{3}, \frac{-1}{3} \right)$

2. Find two mutually orthogonal vectors each of which is orthogonal to the vector $u = (4,2,3)$ of $V_3(R)$ with respect to the standard inner product.

Ans : $(3,-3,-2);(5,17,-18)$

3. Normalize the following vectors in R^3

(i) $u = (2,3,-1)$ Ans: $\hat{u} = \left(\frac{2}{\sqrt{14}}, \frac{3}{\sqrt{14}}, \frac{-1}{\sqrt{14}} \right)$

(ii) $v = \left(\frac{1}{2}, \frac{1}{3}, \frac{-1}{4} \right);$ Ans: $\hat{v} = \left(\frac{6}{\sqrt{61}}, \frac{4}{\sqrt{61}}, \frac{-3}{\sqrt{61}} \right)$

4. Find a unit vector orthogonal to $u_1 = (1,2,1)$ and $u_2 = (3,1,0)$ in R^3 with the standard inner product.

Ans : $\frac{1}{\sqrt{35}}(1,-3,5)$

5. Let V be the vector space over R of all continuous real valued functions defined in $[0,1]$ with inner product defined by $< f, g >= \int_0^1 f(t)g(t)dt$. For each positive integer n, define $f_n(t) = \sqrt{2}\cos(2\pi nt), g_n(t) = \sqrt{2}\sin(2\pi nt)$. Then show that $S = \{1, f_1, g_1, f_2, g_2 ...\}$ is an orthonormal set.

6. Let S consists of the following vectors in R^3 $u_1 = (1,1,1), u_2 = (1,2,-3), u_3 = (5,-4,-1)$

Then 1) Show that S is orthogonal and S is a basis of R^3. (ii) Write $v = (1,5,-7)$ as a linear combination of u_1, u_2, u_3.

Ans : $v = \frac{-1}{3}u_1 + \frac{16}{7}u_2 + \frac{-4}{21}u_3$

7 (a). Find the value of K so that the pair of vectors $u = (1,2,k,3)$ and $v = (3,k,7,-5)$ are orthogonal in R^4.

Ans : $k = \frac{4}{3}$

(b) Find m so that $(m,-3,-4),(m,-m,1)$ may be orthogonal vectors in R^3.

Ans : $m = -4, 1$

8. If $\{(10,1,0),(0,1,0,1),(-1,0,1,0)\}$ is an orthogonal subset of $R^4(R)$ inner product space, obtain the orthonormal set by normalizing.

Ans : $\frac{1}{\sqrt{2}}(1,0,1,0), \frac{1}{\sqrt{2}}(0,1,0,1), \frac{1}{\sqrt{2}}(-1,0,1,0)$

9. If u, v are orthonormal vectors in $V(F)$ prove that $d(u,v) = \sqrt{2}$.

10. Two vectors u, v in an unitary space $V(C)$ are orthogonal if and only if

$$\|au + bv\|^2 = |a|^2 \|u\|^2 + |b|^2 \|v\|^2$$

11. show that $\left\{\left(\frac{1}{\sqrt{3}}, \frac{-1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right), \left(0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right), \left(\frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}, \frac{-1}{\sqrt{6}}\right)\right\}$ is an orthonormal set in R^3.

14.11 Orthonormal Basis:

Definition: A basis of an inner product space that consists of mutually orthogonal unit vectors is called an orthonormal basis.

Ex: i) The basis $S=\{(1,0),(0,1)\}$ of the inner product space $R^2(R)$ is also orthonormal. So S is an orthonormal basis of R_2.

ii) The set $S=\left\{\left(\frac{1}{\sqrt{5}},\frac{2}{\sqrt{5}}\right),\left(\frac{2}{\sqrt{5}},\frac{-1}{\sqrt{5}}\right)\right\}$ is an orthonormal basis of $R^2(R)$.

iii) The set $S=\{(1,0,0),(0,1,0),(0,0,1)\}$ is a basis of the inner product space $R^3(R)$, which is also orthonormal. So S is an orthonormal basis of $R^3(R)$.

iv) The standard ordered basis for the inner product space $V_n(R)$ is also orthonormal. So it is an ortho normal basis.

14.11.1 Finite dimensional inner product space: Definition:

A finite dimensional vector space, in which an inner product is defined is called a finite dimensional inner product space.

We now establish that every finite dimensional inner product space possesses an orthonormal basis. If S is a basis of the finite dimensional inner product space $V(F)$, we construct an orthonormal set S^1 from S such that $L(S)=L(S^1)=V$.

14.11.2 Gram - Schmidt orthogonalisation Process:

Theorem: Show that every finite dimensional inner product space has an orthonormal basis.

Proof: Let $V(F)$ be an n - dimensional inner product space. Let $B=\{u_1,u_2,...u_n\}$ be the basis of $V(F)$. We will now construct an orthonormal set in $V(F)$ with the help of elements of B.

As B is the basis of $V(F)$ so each $u_i(i=1,2,...n)\in B$ is a non zero vector.

Now $u_1 \neq O \rightarrow \|u\| \neq 0$

Further let $\frac{u_1}{\|u_1\|}=v_1(say)\neq O$

This belongs to $V(F)$ and $\|v_1\|^2 = <v_1,v_1>$

$$=<\frac{u_1}{\|u_1\|},\frac{u_1}{\|u_1\|}> \quad =\frac{<u_1,u_1>}{\|u_1\|^2}$$

(Since $\|u\|$ is real)

$$=1$$

So the set $\{v_1\}$ forms an orthonormal set in $V(F)$ and v_1 is in the linear span of u_1. Now we extend the above set by assuming $w_2 = u_2 - <u_2, v_1> v_1$ and $w_2 \in V(F)$ Evidently $w_2 \neq 0$ otherwise $w_2 = O$ would imgly $u_2 =< u_2, v_1 > v_1$ i.e. u_2 is a scalar multiple of v_1 or u_2 is a scalar multiple u_1, or u_1, u_2 are linearly dependent. Which is not possible as being elements of the basis.

Now $\frac{w_2}{\|w_2\|} = v_2(say)(\neq O) \in V(F)$

Evidently $\|v_2\| = 1$ now

$$<v_2,v_1>=<\frac{w_2}{\|w_2\|},v_1>=\frac{1}{\|w_2\|}<w_2,v_1>$$

$$=\frac{1}{\|w_2\|}<u_2-<u_2,v_1>v_1,u_1>$$

So $<v_2,v_1>=\frac{1}{\|w_2\|}<u_2,v_1>-<u_2,v_1>\frac{<v_1,v_1>}{\|w_2\|}$

$$=\frac{<u_2,v_1>}{\|w_2\|}-\frac{<u_2,v_1>}{\|w_2\|} \text{ since } <v_1,v_1>=1$$

$$=0$$

As $<v_2,v_1>=0, v_1, v_2$ are orthogonal to each other and have unit norms implying the set $\{v_1, v_1\}$ is an orthonormal set and consists of distinct vectors v_1 and v_2. As also $w_2 = u_2 - <u_2, v_1> v_1$.

$$\Rightarrow v_2\|w_2\| = u_2 - <u_2,v_1>\frac{u_1}{\|u_1\|} \qquad \text{Since } \left(\frac{w_2}{\|w_2\|}=v_2\right)$$

$$\Rightarrow v_2 = \frac{u_2}{\|w_2\|} - \frac{<u_2, v_1> u_1}{\|w_2\|\|u_1\|}$$

or v_2 is a linear combination of u_1 and u_2 i.e. u_1, u_2 generate v_2 and further more u_1 generates v_1.

Now we extend the set $\{v_1, v_2\}$ by assuming that $w_3 = u_3 - <u_3, v_1> v_1 - <u_3, v_2> v_2 \in V(F)$.

here again $<v_2, v_1> = 0, <v_3, v_2> = 0$ and $<v_3, v_3> = 1$

$$\text{Where } v_3 = \frac{w_3}{\|w_3\|} \neq O \in V(F)$$

This shows that the set $\{v_1, v_2, v_3\}$ is an orthonormal set of distinct vectors v_1, v_2, v_3. Also v_3 is a linear combination of u_3, u_2, u_1 i.e. u_1, u_2, u_3 generate v_3. Thus we have constructed an orthonormal set $\{v_1, v_2, v_3\}$ such that v_1, v_2, v_3 are distinct vectors and $v_j (j = 1, 2, 3)$ is a linear combination of $u_1, u_2, \ldots. u_j$ suppose that, in this way, we have constructed an orthonormal set $\{v_1, v_2, ..., v_k\} (k < n)$ of k distinct vectors such that $v_j (j = 1, 2, ... k)$ is a linear combination of $u_1, u_2, ..., u_j$.

To prove it by induction, we consider the vector

$$w_{k+1} = w_{k+1} - <u_{k+1}, v_1> v_1 - <u_{k+1}, v_2> v_2 .... - <u_{k+1}, v_k> v_k \quad .... (1)$$

Evidently w_{k+1} is orthogonal to each of the vectors $v_1, v_2, ... v_k$ and $w_{k+1} \neq 0_j$ other wise $w_{k+1} = 0$ would mean that u_{k+1} is a linear combination of $v_1, v_2, ... v_k$ and by assumption that v_j is a linear combination of $u_1, u_2, ... u_j$. We infer that u_{k+1} is a linear combination of $u_1, u_2, ... u_k$ which is impossible as $u_1, u_2, ... u_k, u_{k+1}$ is a linearly independent set of vectors. Now we write

$$v_{k+1} = \frac{w_{k+1}}{\|w_{k+1}\|} \Rightarrow \|v_{k+1}\| = 1$$

and also that v_{k+1} is orthogonal to each of the vectors $v_1, v_2, ... v_k$. Where $v_{k+1} \neq v_j (j = 1, 2, ... k)$. Because $v_{k+1} = v_j (j = 1, 2, ... k)$ will imply that u_{k+1} is a linear combination of $u_1, u_2, ... u_k$, which impossible. Also from the above it is clear that v_{k+1} is the linear combination of $u_1, u_2, ... u_{k+1}$. Hence we have constructed an orthonormal set $\{v_1, v_2, ..., v_k, v_{k+1}\}$ of $k+1$ distinct vectors such that $v_j\,(j=1,2,...,k+1)$ is the linear combination of $u_1, u_2, ..., u_j$. Thus by induction we can construct an orthonormal set $\{v_1, v_2, ..., v_n\}$ of n - distinct vectors such that $v_j\,(j=1,2,...,n)$ is the linear combination of $u_1, u_2, ..., u_j$.

As we know that an orthonormal set is always linearly independent, it follows that $\{v_1, v_2, ..., v_n\}$ is a linearly independent set and consequently it is a basis of $V(F)$ as it contains the number of vectors, equal to the dimension of $V(F)$. Further more this basis set is a complete orthonormal set as the maximum number of vectors in an orthonormal set in V can be n.

This method of converting a basis of $V(F)$ into a complete orthonormal set is called Gram - Schmidt orthogonalisation process.

14.11.3 Working Procedure to apply Gram - Schwidt orthogonalization process to numerical problems:

Suppose $B=\{u_1, u_2, ..., u_n\}$ is a given basis of a finite dimensional inner product space V. Let $\{v_1, v_2, ..., v_n\}$ be an orthonormal basis for V, which we are required to construct from the basis B. The vectors $v_1, v_2, ..., v_n$ will be obtained in the following way.

Take $v_1 = \dfrac{u_1}{\|u_1\|}$

$u_2 = \dfrac{w_2}{\|w_2\|}$ where $w_2 = u_2 - <u_2, v_1> v_1$

$v_3 = \dfrac{w_3}{\|w_3\|}$ where $w_3 = u_3 - <u_3, v_1> v_1 - <u_3, v_2> v_2$

$v_n = \dfrac{w_n}{\|w_n\|}$ where $w_n = u_n - <u_n, v_1> v_1 - <u_n, v_2> v_2 ... - <u_n, v_{n-1}> v_{n-1}$

Note: $\{u_1 = w_1, w_2, w_3 ... w_n\}$ is an orthogonal basis.

Worked out Examples:

W.E. 14: Apply Gram - Schmidt process to the vectors $u_1 = (1,0,1); u_2 = (1,0,-1); u_3 = (0,3,4)$ to obtain an orthonormal basis for $V_3(R)$ with standard inner product.

Solution: $u_1 = (1,0,1)$ so $\|u\| = \sqrt{1^2 + 0^2 + 1^2} = \sqrt{2}$

$$v_1 = \frac{u_1}{\|u_1\|} = \frac{1}{\sqrt{2}}(1,0,1) = \left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right)$$

$u_2 = (1,0,-1)$ so $< u_2, v_1 >=< (1,0,-1), \frac{1}{\sqrt{2}}(1,0,1)$

So $< u_2, v_1 >= \frac{1}{\sqrt{2}}\{1(1) + 0(0) + (-1)(1)\} = 0$

$$w_2 = u_2 - < u_2, v_1 > v_1 = (1,0,-1) - 0\left\{\frac{1}{\sqrt{2}}(1,0,1)\right\} = (1,0,-1)$$

$$\|w_2\| = \sqrt{1^2 + 0^2 + (-1)^2} = \sqrt{2}$$

$$v_2 = \frac{w_2}{\|w_2\|} = \frac{1}{\sqrt{2}}(1,0,-1) = \left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right)$$

$u_3 = (0,3,4)$ so $u_3 = (0,3,4) < u_3, v_1 >=< (0,3,4), \frac{1}{\sqrt{2}}(1,0,1) >$

$$= \frac{1}{\sqrt{2}} < (0,3,4), (1,0,1) >$$

$$< u_3, v_1 >= \frac{1}{\sqrt{2}}\{0(1) + 3(0) + 4(1)\} = \frac{4}{\sqrt{2}} = 2\sqrt{2}$$

$$< u_3, v_1 >=< (0,3,4), \frac{1}{\sqrt{2}}(1,0,1) >= \frac{1}{\sqrt{2}} < (0,3,4), (1,0,1) >$$

$$= \frac{1}{\sqrt{2}}\{0(1) + 3(0) + 4(1)\} = 2\sqrt{2}$$

Now $w_3 = u_3 - < u_3, v_1 > v_1 - < u_3, v_2 > v_2$

$$= (0,3,4) - 2\sqrt{2}.\frac{1}{\sqrt{2}}(1,0,1) + 2\sqrt{2}.\frac{1}{\sqrt{2}}(1,0,-1)$$

$= (0,3,4)-(2,0,2)+(2,0,-2) = (0,3,0)$

$\|w_3\| = \sqrt{0^2+3^2+0^2} = 3$

$$v_3 = \frac{w_3}{\|w_3\|} = \frac{1}{3}(0,3,0) = (0,1,0)$$

The required orthonormal basis is $\{v_1, v_2, v_3\}$

$$= \left\{\left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right), \left(\frac{1}{\sqrt{2}}, 0, \frac{-1}{\sqrt{2}}\right), (0,1,0)\right\}$$

W.E. 15: Apply Gram - Schmidt process to obtain an orthonormal basis for $V_3(R)$ with respect to the standard inner product to the vectors $(2,0,1),(3,-1,5),(0,4,2)$.

Solution: Let $\{u_1, u_2, u_3\}$ be the basis of the finite dimensional vector space $V_3(R)$ where $u_1 = (2,0,1), u_2 = (3,-1,5), u_3 = (0,4,2)$

Now $\|u_1\| = \sqrt{2^2+0^2+1^2} = \sqrt{5}$

$$v_1 = \frac{u_1}{\|u_1\|} = \frac{1}{\sqrt{5}}(2,0,1) = \left(\frac{2}{\sqrt{5}}, 0, \frac{1}{\sqrt{5}}\right)$$

$u_2 = (3,-1,5)$

$$< u_2, v_1 > = < (3,-1,5), \frac{1}{\sqrt{5}}(2,0,1) > = \frac{1}{\sqrt{5}} < (3,-1,5),(2,0,1) >$$

$$= \frac{1}{\sqrt{5}}\{3(2)+(-1)(0)+5(1)\} = \frac{11}{\sqrt{5}}$$

$w_2 = u_2 - < u_2, v_1 > v_1$ so $w_2 = (3,-1,5) - \frac{11}{\sqrt{5}}\left(\frac{1}{\sqrt{5}}\right)(2,0,1)$

So $w_2 = (3,-1,5) - \frac{11}{\sqrt{5}}(2,0,1) = \left(3-\frac{22}{5}, -1-0, 5-\frac{11}{5}\right)$

So $w_2 = \left(\frac{-7}{5}, \frac{-5}{5}, \frac{14}{5}\right) = \frac{1}{5}(-7,-5,14)$

$$\|w_2\| = \frac{1}{5}\sqrt{(-7)^2 + (-5)^2 + (14)^2} = \frac{1}{5}\sqrt{49+25+196}$$

So $\|w_2\| = \frac{1}{5}\sqrt{270}$

$$v_2 = \frac{w_2}{\|w_2\|} \frac{5}{\sqrt{270}} \cdot \frac{1}{5}(-7,-5,14) = \frac{1}{\sqrt{270}}(-7,-5,14)$$

$u_3 = (0,4,2)$ so $< u_3, v_1 > = < (0,4,2), \frac{1}{\sqrt{5}}, (2,0,1) >$

$$= \frac{1}{\sqrt{5}} < (0,4,2),(2,0,1) >$$

$$= \frac{1}{\sqrt{5}}\{0(2) + 4(0) + 2(1)\} = \frac{2}{\sqrt{5}}$$

$$< u_3, v_2 > = < (0,4,2), \frac{1}{\sqrt{270}}(-7,-5,14) >$$

$$= \frac{1}{\sqrt{270}} < (0,4,2),(-7,-5,14) >$$

$$= \frac{1}{\sqrt{270}}\{0(-7) + 4(-5) + 2(14)\}$$

$$= \frac{1}{\sqrt{270}}(-20+28) = \frac{8}{\sqrt{270}}$$

$$w_3 = u_3 - < u_3, v_1 > v_1 - < u_3, v_2 > v_2$$

$$= (0,4,2) - \frac{2}{\sqrt{5}} \cdot \frac{1}{\sqrt{5}}(2,0,1) - \frac{8}{\sqrt{270}} \cdot \frac{1}{\sqrt{270}}(-7,-5,14)$$

$$= (0,4,2) - \frac{2}{5}(2,0,1) - \frac{8}{270}(-7,-5,14)$$

$$=\left(0-\frac{4}{5}+\frac{56}{270},4-0+\frac{40}{270},2-\frac{2}{5}-\frac{112}{270}\right)$$

$$=\left(\frac{0-54(4)+56}{270},\frac{1080+40}{270},\frac{540-54\times2-112}{270}\right)$$

$$=\left(\frac{-160}{270},\frac{1120}{270},\frac{320}{270}\right)=\left(\frac{-16}{27},\frac{112}{27},\frac{32}{27}\right)$$

So $w_3=\frac{16}{27}(-1,7,2)$, $\|w_3\|=\frac{16}{27}\sqrt{(-1)^2+7^2+2^2}$

$$\|w_3\|=\frac{16}{27}\sqrt{54}$$

$$v_3=\frac{w_3}{\|w_3\|}=\frac{27}{16\sqrt{54}}.\frac{16}{27}(-1,7,2)$$

So $v_3 == \frac{1}{\sqrt{54}}(-1,7,2)=\frac{1}{3\sqrt{6}}(-1,7,2)$

Hence the required orthonormal basis is $\{v_1,v_2,v_3\}$ where $v_1=\frac{1}{\sqrt{5}}(2,0,1)$ and

$v_2=\frac{1}{\sqrt{270}}(-7,-5,14)$ and $v_3=\frac{1}{3\sqrt{6}}(-1,7,2)$

14.12 Fourier Coefficients :

Definition: Let B be an orthonormal subset (possibly infinite) of an inner product space V and let $v\in V$. We define the fourier coefficients of v relative to B to be the scalar $<v,u>$ where $u\in B$.

Worked out Examples:

W.E. 16: If $S=\{(1,1,1),(0,1,1),(0,0,1)\}$ is a subset of the vector space R^3 and $V=R^3$. Obtain the orthonormal basis B for span (S) and find the fourier coefficients of the vector (1, 0,1) to B.

Solution: Let $u_1 = (1,1,1), \quad u_2 = (0,1,1), \; u_3 = (0,0,1)$

$\|u_1\| = \sqrt{1^2+1^2+1^2} = \sqrt{3}$ so $v_1 = \dfrac{u_1}{\|u_1\|} = \dfrac{1}{\sqrt{3}}(1,1,1)$

$$<v_2, v_1> = <(0,1,1), \frac{1}{\sqrt{3}}(1,1,1)> = \frac{1}{\sqrt{3}} <(0,1,1),(1,1,1)>$$

$$= \frac{1}{\sqrt{3}}\{0(1)+1(1)+1(1)\} = \frac{2}{\sqrt{3}}$$

$$w_2 = u_2 - <v_2, v_1> v_1$$

$$= (0,1,1) - \frac{2}{\sqrt{3}} \cdot \frac{1}{\sqrt{3}}(1,1,1)$$

$$= (0,1,1) - \frac{2}{\sqrt{3}}(1,1,1) = \left(0-\frac{2}{3}, 1-\frac{2}{3}, 1-\frac{2}{3}\right)$$

$$w_2 = \left(\frac{-2}{3}, \frac{1}{3}, \frac{1}{3}\right) = \frac{1}{3}(-2,1,1)$$

So $\|w_2\| = \dfrac{1}{3}\sqrt{(-2)^2+1^2+1^2} = \dfrac{\sqrt{6}}{3}$

$$v_2 = \frac{w_2}{\|w_2\|} = \frac{1}{3}(-2,1,1).\frac{3}{\sqrt{6}} = \frac{1}{\sqrt{6}}(-2,1,1)$$

$$<u_3, v_1> = <(0,0,1), \frac{1}{\sqrt{3}}(1,1,1)> = \frac{1}{\sqrt{3}}\{0(1)+0(1)+1(1)\}$$

$$<u_3, v_1> = \frac{1}{\sqrt{3}}$$

$$<u_3, v_2> = <(0,0,1), \frac{1}{\sqrt{6}}(-2,1,1)> = \frac{1}{\sqrt{6}}\{0(-2)+0(1)+1(1)\}$$

i.e. $<u_3, v_2> = \dfrac{1}{\sqrt{6}}$

$$w_3 = u_3 - <u_3, v_1> - <u_3, v_2> v_2$$

$$= (0,0,1) - \frac{1}{\sqrt{3}}.\frac{1}{\sqrt{3}}(1,1,1) - \frac{1}{\sqrt{6}}.\frac{1}{\sqrt{6}}(-2,1,1)$$

$$= (0,0,1) - \frac{1}{3}(1,1,1) - \frac{1}{6}(-2,1,1)$$

$$= \left(0 - \frac{1}{3} + \frac{2}{6}, 0 - \frac{1}{3} - \frac{1}{6}, 1 - \frac{1}{3} - \frac{1}{6}\right) = \left(0, \frac{-1}{2}, \frac{1}{2}\right)$$

$$w_3 = \frac{1}{2}(0,-1,1)$$

$$\|w_3\| = \frac{1}{2}\sqrt{0+1+1} = \frac{\sqrt{2}}{2} = \frac{1}{\sqrt{2}}$$

$$v_3 = \frac{\varpi_3}{\|\varpi_3\|} = \sqrt{2}\left(\frac{1}{2}\right)(0,-1,1) = \frac{1}{\sqrt{2}}(0,-1,1)$$

So the orthonormal basis of Span (S) is $B = \{v_1, v_2, v_3\}$ where

$$v_1 = \frac{1}{\sqrt{3}}(1,1,1), v_2 = \frac{1}{\sqrt{6}}(-2,1,1), v_3 = \frac{1}{\sqrt{2}}(0,-1,1)$$

To find the fourier coefficients of the vector $v = (1,0,1)$ relative to $S^1 = \{v_1, v_2, v_3\}$:

$$<v, v_1> = <(1,0,1), \frac{1}{\sqrt{3}}(1,1,1)> \quad = \frac{1}{\sqrt{3}}(1)1 + 0(1) + 1(1)$$

$$<v, v_1> = \frac{2}{\sqrt{3}}$$

$$<v, v_1> = <(1,0,1), \frac{1}{\sqrt{6}}(-2,1,1)> \quad = \frac{1}{\sqrt{6}}\{1(-2) + 0(1) + 1(1)\}$$

$$<v, v_2> = \frac{-1}{\sqrt{6}}$$

$< v, v_3 >=< (1,0,1), \frac{1}{\sqrt{2}}(0,-1,1) > \quad = \frac{1}{\sqrt{2}}\{1(0)+0(-1)+1(1)\}$

$< v, v_3 >= \frac{1}{\sqrt{2}}$

Hence the fourier coefficients relative to the set B is $\frac{2}{\sqrt{3}}, \frac{-1}{\sqrt{6}}, \frac{1}{2}$ or $\frac{2\sqrt{3}}{3}, \frac{-\sqrt{6}}{6}, \frac{\sqrt{2}}{2}$

W.E. 17: If $V = L(S)$ where $S = \{(1,i,0),(1-i,2,4i)\}$. Find the orthonormal basis B of V and compute the fourier coefficients of the vector $(3+i,4i,-4)$ relative to B.

Solution: Let $u_1 = (1,i,0) u_2 = (1-i,2,4i,)$

then $\|u_1\|^2 =< u_1, u_1 >=< (1,i,0),(1,i,0) >$

$= 1(\overline{1}) + i(\overline{i}) + 0(\overline{0})$ where $\overline{a}$ is congugate of a.

$= 1(1) + i(-i) + 0(0)$

$= 1 - i^2 + 0 = 1 + 1 = 2$

$\|u_1\| = \sqrt{2}$ so $v_1 = \frac{u_1}{\|u_1\|} = \frac{1}{\sqrt{2}}(1,i,0)$

$< u_2, v_1 >=< (1-i,2,4i), \frac{1}{\sqrt{2}}(1,i,0) >$

$= \frac{1}{\sqrt{2}}\{(1-i)\overline{1} + 2\overline{i}, 4i(\overline{0})\}$

$= \frac{1}{\sqrt{2}}\{(1-i)1 + 2(-i) + 4i(0)\}$

$= \frac{1}{\sqrt{2}}\{1 - i - 2i\} = \frac{1}{\sqrt{2}}(1-3i)$

$w_2 = u_2 - < u_2, v_1 > v_1 = (1-i,2,4i), -\frac{1}{\sqrt{2}}(1-3i).\frac{1}{\sqrt{2}}(1,i,0)$

$$= (1-i, 2, 4i), -\frac{1}{2}\{(1-3i), 3+i, 0\}$$

$$= \left((1-i) - \frac{(1-3i)}{2}, 2 - \frac{(3+i)}{2}, 4i - 0 \right)$$

$$w_2 = \left(\frac{1+i)}{2}, \frac{1-i)}{2}, 4i \right) = \frac{1}{2}\left((1+i), (1-i), 8i\right)$$

$$\|w_2\|^2 = < \frac{1}{2}\left((1+i), (1-i), 8i\right), \frac{1}{2}(1+i, 1-i, 8i) >$$

$$= \frac{1}{2}\left(\frac{1}{2}\right)\left\{(1+i)\overline{(1+i)} + (1-i)\overline{(1-i)} + 8i(8\overline{i})\right\}$$ where $\overline{a}$ is congugate of a.

$$\|w_2\|^2 = \frac{1}{4}\{(1+i)(1-i) + (1-i)(1+i) + 8i(-8i)\}$$

$$= \frac{1}{4}\left\{(1+1) + (1+1) - 64i^2\right\}$$

$$\|w_2\| = \frac{1}{2}\sqrt{2+2+64} = \frac{1}{2}\sqrt{68} = \frac{2}{2}\sqrt{17} = \sqrt{17}$$

Hence $v_2 = \dfrac{w_2}{\|w_2\|} = \dfrac{1}{\sqrt{17}} . \dfrac{1}{2}(1+i, 1-i, 8i)$

So $v_2 = \left(\dfrac{1+i}{2\sqrt{17}}, \dfrac{1-i}{2\sqrt{17}}, \dfrac{4i}{\sqrt{17}} \right)$

So the orthonormal basis $B = \{v_1, v_2\}$

Where $v_1 = \left(\dfrac{1}{\sqrt{2}}, \dfrac{i}{\sqrt{2}}, 0 \right)$ and $v_2 = \left(\dfrac{1+i}{2\sqrt{17}}, \dfrac{1-i}{2\sqrt{17}}, \dfrac{4i}{\sqrt{17}} \right)$

To find the fourier coefficients of $v = (3+i, 4i, -4)$

Now $< v, v_1 > = < (3+i, 4i, -4), \dfrac{1}{\sqrt{2}}(1, i, 0) >$

$$= \frac{1}{\sqrt{2}}\{(3+i)1+(4i)(-i)-4(0)\} \qquad (\because \bar{i} = -i)$$

$$< v, v_1 >= \frac{1}{\sqrt{2}}(3+i+4) = \frac{1}{\sqrt{2}}(7+i)$$

$$< v, v_2 >=< (3+i, 4i, -4), \frac{1}{2\sqrt{17}}(1+i, 1-i, 8i) >$$

$$= \frac{1}{2\sqrt{17}}\left\{(3+i)\overline{(1+i)} + 4i\overline{(1-i)} - 4\overline{(8i)}\right\}$$

$$= \frac{1}{2\sqrt{17}}\{(3+i)(1-i)+4i(1+i)-4(-8i)\}$$

$$= \frac{1}{2\sqrt{17}}\{(3+1-2i)+(4i-4)+32i\}$$

$$= \frac{1}{2\sqrt{17}}\{(0+34i\} = \sqrt{17}i$$

Hence the fowrier coefficients of v relative to B are $< v, v_1 >, < v, v_2 >$ i.e. $\frac{1}{\sqrt{2}}(7+i), \sqrt{17}i$

14.13 Parseval's Identity:

14.13.1 Theorem: Parseval's Identity : If $B = \{u_1, u_2, ...u_n\}$ be an ortho-normal basis of a finite dimensional inner product space $V(F)$ then $< u, v >= \sum_{i=1}^{n} < u, u_i >< u_i, v >$ for all $u, v \in V$.

Proof: $B = \{u_1, u_2, ...u_n\}$ is a basis of V. Let $u, v \in V$. Then there exists scalars $a_1, a_2, ...a_n$, $b_1, b_2, ...b_n \in F$ so that

$$u = a_1u_1 + a_2u_2 + ... + a_nu_n = \sum_{i=1}^{n} a_iu_i$$

$$v = b_1u_1 + b_2u_2 + ... + b_nu_n = \sum_{J=1}^{n} b_ju_j$$

Now $<u,v>=<\sum_{i=1}^{n} a_i u_i, \sum_{j=1}^{n} b_j u_j>,$

$$= \sum_{i=1}^{n} < a_i u_i,\ b_i u_i > \text{ since } < u_i, u_j > 0 \text{ for } i \neq j$$

$$= \sum_{i=1}^{n} a_i \overline{b_i} < u_i, u_i >= \sum_{i=1}^{n} a_i \overline{b_i} \text{ .... (1)}$$

since $< u_i, u_j >= 1$ if $i = j$

But $< u, u_i >= \sum_{i=1}^{n} < a_i u_i, u_i >= \sum_{i=1}^{n} a_i < u_i, u_i >= a_i$ .... (2) Since $< u_i, u_j >= \begin{cases} 0 \text{ if } i \neq j \\ 1 \text{ if } i = j \end{cases}$

and $< u_i, v >=< u_i, \sum_{J=1}^{n} < b_j u_j >$

$$= \overline{b_i} < u_i, u_i > \qquad \text{Since } < u_i, u_j >= \begin{cases} 0 \text{ if } i \neq j \\ 1 \text{ if } i = j \end{cases}$$

$$= \overline{b_i}(1) = \overline{b_i} \text{ .... (3)}$$

Using (2) and (3) in (1) we get

$$< u, v >= \sum_{i=1}^{n} < u, u_i >< u_i, v > \text{ for all } u, v \in V$$

14.13.2 Corollary: If $S = \{u_1, u_2, ...u_n\}$ is a complete orthonormal set in an inner product space $V(F)$ and if $v \in V$, then $\sum_{i=1}^{n} |< v, u_i >|^2 = \|v\|^2$

Proof: By Parseval's Identity if $v \in V$ then $< u, v >= \sum_{i=1}^{n} < u, u_i >< u_i, v >$ for all $v \in V$

So as $v \in V$, $< v, v >= \sum_{i=1}^{n} < v, u_i > < u_i, v >$

$$= \sum_{i=1}^{n} < v, u_i > \overline{< v, u_i >}$$

So $\|v\|^2 = \sum_{i=1}^{n} |< v, u_i >|^2 \ \therefore z\overline{z} = |z|^2$

Thus $\sum_{i=1}^{n} |< v, u_i >|^2 = \|v\|^2$

14.13.3 Theorem:

Bessel's Inequality: Let V be an inner product space and Let $S = \{u_1, u_2, ...u_n\}$ be an orthonormal subset of V. Prove that for any $v \in V$, we have $\sum_{i=1}^{n} |< v, u_i >|^2 \le \|v\|^2$.

Further more the equality holds if and only if v is in the subspace generated by $u_1, u_2, ...u_n$.

Proof: Consider the vector $w = v - \sum_{i=1}^{n} < v, u_i > u_i$

Now $\|w\|^2 =< w, w >=< v - \sum_{i=1}^{m} < v, u_i > u_i, v - \sum_{j=1}^{m} < v, u_j > u_j >$

$$=< v, v > - \sum_{i=1}^{m} < v, u_i >< u_i, v > - \sum_{j=1}^{m} \overline{< v, u_j >} < v, u_j > + \sum_{i=1}^{m} \sum_{j=1}^{m} < v, u_i > \overline{< v_i u_j >} < u_i, u_j >$$

$$=< v, v > - \sum_{i=1}^{m} < v, u_i > \overline{< u_i, u_j >} - \sum_{i=1}^{m} \overline{< v, u_j >} < v, u_j > + \sum_{i=1}^{m} < v, u_i > \overline{< v, u_i >}$$

On summing up with respect to j and remembering $< u_i, u_j >= \begin{cases} 1 \text{ if } i = j \\ 0 \text{ if } i \neq j \end{cases}$

$$= \|v\|^2 - \sum_{i=1}^{m} |< v, u_i >|^2 - \sum_{i=1}^{m} |< v, u_i >|^2 + \sum_{i=1}^{m} |< v, u_i >|^2$$

i.e. $\|w\|^2 = \|v\|^2 - \sum_{i=1}^{m} |< v, u_i >|^2$ .... (1)

Now $\|w\|^2 \geq 0 \Rightarrow \|v\|^2 - \sum_{i=1}^{m} |< v, u_i >|^2 \geq 0$

$$\Rightarrow \sum_{i=1}^{m} |< v, u_i >|^2 \leq \|v\|^2$$

To show that equality hold if and only if v is in the subspace spanned by $u_1, u_2, ...u_m$:

Case i) Let the equality holds good i.e. $\sum_{i=1}^{m} |< v, u_i >|^2 = \|v\|^2$ then from (1)

i.e. $\|w\|^2 = \|v\|^2 - \sum_{i=1}^{m} |< v, u_i >|^2$

We get $\|w\|^2 = 0 \Rightarrow w = O$ (zero vector)

$$\Rightarrow v - \sum_{i=1}^{m} < v, u_i > u_i = O$$

$$\Rightarrow v = \sum_{i=1}^{m} < v, u_i > u_i$$

Thus if the equality holds good then v is a linear combination of $u_1, u_2, ...u_m$

Hence v is in the subspace spanned by $u_1, u_2, ...u_m$

Case ii) Converse : v is in the subspace spanned by $u_1, u_2, ..., u_m$

So v can be expressed as a linear combination of $u_1, u_2, ..., u_m$, by theorem 14.8.2. We know that

$v = \sum_{i=1}^{m} < v, u_i > u_i$ .... (2) But we know

$w = v - \sum_{i=1}^{m} < v, u_i > u_i = O$ using (2)

So $\|w\|^2 = 0$

$\Rightarrow \|v\|^2 - \sum_{i=1}^{m} |< v, u_i >|^2 = 0$

$\Rightarrow \sum_{i=1}^{m} |< v, u_i >|^2 = \|v\|^2$

Thus the equality holds good.

Hence the theorem.

14.13.4 Corollary: Let $\{u_1, u_2, ...u_m\}$ be an orthogonal set of non zero vectors in an inner product space V. If v is any vector in V, then $\sum_{i=1}^{m} \left(\frac{|< v, u_i >|^2}{\|u_i\|^2} \right) \leq \|v\|^2$

Proof: Let $B = \{v_1, v_2, ..., v_m\}$ where $v_i = \frac{u_i}{\|u_i\|}$ $(1 \leq i \leq m)$. Then $\|v_i\| = 1$, so the set B is an orthonormal set. Hence by Bessels's inequality we get $\sum_{i=1}^{m} |< v, u_i >|^2 \leq \|v\|^2$ .... (1)

Also $< v, v_i > = < v, \frac{u_i}{\|u_i\|} > \quad = \frac{1}{\|u_i\|} < v, u_i >$

So $|< v, v_i >|^2 = \frac{1}{\|u_i\|^2} |< v, u_i >|^2$ .... (2)

From (1) and (2) we get

$$\sum_{i=1}^{m} \left(\frac{|< v, u_i >|^2}{\|u_i\|^2} \right) \leq \|v\|^2$$

14.13.5 Theorem: If V is a finite dimensional inner product space, and if $\{u_1, u_2, ...u_m\}$ is an orthonormal set in V, such that $\sum_{i=1}^{m} |< v, u_i >|^2 = \|v\|^2$ for every $v \in V$; prove that $\{u_1, u_2, ...u_m\}$ must be a basis of V.

Proof: Let v be any vector in V. Consider

$$w = v - \sum_{i=1}^{m} < v, u_i > u_i \text{ .... (1)}$$

As in the proof of Bessel's inequality

We have $\|w\|^2 = < w, w >$

$$= \|v\|^2 - \sum_{i=1}^{m} |< v, u_i >|^2$$

$= 0$ by given condition.

$$\Rightarrow w = O \quad \Rightarrow v = \sum_{i=1}^{m} < v, u_i > u_i$$

Thus every vector v in V can be expressed as a linear combination of the vectors in the set $S = \{u_1, u_2, ...u_m\}$ i.e. $L(S) = V$. As S is an orthonormal basis, S is linearly independent.

Hence S is a basis of V.

14.14 Orthogonal Compliment:

14.14.1 Definition: Let S be a non empty subset of an inner product space V. We define $S^{\perp}$ (read S perp) to be the set of all vectors in V; that are orthogonal to every vector in S; i.e.

$S^{\perp} = \{u \in V / < u, v >= 0 \text{ for all } v \in S\}$

$S^{\perp}$ is called the orthogonal complement of S and the symbol is usually read as S perpendicular

Note i) $S^{\perp} \subset V$

ii) $u \in S^{\perp}$, $v \in S$ $\Rightarrow < u, v >= 0$

iii) **O** (zero vector) is an element in V. $v \in S$

then $< O, u >= 0 \forall u \in S \Rightarrow O \in S^{\perp}$. Hence $S^{\perp} \neq \phi$

14.14.2 Theorem: If S is any non empty subset of an inner product space $V(F)$, then $S^{\perp}$ is a subspace of $V(F)$.

Proof: By definition $S^{\perp} = \{u \in V / < u, v >= 0 \forall v \in V\}$

Let $u_1, u_2 \in S^{\perp}$ and $v \in S$. then $< u_1, v >= 0$ .... (1)

and $< u_2, v >= 0$ .... (2) Now for $a, b \in F$ and for each $v \in S$ we have

$$< au_1 + bu_2, v >=< au_1, v > + < bu_2, v >$$

$$= a < u_1, v > + b < u_2, v >$$

$$= a(0) + b(0) \text{ using (1) and (2)}$$

$$= 0$$

$\therefore$ For $u_1, u_2 \in S^{\perp}$ $\quad a, b \in F$

$au_1 + bu_2 \in S^{\perp}$ so $S^{\perp}$ is a subspace of V.

14.14.3 Theorem: If $V(F)$ is an inner product space, **O** is the zero vector in V, then show that $\{O\}^{\perp} =.V$

Proof: Let $v \in V$, then $< v, O >= 0$ by definition of inner product. So $v \in O^{\perp}$ i.e. any element v of V is also an element of $S^{\perp}$. So $V \subseteq \{0\}^{\perp}$ .... (1). Also $\{0\}^{\perp} \subseteq V$ .... (2) from (1) and (2) $\{0\}^{\perp} = V$.

14.14.4 Theorem:

If $V(F)$ is an inner product space, **O** is the zero vector of V; then show that $V^{T} = \{0\}$.

Proof: Let $u \in V^{\perp}$ then $< u, v >= 0 \forall v \in V$ by definition of $V^{\perp}$ when $v = u$ then $< u, u >= 0$

i.e. $\|u\|^2 = 0 \Rightarrow u = O$

Thus **O** is the only vector orthogonal to itself and hence $V^{\perp} = \{O\}$

14.14.5 Theorem:

If S is a subset of an inner product space $V(F)$, then show that $S \wedge S^{\perp} = \{O\}$

Proof: If $u \in S \wedge S^{\perp}$ then $u \in S$ and $u \in S^{\perp}$

$\Rightarrow$ u is orthogonal to u.

$\Rightarrow < u, u >= 0 \Rightarrow u = O$

So $S \wedge S^{\perp} = \{O\}$

14.14.6 Theorem:

If S_1, S_2 are two subsets of an inner product space $V(F)$ then show that $S_1 \subseteq S_2 \Rightarrow S_2^{\perp} \subseteq S_1^{\perp}$

Solution: Let $u \in S^{\perp} \Rightarrow u$ is orthogonal to every vector in S_2.

As $S_1 \subseteq S_2$ so u is orthogonal to every vector in S_1.

$\Rightarrow$ $u \in S_1^{\perp}$ thus $u \in S_2^{\perp} \Rightarrow u \in S_1^{\perp}$

So $S_2^{\perp} \subseteq S_1^{\perp}$

14.14.7 Theorem:

If S is a subset of an inner product space $V(F)$, then show that $S^{\perp} =$ span $(S)^{\perp}$.

Proof: V is an inner product space over the field F.

S is a subset of V So $S \subseteq$ span (S)

We know if S_1, S_2 are two subsets of V, then $S_1 \subseteq S_2 \Rightarrow S_2^{\perp} \subseteq S_1^{\perp}$

Hence by this, [span of $S]^{\perp} \subseteq S^{\perp}$ .... (1)

Let $u \in S^{\perp}$ and $v \in$ span of (S). Hence there exists scalars $a_1, a_2, ..., a_n$ in F such that $v = a_1 w_1 + a_2 w_2 + ... + a_n w_n$ where $w_1, w_2, ... w_n \in S$

As $u \in S^{\perp}$ $< u, v >=< u, a_1 w_1 + a_2 w_2 + ... + a_n w_n >$

So $< u, v >= \bar{a}_1 < u, w_1 > + \bar{a}_2 < u, w_2 > + ... + \bar{a}_n < u, w_n >$

$= \bar{a}_1 (0) + \bar{a}_2 (0) + ... + \bar{a}_n (0) = 0$

As $v \in$ span of S and $<u,v>=0$ where $u \in S^{\perp}$

So $u \in$[span of $S]^{\perp}$

Thus $u \in S^{\perp} \Rightarrow u \in$ (Span $S)^{\perp}$

$\therefore S^{\perp} \subseteq$ (Span $S)^{\perp}$ .... (2)

From (1) and (2) $S^{\perp} =$ (span $S)^{\perp}$

14.14.8 Theorem: If $B = \{u_1, u_2, ...u_m\}$ is an orthonormal subset of the inner product space $V(F)$, then for each $v \in V, w = v - \sum_{j=1}^{m} < v, u_j >$ is a vector of $B^{\perp}$.

Proof: We have proved in theorem 14.8.4 that w is orthogonal to each of $u_1, u_2, ...u_m$. By definition of orthogonal complement $w \in B^{\perp}$.

Hence the theorem.

14.14.9 Orthogonal Compliment of an orthogonal compliment:

Definition: If S is a subset of an inner product space $V(F)$, then $S^{\perp}$ is a sub space V. We define $(S^{\perp})^{\perp}$ written as $S^{\perp\perp}$ (containing those vectors in $V(F)$ which are orthogonal to each vector of $S^{\perp}$) by $(S^{\perp})^{\perp} = S^{\perp\perp} \{u \in V : < u, v > \forall v \in S^{\perp}\}$

Note: Obviously $(S^{\perp})^{\perp}$ is a subspace of V; since we know if S is any set of vectors in an inner product space $V(F)$ then $S^{\perp}$ is a subspace of V.

14.14.10 Theorem:

Show that for any subset S of an inner product space $V(F)$, $S \subseteq S^{\perp\perp}$.

Solution : $V(F)$ is the given inner product space S, is subset of V. then $S^{\perp}, S^{\perp\perp}$ are subspaces of V.

Let $u \in S$ then $< u, v >= 0$ for all $v \in S^{\perp}$

$< u, v >= O < v, u >= O$ for $v \in S^{\perp}$ and $u \in V$

So by definition $u \in S^{\perp\perp}$

Thus $u \in S \Rightarrow u \in S^{\perp\perp}$

So $S \subseteq S^{\perp\perp}$.

W.E. 18: If $V(F)$ is an inner product space and S is any subset of V then show that

i) $S^{\perp} = (L(S))^{\perp}$

ii) $L(S) \subseteq S^{\perp\perp}$

iii) $L(S) = S^{\perp\perp}$ if V is finite dimensional

iv) $S^{\perp} = S^{\perp\perp\perp}$

Solution: $V(F)$ is an inner product space S is a subset of V.

i) To show $S^{\perp} = (L(S))^{\perp}$

As S is a subset of V, $S \subseteq L(S)$

Hence $[L(S)]^{\perp} \subseteq S^{\perp}$ .... (1)

To show $S^{\perp} \subseteq [L(S)]^{\perp}$

Let $v \in L(S)$ then $v = \sum_{i=1}^{n} a_i u_i \; \forall \; u_i \in S$

for $u \in S^{\perp}$ we have $< u, v > = < u, \sum_{i=1}^{n} a_i u_i >$

$= \sum_{i=1}^{n} \overline{a}_i < u, u_i >$ by definition.

$= 0$ since u is perpendicular to each $u_i \in S$.

This implies u is perpendicular to $v \in L(S)$ i.e. $u \in (L(S))^{\perp}$ or $S^{\perp} \subseteq [L(S)]^{\perp}$ ...... (2)

From (1) and (2) $S^{\perp} = [L(S)]^{\perp}$

ii) To show that $L(S) \subseteq S^{\perp\perp}$

Let $u \in L(S)$ and $v \in S^{\perp}$. Then v is orthogonal to every vector of S or inother words v is orthogonal to the linear combination of a finite number of vectors in S. i.e. v is orthogonal to u.

$\Rightarrow u \in (S^{\perp})^{\perp}$ i.e. $u \in S^{\perp\perp}$

Thus $u \in L(S) \Rightarrow u \in S^{\perp\perp}$. So $L(S) \subseteq S^{\perp\perp}$

iii) To show $S^{\perp} = S^{\perp\perp\perp}$

We have $S \subseteq L(S)$ and $L(S) \subseteq S^{\perp\perp}$ from (iii)

So $S \subseteq S^{\perp\perp} \Rightarrow S^{\perp} \subseteq (S^{\perp\perp})^{\perp}$

$\Rightarrow S^{\perp} \subseteq S^{\perp\perp\perp}$ .... (1)

Now $S \subseteq S^{\perp\perp} \Rightarrow S, (S^{\perp\perp})^{\perp} \subseteq S^{\perp}$

$\Rightarrow S^{\perp\perp\perp} \subseteq S^{\perp}$ .... (2)

From (1) and (2) we get $S^{\perp} = S^{\perp\perp\perp}$

14.14.11 Theorem: Let W be a finite dimensional subspace of an inner product space V. Let $v \in V$ then there exists unique vectors $u \in W$ and $w \in W^{\perp}$ such that $v = u + w$.

Proof: Let $B = \{u_1, u_2, ...u_n\}$ be an orthonormal basis of W. Then B is linearly independent and $L(B) = W$. Let u be defined as $u = \sum_{i=1}^{n} < v, u_i > u_i$ and $w = v - u$.

Now $u \in W = L(B)$ and $v = u + w$ now we have to prove that $w \in W^{\perp}$.

As B is an orthonormal basis of the vector space W. $w = v - \sum_{i=1}^{n} < v, u_i > u_i$ a vector of $W^{\perp}$.

So for $v \in V$, there exists $u \in W$ and $w \in W^{\perp}$ such that $v = u + w$.

To show the uniqueness:

Let if possible $v = u' + w'$ where $u' \in W$ and $w' \in W^{\perp}$

Then $u + w = u' + w' \Rightarrow u - u' = w' - w$

But $u \in W, u' \in W', W$ is a subspace $\Rightarrow u - u^1 \in W$

$w' \in W^{\perp}$ and $w \in W^{\perp} \Rightarrow w' - w \in W^{\perp}$ since $W^{\perp}$ is a subspace and as $W \wedge W^{\perp} = \{O\}$ and so $u - u' = W' - w = 0 \Rightarrow u = u', w = W'$

So the representation $v = u + w$ is unique.

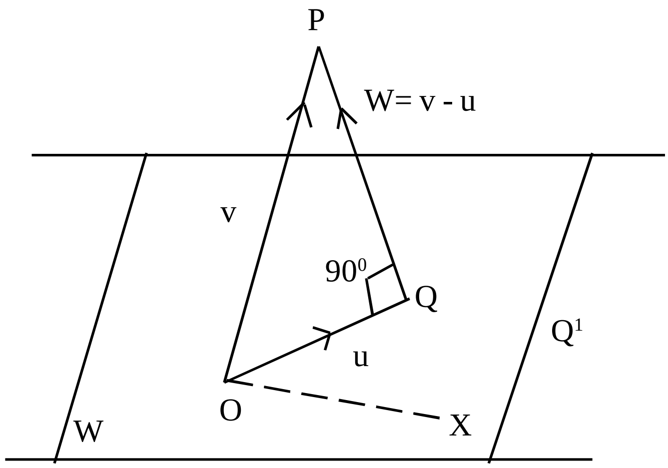
Note: For $u \in W$, there exists $w \in w^{\perp}$ such that $u + w = v \in V$

14.14.12 Closest Vector:

If $V = R^3$ and $S = \{e_3\}$ where $e_3 = (0,0,1)$ there $S^{\perp}$ is equal to xy plane.

Consider the problem in R^3 of finding the distance from a point P to a plane W.

If we let v be the vector determined by O and P, we may restate the problem as follows.



Determine the vector u in W that is ctosest to v. The desired distance is clearly $\|v-u\|$ we observe from the figure that the vector $w = v - u$ is orthogonal to every vector in W and so $w \in W^{\perp}$.

For any $x \in W$, there exists a point $Q' \in W$ plane so that $PQ' > PQ$.

i.e. $|v - x| \geq v - u$

The vector $u \in W$ is clearly the orthogonal projection of $v \in V$ on the plane.

14.14.3 Orthogonal Projection: Let W be a subspace of the finite dimensional inner product space. For $v \in V$ there exists unique vectors $u \in W, w \in W^{\perp}$ such that $v = u + W$.

The vector $u \in W$ that is $u = \sum_{i=1}^{n} < v, u_i > u_i$ where $\{u_1, u_2, ..., u_n\}$ is an orthonormal basis of W, is called the orthogonal projection of $v \in V$ on the subspace W.

14.14.14 Theorem: Let $S = \{v_1, v_2, ...v_k\}$ is an orthonormal set in an n - dimensional inner product space V. Then show that S can be extended to an orthonormal basis $S^1 = \{v_1, v_2, ...v_k, v_{k+1}...v_n\}$ for V.

Proof: $V(F)$ is an n - dimensional inner product space. So $V(F)$ is an n dimensional vector space. $S = \{v_1, v_2, ...v_k\}$ is an orthonormal subset of V. So it can be extended to the set $S_1 = \{v_1, v_2, ...v_k, w_{k+1}, w_{k+2}...w_n\}$, so as to form a basis of the vectorspace $V(F)$.

By applying Gram schmidt orthogonalisation process to S_1 we can obtain an orthonormal basis in which the first k vectors are the vectors in S; the last n - k vectors are obtained after normalising, given by $S' = \{v_1, v_2, ...v_k, v_{k+1}...v_n\}$. So that $L(S') = V$.

Hence the theorem.

14.14.15 Corollary 1: If $S = \{v_1, v_2, ...v_k\}$ is an orthonormal set in an inner product space V and if W is span (S) i.e. W is a subspace of V, then $S_2 = \{v_{k+1}, v_{k+2}, ...v_n\}$ is an orthonormal basis of $W^{\perp}$.

Proof: In the above theorem we have shown that

S can be estended to $S' = \{v_1, v_2, ...v_k, v_{k+1}, ...v_n\}$

So $S_2 = \{v_{k+1}, v_{k+2}..., v_n\} \subseteq S'$ (Basis of V)

Hence S_2 is linearly independent, being the subset of basis of V.

We are given $L(S) = W$ and $S \wedge S_2 = \phi$,so $S_2 \subset W^{\perp}$ we will now show that $(S_2) = W^{\perp}$.

As $L(S') = V$ for any $u \in V$, we have $u = \sum_{i=1}^{n} < u, u_i > u_i$

Let $u \in W^{\perp} \Rightarrow < u, u_i >= 0$ for $i = 1, 2, ..k$

since $S \subset W$.

$\therefore$ for each $u \in W^{\perp}$,

$$u =< u, u_1 > u_1 + < u, u_2 > u_2 + ... + < u, u_k > u_k + \sum_{i=k+1}^{n} < u, u_i > u_i$$

$$= O + O + ... + O + \sum_{i=k+1}^{n} < u, u_i > u_i$$

$$= \sum_{i=k+1}^{n} < u, u_i > u_i$$

So $u \in W^{\perp} \to u \in L(S_2) \Rightarrow W^{\perp} \subset L(S_2)$ .... (2)

From (1) and (2) $L(S_2) = W$

As S_2 is linearly independent $L(S_2) = W$

$S_2 = \{u_{k+1}, u_{k+2}, ...u_n\}$ is a basis of $W^{\perp}$.

14.14.16 Corollary 2: V is an n dimensional inner product space and if W is any subspace of V, then show that $\dim(V) = \dim(W) + \dim(W^{\perp})$.

Proof: As V is a finite dimensional inner product space and W is a subspace V, so W is finite dimensional inner product space. So it has an orthonormal basis $S = \{v_1, v_2, ...v_k\}$. So it can be extended to an orthonomal basis $S_1 = \{v_1, v_2, ...v_k, v_{k+1}, ...v_n\}$ and by the above corollary 1, $S_1 = \{.v_{k+i}, v_{k+2}, ...v_n\}$ is an orthonormal basis of $W^{\perp}$.

So dim $V = n = k + (n-k)$

$$= \dim(W) + \dim(W^{\perp})$$

Hence $\dim V = \dim(W) + \dim(W^{\perp})$.

14.14.17 Projection Theorem:

If W is any subspace of a finite dimensional inner product space then show that

i) $v = W \oplus W^{\perp}$ where $W^{\perp}$ is an orthogonal complement of W.

ii) $(W^{\perp})^{\perp} = W$

Proof: W being a subspace of a finite dimensional vector spae $V(F)$ of dimension n, is also finite dimensional say of dimension k.

Thus we can find $B_1 = \{u_1, u_2, ...u_k\}$ as an orthonormal set in W which is also a basis of W. This can be exteneded to give $B = \{u_1, u_2, ...u_k, u_{k+1}...u_n\}$ as an orthonormal basis for $V(F)$.

As W is a subspace of V, then $W^{\perp}$ is also a subspace of V; and $W \wedge W^{\perp} = \{0\}$ ...... (1) where 0 is the zero vector.

We will now prove $V = W + W^{\perp}$

Now consider the vector $w = v - \sum_{i=1}^{k} < u, u_i > u_i$ .... (2)

Now $< w, u_j > = < v - \sum_{i=1}^{k} < v, u_i > u_i, u_j > 1 \le j \le k$

$$= < v, u_j > - \sum_{i=1}^{k} < v, u_i > < u_i, u_j >$$

$$= < v, u_j > - < v, u_j > < u_j, u_i >$$

Since $< u_i, u_j > = \begin{cases} 1 \text{ if } i = j \\ 1 \text{ if } i \neq j \end{cases}$

So $< w, u_j > = < v, u_j > - < v, u_j > = 0$

Showing that W is orthogonal to each of the vectors $u_1, u_2, ...u_k$. i.e. orthogonal to the sub-space W spanned by there vectors and hence it belongs to $W^{\perp}$.

Also, the vectors $\sum_{i=1}^{k} < v, u_i > u_i$ being a linear combination of elements of $B_1 \in W$.

Hence from (2) i.e. $w = v - \sum_{i=1}^{k} < v, u_i > u_i$

each $v \in V$ we have

$$v = \left[\sum_{i=1}^{k} < v, u_i > u_i \right] + w$$

= an elements of W + an elements of $W^{\perp}$.

So $V = W + W^{\perp}$ .... (3)

and we have $W \wedge W^{\perp} = \{0\}$ .... (1)

So from (1) and (3) $V = W_1 \oplus W_2$

ii) To prove that $(W^{\perp})^{\perp} = W$ when W is a subspace of an inner product space of finite dimension.

Proof: By definition $(W^{\perp})^{\perp} = W^{\perp\perp} = \{w \in V :< w, v >= 0 \forall v \in W^{\perp}\}$

Let $u \in W \Rightarrow < u, v >= 0 \forall \vee \in W^{\perp}$

Hence from definition of $W^{\perp\perp}, u \in W^{\perp\perp}$

Thus $u \in W \Rightarrow u \in W^{\perp\perp}$ so $W \subseteq W^{\perp\perp}$

We have $V = W \oplus W^{\perp}$ .... (1)

So $\dim V = \dim W + \dim W^{\perp}$ .... (2)

Putting $W^{\perp}$ fro W in (1) we get

$V = W^{\perp} \oplus W^{\perp\perp}$ .... (3)

So $\dim V = \dim(W^{\perp}) + \dim(W^{\perp\perp})$ .... (4)

From (2) and (4) we get $\dim W = \dim W^{\perp\perp}$

Now $W \subseteq W^{\perp\perp}$. Hence W is a subspace of $W^{\perp\perp}$ and $\dim W = \dim W^{\perp\perp}$ so $W = W^{\perp\perp}$

Thus $(W^{\perp})^{\perp} = W$.

Note: If W is a subspace of any finite dimensional inner product space V(F), then $V = W \oplus W^{\perp}$

$\Rightarrow \dim V = \dim W + \dim W^{\perp}$

$\Rightarrow \dim W^{\perp} = \dim V - \dim W$

14.15 Worked Out Examples:

W.E.19: If ϖ_1 and ϖ_2 are two subspaces of a finite dimensional inner product space V(F) then show that i) $(W_1 + W_2)^{\perp} = W_1^{\perp} \wedge W_2^{\perp}$

and ii) $(W_1 + W_2)^{\perp} = W_1^{\perp} + W_2^{\perp}$

Solution: i) To show $(W_1+W_2)^{\perp}=W_1^{\perp}\wedge W_2^{\perp}$

We know $S_1\subseteq S_2\Rightarrow S_2^{\perp}\subseteq S_1^{\perp}$

Now $W_1\subseteq W_1+W_2\Rightarrow(W_1+W_2)^{\perp}\subseteq W_1^{\perp}$

and $W_2\subseteq W_1+W_2\Rightarrow(W_1+W_2)^{\perp}\subseteq W_1^{\perp}$

Hence from the above two, it follows

$(W_1+W_2)^{\perp}\subseteq W_1^{\perp}\wedge W_2^{\perp}$ .... (1)

To show that $W_1^{\perp}\wedge W_2^{\perp}\subseteq(W_1+W_2)^{\perp}$

Let $w\in W_1^{\perp}\wedge W_2^{\perp}\Rightarrow w\in W_1^{\perp}$, and $w\in W_2^{\perp}$

$\Rightarrow< w,u >=0 \forall u\in W_1$

and $< w,v >=0 \forall v\in W_2$

Thus $< w,u >=0$ and $< w,v >=0\Rightarrow< w,u+v >=0$ where $u+v\in W_1+W_2$

So $w\in(W_1+W_2)^{\perp}$

Thus $w\in W_1^{\perp}\wedge W_2^{\perp}\Rightarrow w\in(W_1,+W_2)^{\perp}$

So $W_1^{\perp}\wedge W_2^{\perp}\subseteq(W_1,+W_2)^{\perp}$ .... (2)

From (1) and (2) $(W_1,+W_2)^{\perp}=W_1^{\perp}\wedge W_2^{\perp}$.

ii) To show that $(W_1,+W_2)^{\perp}=W_1^{\perp}+W_2^{\perp}$

As W_1,W_2 are subspace of $V;W_1^{\perp},W_2^{\perp}$ are also subspaces of V. Hence replacing W_1 and W_2 by $W_1^{\perp}$ and $W_2^{\perp}$ respectively in the above i.e.

$(W_1+W_2)^{\perp}=W_1^{\perp}\wedge W_2^{\perp}$

$(W_1^{\perp}+W_2^{\perp})^{\perp}=(W_1^{\perp})^{\perp}\wedge(W_2^{\perp})^{\perp}$

$\Rightarrow(W_1^{\perp}+W_2^{\perp})^{\perp}=W_1\wedge W_2$

Since $W_1^{\perp\perp} = W_1$ and $W_2^{\perp\perp} = W_2$

$\Rightarrow (W_1^{\perp} + W_2^{\perp})^{\perp\perp} = (W_1 \wedge W_2)$

So $(W_1 \wedge W_2)^{\perp} = W_1^{\perp} + W_2^{\perp}$ since $W^{\perp\perp} = W$

W.E.20: For $R^3(R)$ space $W = L\left(\{(1,0,0),(0,1,0)\}\right)$ is a subspace Let $v = (2,3,4) \in R^3$. Orthonormal basis of $W = \{u_1, u_2\}$ where $u_1 = (1,0,0), u_2 = (0,1,0)$.

Find the orthogonal projection of v and w.

Solution: The orthogonal projection of v on w is equal to $\sum_{i=1}^{2} < v, u_i > u_i = < v, u_1 > u_1 + < v, u_2 > u_2$

$= < (2,3,4),(1,0,) > (1,0,0) + < (2,3,4),(0,1,0) > (0,1,0)$

$= 2(1,0,0) + 3(0,1,0)$

$= (2,3,0)$

W.E. 21 :

Let $V = P_3(R)$ be the inner product space of atmost 3rd degree polynomials continues on $[-1,1]$. Let the inner product be defined as $< f, g > = \int_{-1}^{1} f(t)g(t)dt$ where $f, g \in V$, If $W = P_2(R)$ is a sub space of V, with standard basis $B = \{1, x, x^2\}$.

i) Obtain orthonormal basis by Gram-schmidt process.

ii) Represent the polynomial $f(x) = 1 + 2x + 3x^2 \in P_2(R)$ as a linear combination of the orthonomial basis obtained above.

iii) Obtain the orthogonal projection of $f(x) = x^3$ belonging to $P_2(R)$, the subspace of $P_3(R)$.

Solution: i) To obtain on orthonormal basis by Gram schmidt process.

The standard basis of $P_3(R)$ is $B = \{u_1, u_2, u_3\}$ given $u_1 = 1$, $u_2 = x$, $u_3 = x^2$

Let $B' = \{v_1, v_2, v_3\}$ be the corresponding orthonormal basis.

$u_1 = 1 \quad \|u_1\|^2 = < u_1, u_1 > = \int_{-1}^{1} (1)(1) dt . [t]_{-1}^{1} \leq 1 + 1 = 2, \quad \|u\| = \sqrt{2}$

$$v_1 = \frac{u_1}{\|u_1\|} = \frac{1}{\sqrt{2}}$$

$$< u_2, v_1 >= \int_{-1}^{1} (t).\frac{1}{\sqrt{2}} dt = \frac{1}{\sqrt{2}}\left[-\frac{t^2}{2}\right]_{-1} = \frac{1}{2\sqrt{2}}(1-1) = 0$$

$$w_2 = u_2 - < u_2, v_1 > v_1$$

$$w_2 = x - 0 = x$$

$$\|w_2\|^2 =< w_2, w_2 >= \int_{-1}^{1} t.tdt = 2\int_{0}^{1} t^2 dt = 2\left[\frac{t^3}{3}\right]_0^1 = \frac{2}{3}$$

$$\|w_2\| = \sqrt{\frac{2}{3}}$$

$$v_2 = \frac{w_2}{\|w_2\|} = \sqrt{\frac{3}{2}}(x) = \sqrt{\frac{3}{2}}(x)$$

$$< u_3, v_1 >= \int_{-1}^{1} t^2.\frac{1}{\sqrt{2}} dt = \frac{2}{\sqrt{2}}\int_{0}^{1} t^2 dt = \sqrt{2}\left[\frac{t^3}{3}\right]_0^1 = \frac{\sqrt{2}}{3}$$

$$< u_3, v_2 >= \int_{-1}^{1} t^2.\sqrt{\frac{3}{2}} t dt = \sqrt{\frac{3}{2}}\int_{0}^{1} t^3 dt = \sqrt{\frac{3}{2}}\left[\frac{t^4}{4}\right]_0^1$$ i.e. $< u_3, v_2 >= 0$

$$w_3 = u_3 - < u_3, v_2 > v_2 - < u_3, v_1 > v_1$$

$$= x^2 - 0\sqrt{\frac{3}{2}}x - \frac{\sqrt{2}}{3}\frac{1}{\sqrt{2}} = x^2 - \frac{1}{3}$$

$$\|w_3\|^2 =< w_3, w_3 >= \int_{-1}^{1} (t^2 - \tfrac{1}{3})^2 dt = \int_{-1}^{1} (t^4 - \tfrac{2}{3}t^2 + \tfrac{1}{9})dt$$

$$= 2\int_{-1}^{1} (t^4 - \tfrac{2}{3}t^2 + \tfrac{1}{9})dt$$

$$= 2\int_{-1}^{1}\left[\tfrac{t^5}{5} - \tfrac{2}{3}\tfrac{t^3}{3} + \tfrac{1}{9}t\right]_0^1$$

$$\|w_3\|^2 = 2\left\{\tfrac{1}{5} - \tfrac{2}{9} + \tfrac{1}{9}\right\} = 2\left(\tfrac{9-10+5}{45}\right) = \tfrac{8}{45}$$

$$\|w_3\| = \sqrt{\frac{8}{45}}$$

$$v_3 = \frac{w_3}{\|w_3\|} = \sqrt{\frac{45}{8}}\,\frac{(3x^2-1)}{3} = \sqrt{\frac{5}{8}}(3x^2-1)$$

Hence the orthonormal basis of the subspace is $\{v_1, v_2, v_3\} = \left\{\frac{1}{\sqrt{2}}, \sqrt{\frac{3}{2}}x, \sqrt{\frac{5}{8}}(3x^2-1)\right\}$

ii) The given polynomial is $f(x) = 1 + 2x + 3x^2 \in P_2(R)$

To express $f(x)$ as a linear combination of the vectors in the orthonormal basis:

The linear combination of $f(x) = \sum_{i=1}^{3} < f, v_i > v_i$

$$< f, v_1 >= \int_{-1}^{1}(1+2t+3t^2).\frac{1}{\sqrt{2}}dt$$

$$= \frac{2}{\sqrt{2}}\int_{0}^{1}(1+3t^2)dt + \frac{2}{\sqrt{2}}\int_{-1}^{1}tdt$$

$$< f, v_1 >= \sqrt{2}\left[t + 3.\frac{t^3}{3}\right]_0^1 + 0 = \sqrt{2}(2)$$

$$< f, v_2 >= \int_{-1}^{1}(1+2t+3t^2)\sqrt{\frac{3}{2}}tdt$$

$$= \sqrt{\frac{3}{2}}\int_{-1}^{1}(t+2t^2+3t^3)dt$$

$$= \sqrt{\frac{3}{2}}.4\int_{-1}^{1} t^2 dt + \frac{\sqrt{3}}{2}\int_{-1}^{1}(t+3t^3)dt$$

$$< f, v_2 >= 2\sqrt{2}\sqrt{3}\left[\frac{(t)^3}{3}\right]_0^1 + 0 = \frac{2\sqrt{6}}{3}$$

$$< f, v_3 >= \int_{-1}^{1}(1+2t+3t^2)\sqrt{\frac{5}{8}}(3t^2-1)dt$$

$$= \sqrt{\frac{5}{8}} \times 2\int_{0}^{1}(-1+9t^4)dt$$

$$= 2\sqrt{\frac{5}{8}}\left\{[-t]_0^1 + 9.\left[\frac{t^5}{5}\right]_0^1\right\} = 2\sqrt{\frac{5}{8}}\left\{-1+\frac{9}{5}\right\}$$

$$= \frac{2}{2}\sqrt{\frac{5}{2}}\frac{(-5+9)}{5}$$

So $< f, v_3 >= \frac{2\sqrt{10}}{5}$

So $f(x) =< f, v_1 > v_1 + < f, v_2 > v_2 + < f, v_3 > v_3$

$$= 2\sqrt{2}\left(\frac{1}{\sqrt{2}}\right) + \frac{2\sqrt{6}}{3}\left(\sqrt{\frac{3}{2}}x\right) + \frac{2\sqrt{10}}{5}\left(\sqrt{\frac{5}{8}}(3x^2-1)\right)$$

Where $f(x)$ is a linear combination of vector of orthonormal basis

iii) To find the orthogonal projection of $f(x) = x^3 \in P_3(R)$ on $W = P_2(R)$:

Solution : The orthogonal projection of $f(x) = x^3$ on $W = \sum_{i=1}^{3} < f, v_i > v_i$.

$$=< f, v_1 > v_1 + < f, v_2 > v_2 + < f, v_3 > v_3$$

$$< f, v_1 >= \int_{-1}^{1} t^3 \frac{1}{\sqrt{2}} dt = 0$$

$$< f,v_2 >= \int_{-1}^{1} t^3 \sqrt{\frac{3}{2}} t dt = 2\sqrt{\frac{3}{2}} \int_{0}^{1} t^4 dt$$

So $< f,v_2 >= \sqrt{6}\left[\frac{t^5}{5}\right]_0^1 = \frac{\sqrt{6}}{5}$

$$< f,v_3 >= \int_{-1}^{1} t^3 . \sqrt{\frac{5}{8}} (3t^2 - 1) dt$$

$$= \sqrt{\frac{5}{8}} \int_{-1}^{1} (3t^5 - t^3) dt = \sqrt{\frac{5}{8}} \left[3.\frac{t^6}{6} - \frac{t^4}{4}\right]_{-1}^{1}$$

So $< f,v_3 >= 0$

Hence the orthogonal projection of $f(x) = x^3$ on $W = P_2(R)$ is $< f,v_1 > v_1 + < f,v_2 > v_2 + < f,v_3 > v_3$

$$= 0\left(\frac{1}{\sqrt{2}}\right) + \frac{\sqrt{6}}{5}\left(\sqrt{\frac{3}{2}} x\right) + 0.\frac{\sqrt{5}}{8}(3x^2 - 1)$$

$$= \frac{3}{5} x$$

W.E. 22 : Compute $S^{\perp}$ if $S = \{(1,0,i),(1,2,1)\}$ in the inner product space $C^3(C)$.

Solution: Let $S = \{u_1, u_2\}$ so that $u_1 = (1,0,i)$; $u_2 = (1,2,1)$

By definition $S^{\perp} = \{v \in C^3 / < v,u >= 0 \forall u \in S\}$

Let $u = (a,b,c) \in C^3$ where a, b, c are scalars belonging to C.

$v \in S^{\perp} \Rightarrow < v,u_1 >= 0$ and $< v,u_2 >= 0$

i.e. $< v,u_1 >= 0 \Rightarrow < (a,b,c),(1,0,i) >= 0$

$\Rightarrow a(1) + b(0) + c(\overline{i}) = 0$ where $\overline{i}$ is the congugate of i.

So $a - ci = 0$ .... (1) since $\overline{i} = -i$

and $< v, u_2 >= 0 \Rightarrow < (a,b,c),(1,2,1) >= 0$

$$\Rightarrow a(1)+b(2)+c(1)=0$$

$$\Rightarrow a+2b+c=0$$

Let $c=1$, then from (1) $a=i$

using in (2), $i+2b+1=0 \Rightarrow b=\dfrac{-(i+1)}{2}$

So $v=(a,b,c)=\left(i,\dfrac{-(1+i)}{2},1\right)$

So $S^{\perp}=\left\{i,\dfrac{-(1+i)}{2},1\right\}$

14.16 Summary:

In this lesson we discussed about orthogonality of vectors. Orthonormality. properties of orthogonality and orthonormality, Gram-Schmidt orthogonalization process to obtain orthonormal bases. Parseval's identity, Bessels inequality orthogonal complements closest vectors projection.

14.17 Technical Terms:

Orthogonality of vectors, orthonormality of vectors, orthogonalization, orthogonal compliment orthogonal projection.

14.18 Model Questions:

1. Define i) Orthogonal set ii) Orthonormal set in an inner product space.

2. If u, v, are two vectors in a real inner product space $V(F)$ such that $\|u\|=\|v\|$ then show that $(u+v)$ is orthogonal to $u-v$.

3. Show that the vectors $(-1,0),(0,-1)$ in R^2 form an orthonormal basis over R under usual inner product on R^2.

4. Prove that every orthogonal set of non zero vectors in an inner product space $V(F)$ is linearly independent.

5. State and prove Parsvel's identity.

6. Apply Gram -Schmidt process to obtain an orthonormal basis for $V_3(R)$ with the standard inner product to the vectors.

i) $(2,1,3),(1,2,3),(1,1,1)$ Ans: $\frac{1}{\sqrt{14}}(2,1,3),\frac{1}{\sqrt{42}}(4,5,1)\ ,\frac{1}{\sqrt{3}}(1,1,-1)$

ii) $(1,-1,0),(2,-1,-2),(1,-1,-2)$ Ans: $\frac{1}{\sqrt{2}}(1,-1,0),\frac{1}{3\sqrt{2}}(1,1,-4),\frac{1}{3}(-2,-2,-1)$

7. State and prove Bessel's inequality.

14.19 Exercise:

1. Apply the Gram - Schmidt process to obtain an orthonormal basis for $V_3(R)$ with the standard inner product to the vectors

i) $(2,1,3),(1,2,3),(1,1,1)$

Ans : $\frac{1}{\sqrt{14}}(2,1,3),\frac{1}{\sqrt{42}}(-4,5,1),\frac{1}{\sqrt{3}}(1,1,-1)$

ii) $(1,0,1),(1,0,-1),(0,1,4)$

Ans : $\frac{1}{\sqrt{2}}(1,0,1),\frac{1}{\sqrt{2}}(1,0,-1),(0,1,0)$

iii) $(1,-1,0),(2,-1,-2),(1,-1,-2)$

Ans : $\frac{1}{\sqrt{2}}(1,-1,0),\frac{1}{3\sqrt{2}}(1,1,-4),\frac{1}{3}(-2,-2,-1)$

2. In each part apply Gram - Schmidt process to the given subset S of the inner product space V to obtain an orthogonal basis for the span (S). Then normalise there vectors in this basis, to obtain an orthonormal basis B for span (S) and compute the fourier coefficients of the given vector relative to B.

i). $v = R^4; S = \{(2,-1,-2,4),(-2,1,-5,5),(-1,3,7,11)\}$ and $v = (-11,8,-4,18)$.

Ans: $\left\{\frac{1}{5}(2,-1,-2,4),\frac{1}{\sqrt{30}}(-4,2,-3,1),\frac{1}{\sqrt{155}}(-3,4,9,7)\right\}$

$10, 3\sqrt{30}, \sqrt{155}$

ii) $v = C^4; S = \{(-4,3-2i,i,1-4i),(-1-5i,5-4i,-3+5i,7-2i),(-27-i,-7-6i,-15+25i,-7-6i)\}$ and $v = (-13-7i,-12+3i,-39-11i,-26+5i)$.

Ans: $\left\{\frac{1}{\sqrt{47}}(-4,3-2i,i,1-4i),\frac{1}{\sqrt{60}}(3-i,-5i,-2+4i,2+i),\frac{1}{\sqrt{1160}}(-17-i,-9+8i,-18+6i,-9+8i)\right\}$

$\sqrt{47}(-1-i),\sqrt{60}(-1+2i),\sqrt{1160}(1+i)$

iii) $V=P_2(R)$ with the inner product $<f(x),g(x)>=\int_0^1 f(t)g(t)dt; S=\{1,x,x^2\}$; $h(x)=(1+x)$

Ans : $\left\{1,2\sqrt{3}(x-\frac{1}{2}),6\sqrt{5}(x^2-x+\frac{1}{6}\right\},\frac{3}{2},\frac{\sqrt{3}}{6},0$

iv) $V=M_{2\times2}(R), S=\left\{\begin{bmatrix}3&5\\-1&1\end{bmatrix},\begin{bmatrix}-1&9\\5&-1\end{bmatrix},\begin{bmatrix}7&-17\\2&-6\end{bmatrix}\right\}$ and $A=\begin{bmatrix}-1&27\\-4&8\end{bmatrix}$

Ans: $\left\{\frac{1}{6}\begin{bmatrix}3&5\\-1&1\end{bmatrix},\frac{1}{6\sqrt{2}}\begin{bmatrix}-4&4\\6&-2\end{bmatrix},\frac{1}{9\sqrt{2}}\begin{bmatrix}9&-3\\6&-6\end{bmatrix}\right\}$

$24,6\sqrt{2},-9\sqrt{2}$

3. In each of the following parts find the orthogonal projection of the given vectors on the given subspace W of the inner product space V.

i) $V=R^2;u=(2,6)$; and $W=\{(v,w):w=4v\}$

Ans : $\frac{1}{17}\begin{bmatrix}2&6\\10&4\end{bmatrix}$

ii) $V=R^3;u=(2,1,3)$ and $W=\{(u_1,u_2,u_3)/u_1+3u_2-2u_3=0\}$

Ans : $\frac{1}{14}\begin{bmatrix}29\\17\\40\end{bmatrix}$

4. Find the distance from the vector $u=(2,1,3)$ to the sub space $W=\{(u_1,u_2,u_3)/u_1+3u_2-2u_3=0\}$ of the vector space R^3.

Ans: $\frac{1}{\sqrt{14}}$

5. If W be a sub space of the inner product space $V_3(R)$ spanned by $B_1 = \{(1,0,1),(1,2,-2)\}$ then find a basis of orthogonal compliment of $W^{\perp}$.

Ans : $\{(2,-3,-2)\}$ is the basis of $W^{\perp}$

6. If $W = L(\{(1,2,3,-2),(2,4,5,-1)\})$ the subspace of $R^4(R)$; find a basis of the orthogonal compliment $W^{\perp}$.

Ans : $\left\{(2,-1,0,0),(0,\frac{-7}{2},3,1)\right\}$

7. If $V = L(S)$ with inner product $< f,g >= \int_0^{\pi} f(t)g(t)dt$ and $S = \{\sin t, \cot, 1, t\}$. Find an orthogonal basis and compute the fourier coefficients of $h(t) = 2t+1$

Ans : $\left\{\sin t, \cos t, 1-\frac{4}{\pi}\sin t, t+\frac{4}{\pi}\cos t-\frac{\pi}{2}\right\}$

14.20 Reference Books:

1. Linear Algebra 4th edition Stephen H. Friedberg, Arnold J. Insel, Lawrence E. Spence

2. Schaum's outline; Beginning Linear Albegra Seymour Lipschutz

3. Topics in Algebra I.N. Herstein

4. Linear Algebra J.N. Sharma & A.R. Vasishtha

- A. Mallikharjana Sarma

LESSON - 15

LINEAR OPERATORS

15.1 Objective of the Lesson:

We are familiar with the conjugate transpace of a matrix $A*$ of A. If A is $\left[a_{ij}\right]_{m+n,}$ then $A* = \left[b_{ij}\right]_{n+n}$ where $b_{ij} = \overline{a}_{ji}$ i.e. $A*$ is the transpace of the matrix formed with the conjugate complex numbers of the elements of A.

For a linear operator T on an inner product space V, we now define a related linear operator on V, called the adjoint of T, whose matrix representation with repect of any orthonormal basis B for V is $\left[T\right]_B^*$. The analogy between conjugate complex numbers and adjoint of a linear operator will become apparent.

As V is an inner product space in this chapter, we study the condition which guarantee that V has an orthonormal basis.

15.2. Structure of the Lesson:

This lesson contains the following items.

15.3 Introduction

15.4 Basic definitions

15.5 Theorems on linear transformations

15.6 Worked Out Examples

15.7 Adjoint of an operator - Definition

15.8 Theorems on adjoint operators

15.9 Worked out examples

15.10 Some properties of adjoint operators - theorems

15.11 Worked out Examples

15.12 Exercise

15.13 Normal and self adjoint operators definitions and theorems

15.14 T - In variance and polynomial split definitions

15.15 Schur Theorem and other theorems

15.3 Introduction:

Here we shall consider linear functionals defined on an inner product space $V(F)$. Since an inner product space is also a vector space so all concepts of linear functionals on vector spaces are also applicable to inner product space. So give some basic definitions that are useful in inner product spaces.

15.4 Basic Definitions:

1. Linear Operators : Let $V(F)$ be a vector space. A linear transformation from $V(F)$ to $V(F)$ is called a linear operator.

Also a linear functional f over $V(F)$ is a mapping i.e. $f:V \to F$ that assigns to every vector v in V, an element $f(v)$ in F; such that f is linear.

In other words $f(u+v)=f(u)+f(v)$ for every $u,v \in V$; $f(au)=af(u)$ for every $a \in F$. These two can be clubbed together as $f(au+bv)=af(u)+bf(v) \forall u,v \in V$ and $a,b \in F$

Or

$$f(au+v)=af(u)+f(v) \forall u,v \in V \text{ and } \forall a \in F$$

2. Inner Product: An inner product on V is a function f that assigns to every order pair of vectors u, v in V a scalar $<u,v>$ in $F(=R$ or C)

Also $f(au_1+bu_2,v) = <au_1+bu_2,v>$

$$= a<u_1,v> + b<u_2,v> \text{ for } u_1,u_2 \in V, a,b \in F$$

Hence f is also linear and hence f is a linear functional on V.

If $V(F)$ is a finite dimensional inner product space; then it will have an orthonormal basis.

15.5 Theorems on Linear Transformations:

15.5.1 Theorems: Let V be a finite dimensional inner product space and f is a linear transformation from $V \to F$. Then there exists a unique vector v in V such that $f(u) = < u, v > \forall u \in V$.

Proof: Let $B = \{u_1, u_2, ..., u_n\}$ is an orthonormal basis for V; and f is a linear transformation from $V \to F$.

Let $v = \sum_{j=1}^{n} \overline{f(u_j)} \quad u_j$ for each $f(\bar{u}_j) \in F$ .... (1)

$f\overline{\left[u_j\right]}$ simply denotes the conjugate of $f(u_j)$

Then $v \in V$, further more let g be a function from V to F defined by

$$g(u) = < u, v > \forall u \in V \text{ .... (2)}$$

To show that g is a linear functional on V;

Let $a, b \in F$, $w_1, w_2 \in V$, we have

of $(aw_1 + bw_2) = < aw_1 + bw_2, v >$ by (2)

$$= a < w_1, v > + b < w_2, v >$$

$$= ag(w_1) + bg(w_2)$$

Thus g is a linear functional on V.

Now we will show g = f.

Let $u_k \in B$ then $g(u_k) = < u_k, v >$ .... (3)

Now substituting the value of v, from (1), we get

$$g(u_k) = < u_k, \sum_{j=1}^{n} f\overline{(u_j)} \quad u_j >$$

$$= \overline{\sum_{j=1}^{n} f\overline{(u_j)}} < u_k, u_j >$$

$$= \sum_{j=1}^{n} f(u_j) < u_k, u_j >= f(u_k) \qquad \therefore \{< u_k, u_j >\} = \begin{cases} 0 \text{ if } j \neq k \\ 1 \text{ if } j = k \end{cases}$$

i.e. $g(u_k) = f(u_k)$

Thus f and g agree on a basis for $V(F)$ and hence $g = f$.

Inother words we say that there exists a vector $v \in V$ corresponding to the linear functional f on V :

$$f(u) =< u, v > \forall u \in V$$

Uniqueness of v:

Suppose there exists w in V such that

$$f(u) =< u, w > \forall u \in V$$

Thus $< u, v >=< u, w > \forall u \in V$

$$\Rightarrow < u, v > - < u, w >= 0 \forall u \in V$$

$\Rightarrow < u, v - w >= 0$ for all $u \in V$

$\Rightarrow < v - w, v - w >= 0$ Substituting $v - w$ for u a particular value

$$\Rightarrow v = w = 0$$

$$\Rightarrow v = w$$

So v is unique. Hence the theorem.

15.5.2 Theorem: For any linear operator T on a finite dimensional inner product space V, then there exists a unique linear operator $T*$ on V such that

$$< T(u), v >=< u, T*(v) > \forall u, v \in V$$

Proof: Let T be a linear operator on a finite dimensional inner product space V; over the field F let v be a vector in V. Let f be a function from V into F defined by $f(u) =< T(u), v > \forall u \in V$ ..... (1)

To show that f is a linear functional on V.

Let $a, b \in F; u_1, u_2 \in V$, then

$f(au + bv) =< T(au_1 + bu_2), v >$ from (1)

$=< aT(u_1)+bT(u_2), v >$ Since T is linear.

So $f < au_1 + bu_2 >= af(u_1)+bf(u_2)$ from (1)

Thus f is a linear functional on V.

So there exists a unique vector v' in V.

such that $f(u) =< u, v' > \forall u \in V$ .... (2)

Hence from (1) and (2) we observe that if T is a liear operator on V, then corresponding to every v in V, there is a uniquely determined vector v' in V such that $< T(u), v >=< u, v' >$ for all $u \in V$. Let $T*$ be the rule by which we associate v with v'. i.e. let $T*(v) = v'$.

Then $T*$ is a function from V to V defined by $< T(u), v >=< u, T*(v) \ \forall u, v \in V$ .... (3)

1. To show that $T*$ is linear:

Let $a, b \in F; v_1, v_2 \in V$. then for every u in V we have $< u, T*(av_1 + bv_2) >$

$=< T(u), (av_1 + bv_2) >$ from (3)

$= \bar{a} < T(u), v_1 > + \bar{b} < T(u), v_2 >$

$= \bar{a} < u, T*(v_1) > + \bar{b} < u, T*(v_2) >$ from (3)

$=< u, aT*(v_1) > + < u, bT*(v_2) >$

$=< u, aT*(v_1) + bT*(v_2) >$

So $T*(av_1 + bv_2) = aT*(v_1) + bT*(v_2)$

since in an inner product space $< u, v >=< u, w > \Rightarrow v = w$ for every u.

So $T*$ is a linear operator on V.

So corresponding to a linear operator T on V there exists a linear operator $T*$ on V, such that $< T(u), v >=< u, T*(v) > \forall u, v \in V$.

To show that $T*$ is unique:

Let F be a linear operator on V such that $< T(u), v >=< u, F(v) > \forall u, v \in V$

$\Rightarrow < u, T*(v) =< u, F(v) > \forall u, v \in V$

$\Rightarrow T^* = F$

So T is unique.

Hence the theorem.

Note i) The symbol T^* is red as T star.

ii) $T(u)$ can be taken as Tu and $T^*(v)$ as Tv.

iii) T is a linear operator on a finite dimensional inner product space V. If T has an eigen vector, then T^* does so.

iv) T^* is called an adjoint operator. Which we deal later in a detailed manner.

15.6 Worked Out Examples:

W.E.1:

1. For each of the inner product space $V(F)$ and linear transformation (linear functionals) $f: V \to F$ find a vector v, such that $f(u) = <u, v>$ for all $v \in V$.

i) $V = R^2; F = R; f(a_1, a_2) = 2a_1 + a_2$

ii) $V = C^2; F = C; f(z_1, z_2) = z_1 - 2z_2$

Solution: Given $V = R^2; F = R, f(a_1, a_2) = 2a_1 + a_2$ to find v.

$V \to F$ i.e. $R^2 \to R$ is an inner product space of $\dim = 2$

$V(F)$ i.e. $R_2(R)$ has an orthonormal basis $\{u_1 = (1,0), u_2 = (0,1)\}$

Such that $<u_1, u_1> = 1, <u_2, u_2> = 1, <u_1, u_2> = 0, <u_2, u_1> = 0$

$f: V \to F$ is a linear functional such that $f(u) = f(u_1, u_2) = 2u_1 + u_2$ for $u = (u_1, u_2)$

Let $u \in V$ then $u = a_1u_1 + a_2u_2$ for $a_1, a_2 \in R$ and $f(u) = f(a_1u_1 + a_2u_2) = af(u_1) + a_2 f(u_2)$ .... (1) since f is linear. We have $v \in V$ such that $f(v) = <u, v>$.

$\therefore v = b_1u_2 + b_2u_2$ for $b_1, b_2 \in R$ .... (2)

$\therefore <u, v> = <a_1u_1 + a_2u_2, b_1u_1 + b_2u_2>$

$= a_1b_2 <u_1, u_1> + a_2b_2 <u_1, u_2> + a_2b_1 <u_2, u_1> + a_2b_2 <u_2, u_2>$

$= a_1b_1 + a_2b_2$ .... (3)

Since $f(u) = <u,v>$ from (1) and (3) we get $a_1b_1 + a_2b_2 = a_1 f(u_1) + a_2 f(u_2)$

Comparing both sides of the above we get

$b_1 = f(u_1), b_2 = f(u_2)$

$\Rightarrow b_1 = f(1,0) = 2(1) + 0 = 2$

$b_2 = f(0,1) = 2(0) + 1 = 1$

Hence $v = b_1u_1 + b_2u_2 = 2(1,0) + 1(0,1) = (2,1)$ is the required vector.

ii) $V \to F$ is $C^2 \to C$ defined by $f(z_1, z_2) = z_1 - 2z_2$ to find v.

Solution: $V \to F$ in $C^2 \to C$ is a linear functional such that $f(u) = f(z_1, z_2) = z_1 - 2z_2$

Let $u \in V$, then $u = a_1z_1 + a_2z_2$ for some scalar belonging to C.

$\therefore f(u) = f(a_1z_1 + a_2z_2) = a_1 f(z_1) + a_2 f(z_2)$ ....(1) for $a_1, a_2 \in C$

$\therefore$ since f is linear.

We have $v \in V$ such that $f(u) = <u,v>$

So $v = b_1z_1 + b_2z_2$ for some $b_1, b_2 \in C$

$<u,v> = <a_1z_1 + a_2z_2, b_1z_1 + b_2z_2>$

$= a_1\overline{b_1} <z_1, z_1> + a_1\overline{b_2} <z_1, z_2> + a_2\overline{b_1} <z_2, z_1> + a_2\overline{b_2} <z_2, z_2>$

$= a_1\overline{b_1} + a_2\overline{b_2}$ .... (3)

Since $f(u) = <u,v>$ from (3) and (1)

We have $a_1\overline{b_1} + a_2\overline{b_2} = a_1 f(z_1) + a_2 f(z_2)$

Comparing on both sides we get

$\overline{b_1} = f(z_1);\quad \overline{b_2} = f(z_2)$

So $\overline{b_1} = f(1,0) = 1 - 2(0) = 1$

$\overline{b_2} = 0 - 2(1) = -2$. Hence $b_1 = 1, b_2 = -2$

Hence $v = b_1 z_1 + b_2 z_2 = 1(1,0) - 2(0,1) = (1,-2)$

Which is the required vector.

W.E.2: For each of the inner product space $V(F)$ and linear functional g: $V \to F$ find a vector v such that $g(u) = <u, v>$ for all $u \in V$.

i) $V = R^3; g(a_1, a_2, a_3) = a_1 - 2a_2 + 4a_3$

ii) $V(F) = V(C); x = (a_1, a_2, a_3); g(x) = \frac{1}{3}(a_1 + a_2 + a_3)$

Solution: (1) $V = R^3$ $g(a_1, a_2, a_3) = a_1 - 2a_2 + 4a_3$ then to find v.

$V \to F. i.e. R^3 \to R; V$ is an inner product space of $\dim = 3$.

$V(R)$ has an orthonormal basis $\{u_1 = (1,0,0), u_2 = (0,1,0), u_3 = (0,0,1)\}$

Such that $<u_i, u_j> = \begin{cases} 1 \text{ if } i = j \\ 0 \text{ if } i \neq j \end{cases}$

$g : V \to R$ i.e. $g : R^3 \to R$ is a linear functional

such that $g(u) = g(a_1, a_2, a_3) = a_1 - 2a_2 + 4a_3$ for $u = (a_1, a_2, a_3)$

Let $u \in V$, then $u = a_1u_1 + a_2u_2 + a_3u_3 = \sum_{i=1}^{3} a_i u_i$ for $a_i \in R$

So $g(u) = g(a_1u_1 + a_2u_2 + a_3u_3) = a_1 g(u_1) + a_2 g(u_2) + a_3 g(u_3)$ .... (1)

Since g is linear.

We have $v \in V$ such that $g(u) = <u, v>$

So $v = b_1u_1 + b_2u_2 + b_3u_3 = \sum_{j=1}^{3} b_j u_j$ .... (2) for $b_j \in R$

So $<u,v>=<\sum_{i=1}^{3} au_i, \sum_{j=1}^{3} b_j u_j> = a_1b_1 + a_2b_2 + a_3b_3$ .... (3)

As $g(u) = <u,v>$ from (1) and (3) we have

$$a_1b_1 + a_2b_2 + a_3b_3 = a_1 g(u_1) + a_2 g(u_2) + a_3 g(u_3)$$

comparing $b_1 = g(u_1), b_2 = g(u_2), b_3 = g(u_3)$

$\Rightarrow b_1 = g(1,0,0) = 1 - 2(0) + 4(0) = 1$

$b_2 = g(0,1,0) = 0 - 2(1) - 4(0) = -2$

$b_3 = g(0,1,0) = 0 - 2(0) + 4(1) = 4$

$\therefore v = b_1v_1 + b_2v_2 + b_3v_3$ from (2)

$= 1(1,0,0) - 2(0,1,0) + 4(0,0,1)$

$= (1,-2,4)$ which is the required vector.

ii) $g: V \to F$ i.e. $V_3 \to C$ is a linear functional $V(F)$ is an inner product space of $\dim = 3$.

V_3 has an orthonormal basis $\{u_1, u_2, u_3\}$

Such that $< u_i, u_j > = \begin{cases} 1 \text{ if } j = i \\ 0 \text{ if } j \neq i \end{cases}$

given $g: V_3 \to C$ is a linear transformation

Such that $g(u) = g(a_1, a_2, a_3) = \frac{1}{3}(a_1 + a_2 + a_3)$ for $u = (a_1, a_2, a_3)$

Let $u \in V_3$ then $u = a_1u_1 + a_2u_2 + a_3u_3 = \sum_{i=1}^{3} a_i u_i$ for $a_i \in C$.

$\therefore g(u) = g(a_1u_1 + a_2u_2 + a_3u_3) = a_1 g(u_1) + a_2 g(u_2) + a_3 g(u_3)$ .... (1)

we have $v \in V_3$ such that $g(u) = <u,v>$

$\therefore v = b_1u_1 + b_2u_2 + b_3u_3 = \sum_{J=1}^{3} b_j u_j$ ..... (2) for same $b_j \in C$

So $< u, v > = < \sum_{i=1}^{3} a_i u_i, \sum_{i=1}^{3} b_j u_j >$

$< u, v > = a_1\overline{b_1} + a_2\overline{b_2} + a_3\overline{b_3}$ .... (3) Since $< u_i, u_j > = \begin{cases} 0 \text{ if } i \neq j \\ 1 \text{ if } i = j \end{cases}$

Since $g(u) < u, v >$ from (1) and (3)

$$a_1\overline{b_1} + a_2\overline{b_2} + a_3\overline{b_3} = a_1 g_1(u_1) + a_2 g_2(u_2) + a_3 g_3(u_3)$$

Comparing $\overline{b_1} = g_1(u_1); \overline{b_2} = g_2(u_2); \overline{b_3} = g_3(u_3)$

$$\Rightarrow b_1 = \overline{g_1(u_1)}, \quad b_2 = \overline{g_2(u_2)}, \quad b_3 = \overline{g_3(u_3)}$$

If the orthonormal basis is taken as the standard basis we can have $u_1 = (1,0,0), u_2 = (0,1,0), u_3 = (0,0,1)$

We have since $g(u) = g(a_1, a_2, a_3) = \frac{1}{3}(a_1 + a_2 + a_3)$

we have $g(u_1) = g(1,0,0) = \frac{1}{3}(1+0+0) = 1/3$

$g(u_2) = g(0,1,0) = \frac{1}{3}(0+1+0) = 1/3, \quad g(u_3) = g(0,0,1) = \frac{1}{3}(0+0+1) \qquad g(u_3) = 1/3$

$\therefore b_1 = \overline{g(u_1)} = 1/3, b_2 = \overline{g(u_2)} = 1/3, b_3 = \overline{g(u_3)} = 1/3$

So $v = b_1u_1 + b_2u_2 + b_3u_3$ from (1)

$$\Rightarrow v = \frac{1}{3}(1,0,0) + \frac{1}{3}(0,0,1) = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$$

Which is the required vector.

W.E.3 : If $V_3(F)$ is an inner product space with orthonormal basis $\{u_1, u_2, u_3\}$ where $u_1 = \frac{1}{\sqrt{2}}(1,1,0)$, $u_2 = \frac{1}{\sqrt{2}}(1,-1,0)$, $u_3 = (0,0,1)$. If f is a linear functional on $V_3(F)$ such that $f(u_1) = -2, f(u_2) = 1, f(u_3) = 1$. Find the vector v such that $f(u) = <u,v> \forall u \in V_3(F)$.

Solution: $V_3 \to F$ i.e. $V_3 \to C$ is an inner product.

$V_3(F)$ is an inner product space with dimension 3.

V_3 has the orthonormal basis $\{u_1, u_2, u_3\}$ where $u_1 = \frac{1}{\sqrt{2}}(1,1,0); u_2 = \frac{1}{\sqrt{2}}(1,-1,0)$ and $u_3 = (0,0,1)$. such that

$$<u_i, u_j> = \begin{cases} 1 \text{ if } j = i \\ 0 \text{ if } j \neq i \end{cases}$$

$f: V_3 \to C$ is a linear functional such that $f(u) = f(a_1, a_2, a_3)$.

Let $u \in V$ so $u = a_1u_1 + a_2u_2 + a_3u_3 = \sum_{i=1}^{3} a_iu_i$ for all $a_i \in C$.

$f(u) = f(a_1u_1 + a_2u_2 + a_3u_3) = a_1 f(u_1) + a_2 f(u_2) + a_3 f(u_3)$

We have $v \in V$ so $f(u) = <u,v>$

$\therefore v = b_1u_1 + b_2u_2 + b_3u_3 = \sum_{j=1}^{3} b_ju_j$ for $b_j \in C$ .... (2)

$<u,v> = <\sum_{i=1}^{3} a_iu_i, \sum_{j=1}^{3} b_ju_j> = a_1\overline{b_1} + a_2\overline{b_2} + a_3\overline{b_3}$ .... (3)

Since $f(u) = <u,v>$ from (1) and (3):

$$a_1\overline{b_1} + a_2\overline{b_2} + a_3\overline{b_3} = a_1 f(u_1) + a_2 f(u_2) + a_3 f(u_3)$$

Comparing $\overline{b_1} = f(u_1), \overline{b_2} = f(u_2), \overline{b_3} = f(u_3)$

$\Rightarrow b_1 = \overline{f(u_1)}; b_2 = \overline{f(u_2)}; b_3 = \overline{f(u_3)}$

$\Rightarrow b_1 = -2, b_2 = 1, b_3 = 1$

So $v = b_1u_1 + b_2u_2 + b_3u_3 = -2\left(\frac{1}{\sqrt{2}}\right)(1,1,0) + 1.\frac{1}{\sqrt{2}}(1,-1,0) + 1(0,0,1)$

so $v = \left(\frac{-1}{\sqrt{2}}, \frac{-3}{\sqrt{2}}, 0\right)$ is the required vector.

15.7 Definition:

Adjoint of an operator: Let T be a linear operator in an inner product space V (finite dimensional or not). We say that T has an adjoint $T*$, if there exists a linear operator $T*$ on V; such that $<T(u), v> = <u, T*(v)> \forall u, v \in V$.

In theorem 15.5.2, we have proved that every linear operator on a finite dimensional inner product space possess an adjoint. But it should be noted that if V is not finite dimensional then some linear operator may possess an adjoint, while the other may not. In any case if T possess an adjoint $T*$, it is unique as we have proved in that theorem.

Note: We have $<u, T(v)> = \overline{<T(v), u>} = \overline{<v, T*(u)>} = <T*(u), v>$.

Hence $<u, T(v)> = <T*(u), v>$ for all $u, v \in V$.

15.8.1 Theorem: Let V be a finite dimensional inner product space. Let $B = \{u_1, u_2, ... u_n\}$ be an orthonormal basis for V. Let T be a linear operator on V; with respect to the ordered basis B. Then $a_{ij} = <T(u_j), u_i>$.

Proof: As B is an orthonormal basis of V; and if v is any vector in V; then $v = \sum_{i=1}^{n} <v, u_i> u_i$

Taking $T(u_j)$ in place of v, in the above, we get $T(u_j) = \sum_{i=1}^{n} <T(u_j), u_i> u_i$ ......(1)

where $j = 1, 2, ... n$.

Now if $A = \left[a_{ij}\right]_{n+n}$ be the matrix of T in the ordered basis B; then we have $T(u_j) = \sum_{i=1}^{n} a_{ij}u_i$;

$j = 1, 2, ... n$. As the expression for $T(u_j)$ as a linear combination of the vectors in B is unique and so from (1) and (2) we have $a_{ij} = < T(u_j), u_i >$ where $i = 1, 2, ...n$ and $j = 1, 2, ...n$.

15.18.2 Theorem:

Let V be a finite dimensional inner product space let T be a linear operator on V. Let B be any orthonormal basis for V. Then the matrix $T*$ is the conjugate transpose of the matrix T i.e. $\left[T^*\right]_B = [T]_B^*$

Proof: Let $B = \{u_1, u_2, ...u_n\}$ be an orthonormal basis of V. Let $A = \left[a_{ij}\right]_{n+n}$ be the matrix of T with respect to the ordered basis B. Then $a_{ij} = < T(u_j), u_i >$ .... (1)

Now $T*$ is also a linear operator on V.

Let $C = \left[c_{ij}\right]_{n+n}$ be the matrix of $T*$ in the ordered basis B. Let $c_{ij} = < T*(u_j), u_i >$ ..... (2)

Where $c_{ij} = < T*(u_j), u_i > = \overline{< u_i, T*(u_j) >}$ since $< u, v > = \overline{< v, u >}$

$= \overline{< T(u_j), u_j >}$ by def. of T

$= \overline{a_{ji}}$ by (1)

$\therefore C = \left[\overline{a_{ij}}\right]_{n+n}$; so $C = A*$ where $A*$ is the conjugate transpose of A. So $\left[T^*\right]_B = [T]_B^*$

Note: Here the basis B is the orthonormal basis but not an ordinary basis.

15.8.3 Corollary: If A and B are n x n matrices, then

(i) $(A + B)* = A* + B*$ (ii) $(CA)* = \bar{C}A*$ for all $C \in F$

(iii) $(AB)* = B*A*$ (iv) $A** = A$

(v) $I* = I$ (vi) $0_{n\times n}^* = 0_{n\times n}$ (null matrix order n)

15.9 Worked out examples:

W.E.4: Let T be a linear operator on C_2, defined by $T(a_1, a_2) = (2ia_1 + 3a_2, a_1 - a_2)$ if B is the standard ordered B as is for C^2 then find $T*(a_1, a_2)$.

Solution: We are given $T(a_1, a_2) = (2ia_1 + 3a_2, a_1 - a_2)$

The standard ordered basis is $B = \{(1, 0), (0, 1)\}$

$$T(1,0)=(2i(0)+3(0),1-0)=(2i,1)$$

$$=2i(1,0)+1(0,1)$$

$$T(1,0)=(0+3,0-1)=(3,-1)=3(1,0)-1(0,1)$$

$$[T]_B=\begin{bmatrix}2i & 3\\ 1 & -1\end{bmatrix}$$

$\left[T^*\right]_B=[T]_B^*=\begin{bmatrix}-2i & 1\\ 3 & -1\end{bmatrix}$. Hence the coordinate matrix of

$[T^*](a_1,a_2)=\begin{bmatrix}-2i & 1\\ 3 & -1\end{bmatrix}\begin{bmatrix}a_1\\ a_2\end{bmatrix}=(-2ia_1+a_2,3a_1-a_2)$ in the same basis is

$$\Rightarrow T^*(a_1,a_2)=(-2ia_1+a_2)(1,0)+(3a_1-a_2)(0,1)=(-2ia_1+a_2,3a_1-a_2)$$

W.E.5 : Let A be a n x n matrix, then show that $L_{A^*}=[L_A]^*$

Solution: Let B be the standard basis for F^n. then we have $[L_A]_B=A$. Hence

$$\left[[L_A]^*\right]_B=[L_A]_B^*=A^*=\left[L_{A^*}\right]_B$$

So $\left(L_A\right)^*=L_{A^*}$

W.E.6 : Example: If the linear transformation T on $V_3(C)$ is defined by

$T(a,b,c)=\left(2a+(1-i)b,(3+2i)a-4ic,2ia+(4-3ib-3c)\right)$ for any $(a,b,c)\in V_3(R)$ then find $T^*(a,b,c)$ with respect to standard basis.

Solution: The matrix of T relating to the standard basis of $V_3(C)$ which is also an orthonormal basis is given by

$$T=\begin{bmatrix}2 & 1-i & 0\\ 3+2i & 0 & -4i\\ 2i & 4-3i & -3\end{bmatrix}=\left[a_{ij}\right]_{3\times 3}$$

If $T*$ is the adjoint of T; then the matrix of $T*$ relative to the standard basis B is

$$[T*]=\left[\overline{a_{ji}}\right]=\begin{bmatrix} 2 & 3-2i & -2i \\ 1+i & 0 & 4+3i \\ 0 & 4i & -3 \end{bmatrix}$$

Thus showing that

$T*(a,b,c)=\left(2a+(3-2i)b-2ic,(1+i)a+(4+3i)c,4ib-3c\right)$ for each $(a,b,c)\in V_3(C)$

W.E.7: Let T is a linear operator on $V_3(F)$ defined by $T(a,b,c)=(a+b,b,a+b+c)$ for $a,b,c\in F$, then find $T*$.

Solution: Let $(x,y,z)\in V_3(F)$ and T is a linear operator on V_3. By definition.

$$<(a,b,c),T*(x,y,z)>=<T(a,b,c),(x.y,z)>$$

$=<(a+b,b,a+b+c),(x,y,z)>$

$=(a+b)\overline{x}+b\overline{y}+(a+b+c)\overline{z}$

$=a\overline{x}+b\overline{x}+b\overline{y}+a\overline{z}+b\overline{z}+c\overline{z}$

$=a(\overline{x}+\overline{z})+b(\overline{x}+\overline{y}+\overline{z})+c\overline{z}$

$=a\overline{(x+z)}+b\overline{(x+y+z)}+c\overline{z}$

$=<(a,b,c),(x+z,x+y+z,z)>$

So $T*(x,y,z)=(x+z,x+y+z,z)\,\forall\,(x,y,z)\in T*$

15.10 Some Properties of adjoint operators:

15.10.1 Theorem: Suppose S and T are linear operators on an inner product space V, and C is a scalar.

If S and T possess adjoints on V, then

i) $(S+T)*=S*+T*$ (ii) $(CT)*=\overline{C}T*$ where C is a scalar

iii) $(ST)*=T*S*$ (iv) $(T*)*=T$

v) $O^*=O$ where O is the zero operator

$I*=I$ where I is the identity operator

vi) If T is invertible then T^{-1} is also in vertible and in this case $\left(T^*\right)^{-1}=\left(T^{-1}\right)^*$

vii) $T^{**}=T$

Proof: (1) To show that $(S+T)^*=S^*+T^*$

As S and T are linear operators on V, $S+T$ is also a linear operator on V, For every $u,v\in V$.

$<u,(S+T)^*(v)>=<(S+T)(u),v>$

$=<S(u)+T(u),v>$

$=<S(u),v>+<T(u),v>$

$=<u,S^*(v)>+<u,T^*(v)>$

$=<u,S^*(v)+T^*(v)>$ by definition of adjoint.

$=<u,(S^*+T^*)(v)>$

Thus for the linear operator $S+T$ on v, there exists operator S^*+T^* on V such that

$<(S+T),(u),v>=<u,(S^*+T^*)v> \forall u,v\in V$

or $<u,(S+T)^*(v)>=<u,(S^*+T^*)v>$

By uniqueness of adjoint $(S+T)^*=S^*+T^*$

ii) To show that $(CT^*)=\bar{C}T^*$ where C is a scalar in F.

As T is a linear operator on V; so CT is also a linear operator on V; for every u, v in V, we have $<u,(CT)^*,v>=<(CT)u,v>=<CT(u),v>=C<<T(u),v>$

$=C<u,T^*(v)>$

$=<u,\bar{C}T^*(v)>$

$=<u,(\bar{C}T^*)v>$

So by the uniqueness of the adjoint we get $(CT)^*=\bar{C}T^*$.

iii) To show that $(ST)^*=T^*S^*$

As S, T are linear operators on V, so ST is also a linear operator on V. For every u, v in V.

We have $<u,(ST)^*(v)>=<(ST)(u),v>$

$=<ST(u),v>$ by definition of product of two operators

$=<T(u),S^*(v)>$ by definition of adjoint

$=<u,T^*\left(S^*(v)\right)>$

$=<u,\left(T^*S^*\right)(v)>$

i.e $<u,(ST)^*(v)>=<u,(T^*S^*)(v)>$

$\Rightarrow (ST)^*=T^*S^*$ as the adjoint operator is unique.

iv) To show that $(T^*)^*=T$

$<u,\left((T)^*\right)^*(v)>=<T^*(u),v>$

$=\overline{<v,T^*(u)>}$ (Since $<u,v>=\overline{<v,u>}$)

$=\overline{<T(v),u>}$ by definition of adjoint

$=<u,T(v)>$ Since $<u,v>=\overline{<v,u>}$)

Thus for a linear operator T^*, there exists a linear operator T on V, such that

$<u,(T^*)^*(v)>=<u,T(v)>$

Hence $(T^*)^*=T$ by the uniqueness of adjoint.

v) a) To show that $O^*=O$ where O is the zero operator

O is the zero operator in V. For every u, v in V.

We have $<u,O^*(v)>=<Ou,v>$

$=0=<u,O(v)>$

Thus $<u,O^*(v)>=<u,O(v)>$ for all $u,v\in V$

$\Rightarrow O^*=O$ (Where O is the zero operator) by uniqueness of adjoint.

(b) To show that $I^*=I$

For every $u,v\in V$ we have

$< u, I^*(v) > = < I(u), v = < u, v >$

$= < u, I(v) >$

Thus for all $u, v \in V$, $< u, I^*(v) > = < u, I(v) >$

So $I^* = I$ by uniqueness of adjoint.

vi) To show that $(T^*)^{-1} = (T^{-1})^*$

T is an invertible operator on V. So we have

$TT^{-1} = T^{-1}T = I$

$\Rightarrow (TT^{-1})^* = (T^{-1}T)^* = I^*$

$\Rightarrow (T^{-1})^* T^* = T^*(T^{-1})^* = I$ since $I^* = I$

This shows that T^* is invertible and the inverse of T^* i.e. $(T^*)^{-1} = (T^{-1})^*$

i.e. inverse of adjoint of T is adjoint of inverse of T.

vii) for $u, v \in V, < u, T(v) > = < T^*(u), v > = < u, T^{**}(v)$

$\Rightarrow T^{**} = T$ by uniqueness of adjoint.

Note: T is a linear operator on an innerproduct space V and $U_1 = T + T^*, U_2 = TT^*$

then $U_1^* = (T + T^*)^* = T^* + (T^*)^* = T^* + T \oplus T + T^* = U_1$

and $U_2^* = (TT^*)^* = (T^*)^* T^* = TT^* = U_2$

Worked Out Examples:

W.E.8: V is the inner product space R^2, T is a linear operator in V; evaluate T^* at $u = ((3, 5)$ of V where $T(a, b) = (2a + b, a - 3b)$

Solution: $(a, b) \in R^2$, Let $(a_1, b_1) \in R^2$

Then by defination of T^*

$< (a, b), T^*(a_1, b_1) > = < T(a, b), (a_1, b_1) >$

$= < (2a + b, a - 3b), (a_1, b_1) >$ by definition of T.

$= (2a+b)a_1 + (a-3b)b_1$

$= (2a_1+b_1)a + (a_1-3b_1)b$

So $<(a,b), T^*(a_1,b_1)> = <(a,b),(2a_1+b_1, a_1-3b_1)>$

as $(a,b),(a_1,b_1)$ are arbitrary elements in R^2, then we have $T^*(a_1,b_1) = (2a_1+b_1, a_1-3b_1)$

or $T^*(a,b) = (2a+b, a-3b)$

So $T^*(3,5) = (2\times 3+5, 3-3\times 5)$

i.e. $T^*(3,5) = (11,-12)$

W.E.9: Let V be the vector space $V_2(C)$ with standard inner product. Let T be the linear operator defined by $T(1,0) = (1,-2).T(0,1) = (i,-1)$. If $u = (a,b)$ then find $T^*(u)$.

Solution: Let $B = \{(1,0),(0,1)\}$. Then B is the standard basis for V. It is an orthonormal basis for V. Let us find $[T]_B$. i.e. the matrix of T in the ordered basis B.

We have $T(1,0) = (1,-2) = 1(1,0) - 2(0,1)$

$T(0,1) = (i,-1) = i(1,0) - 1(0,1)$

$$[T]_B = \begin{bmatrix} 1 & i \\ -2 & -1 \end{bmatrix}$$

The matrix of T^* in the ordered basis B is the conjugate transpace of the matrix $[T]_B$.

So $[T]_B = \begin{bmatrix} 1 & -2 \\ -i & -1 \end{bmatrix}$

Now $(a,b) = a(1,0) + b(0,1)$

The coordinate $T^*(a,b)$ in the basis B.

$$= \begin{bmatrix} 1 & -2 \\ -i & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} a & -2b \\ -ia & -b \end{bmatrix}_{2\times 1}$$

$T^*(a,b) = (a-2b)(1,0) + (-ia-b)(0,1)$

$= (a-2b, -ia-b)$

W.E.10: The inner product space V is C^2 and T is the linear operator on V, defined by

$T(z_1, z_2) = (2z_1 + iz_2, (1-i)z_1)$. Find T^* at $u = (3-i, 1+2i)$.

Solution: Let $B = \{(1,0),(0,1)\}$. Then B is the standard ordered basis for V. It is an orthonormal basis. Let us find $[T]_B$.

$T(z_1, z_2) = (2z_1 + iz_2, (1-i)z_1)$

$T(1,0) = (2,(1-i))$

$= 2(1,0) + (1-i)(0,1)$

$T(0,1) = (i,0) = i(1,0) + 0(0,1)$

$$[T]_B = \begin{bmatrix} 2 & i \\ 1-i & 0 \end{bmatrix}$$

$$[T^*]_B = \begin{bmatrix} 2 & -(1-i) \\ -1 & 0 \end{bmatrix}$$

Now $(z_1, z_2) = z_1(1,0) + z_2(0,1)$

The coordinate matrix of $T^*(z_1, z_2)$ in the basis B is $\begin{bmatrix} 2 & i-1 \\ -i & 0 \end{bmatrix}\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} 2z_1 + z_2(i-1) \\ -iz_1 + 0z_2 \end{bmatrix}$

$T^*(z_1, z_2) = (2z_1 + z_2(i-1))(1,0) - iz_1(0,1)$

$= (2z_1 + z_2(i-1), -iz_1)$

$T^*(3-i, 1+2i) = (2(3-i) + (1+2i)(i-1) - i(3-i))$

$= 6 - 2i + i - 2 - 1 - 2i$

$= ((3-3i), -(3i+1))$

W.E.11: Let T be the linear operator on $V_2(C)$, defined by $T(1,0) = (1+i, 2)$

$T(0,1) = (i, i)$ using the standard

inner product. Find the matrix T^* in the standard ordered basis. Does T commute with T^*.

Solution: Let $B=\{(1,0),(0,1)\}$. B is the standard ordered basis for T. It is orthonormal basis

We have $T(1,0)=(1+i,2)=(1+i)(1,0)+2(0,1)$

$T(0,1)=(i,i)=i(1,0)+i(0,1)$

So $[T]_B=\begin{bmatrix}1+i & i\\ 2 & i\end{bmatrix}$

So $[T^*]_B$ = The conjugate transpose of the matrix $[T]_B$.

$=\begin{bmatrix}1-i & 2\\ -i & -i\end{bmatrix}$

We have $[T]_B.[T^*]_B=\begin{bmatrix}1+i & i\\ 2 & i\end{bmatrix}\begin{bmatrix}1-i & 2\\ -i & -i\end{bmatrix}$

So $[T]_B[T^*]_B=\begin{bmatrix}3 & 3+2i\\ 3-2i & 5\end{bmatrix}$ .... (1)

Also $[T^*]_B[T]_B=\begin{bmatrix}1-i & 2\\ -i & i\end{bmatrix}\begin{bmatrix}1+i & i\\ 2 & i\end{bmatrix}$

$=\begin{bmatrix}6 & 3i+1\\ -3i+1 & 2\end{bmatrix}$ .... (2)

Now from (1) and (2) $[T]_B[T^*]_B\neq[T^*]_B[T]_B$

$\Rightarrow[TT^*]_B\neq[T^*T]_B$

So $TT^*\neq T^*T$

So T does not commute with T^*.

15.12 Exercise:

1. For each of the following inner product spaces V over F and linear transformations $g:V\to F$ find a vector v such that $g(u)=<u,v>$ for all $u\in V$.

i) $V = R^3 ; g(a_1, a_2, a_3) = a_1 - 2a_2 + 4a_3$

Ans: $v = (1, -2, 4)$

ii) $V = P_2(R)$ with $< f, h >= \int_0^1 f(t)h(t)dt$

$g(f) = f(0) + f^{'}(1)$

Ans : $v = 210u^2 - 204u + 33$

2. $V = P_1(R)$ is an inner product space T is a linear operator on V. The inner product in V is given by $< f, g >= \int_{-1}^1 f(t)g(t)dt; \quad T(f) = f^{'} + 3f$

$f(t) = 4 - 2t$ Evaluate $T*$.

Ans: $T*[f(t)] = 12 + 6t$

3. A linear operator T on $R^2(R)$ is given by $T(x, y) = (x + 2y, x - y)$ for all $x, y \in R$. If the inner product on R^2 is the standard one, find the adjoint $T*$.

Ans : $T*(x, y) = (x + y, 2x - y)$

4. Let T be a linear operator on $V_2(C)$ defined by $T(1,0) = (1,-2); T(1,0) = (i,-1)$ using the standard inner product find $T*(u)$. Where $u = (a,b)$

Ans : $(a - 2b, -ia - b)$

15.13 Normal and Self-Adjoint Operators:

15.13.1 Definition: Self Adjoint Operator: A Linear operator T on an inner product space $V(F)$ is called self adjoint operator if and only if $T = T*$.

Note: 1) A self adjoint operator is called a symmetric according as the space is called Eucledian i.e. $F = R$ ·

2. A self adjoint operator is called Hermitian when the vector space is unitary i.e. $F = C$.

3. In an inner product space if T is self adjoint operator. Then

$< T(u), v >=> u, T*(v) >=< u, T(v) > \forall u, v \in V$.

4. If $\hat{0}$ is the zero operator in V, I is the identity operator in V, then $\hat{0}^* = \hat{0}, I^* = I$. So $\hat{0}$ and I are self adjoint operators.

5. An n x n real or complex matrix A is self adjoint if $A = A^*$.

6. A self adjoint operator is also called as Hermitian operator. A self adjoint matrix is also called as Hermitian matrix.

15.13.2 Theorem:

Every linear operator T on a finite dimensional complex inner product space V can be uniquely expressed as $T^* = T_1 + iT_2$ where T_1 and T_2 are self adjoint linear operators on V.

Proof: Let $T_1 = \frac{1}{2}(T + T^*)$ and $T_2 = \frac{1}{2i}(T - T^*)$ .... (1)

then $T_1^* = \left\{\frac{1}{2}(T + T^*)\right\}^* = \frac{1}{2}\left\{T^* + (T^*)^*\right\}$

$= \frac{1}{2}(T^* + T) = \frac{1}{2}(T + T^*) = T_1$

i.e. $T_1^* = T_1$ so T_1 is self adjoint.

Also $T_2^* = \left\{\frac{1}{2i}(T - T^*)\right\}^*$

$= \left(\frac{1}{2i}\right)(T - T^*)^* = \frac{-1}{2i}(T^* - T) = T_2$ .... (2)

i.e. $T_2^* = T_2$ so T_2 is self adjoint.

From the above we get $T + T^* = 2T_1$

$$T - T^* = 2iT_2$$

Adding $2T = 2(T_1 + iT_2) \Rightarrow T = T_1 + iT_2$ .... (3)

Subtracting $2T^* = 2(T_1 - iT_2) \Rightarrow T^* = T_1 - iT_2$

Whence T_1 and T_2 are self adjoint.

Uniqueness resolution of T:

Let $T = U_1 + iU_2$ where U_1 and U_2 are self adjoint operators. Then

$$T^* = (U_1 + iU_2)^* = U_1^* + (iU_2)^*$$

$$= U_1^* + \bar{i}U_2^*$$

$$= U_1^* - iU_2^*$$

$$= U_1 - iU_2$$

$$T + T^* = (U_1 + iU_2) + (U_1 - iU_2) = 2U_1$$

$$\Rightarrow U_1 = \frac{1}{2}(T + T^*) = T_1$$

Also $T - T^* = (U_1 + iU_2) - (U_1 - iU_2) = 2U_2$

$\Rightarrow U_2 = \dfrac{1}{2i}(T - T^*) = T_2$. Hence the expression $T = T_1 + iT_2$ is unique.

15.13.3 Prove that the product of two self adjoint operators on an inner product space is self adjoint, iff they commute.

Proof: Let T_1 and T_2 be two self adjoint operators on an inner product space V;

Case i) Let the product of T_1 and T_2 i.e. T_1T_2 is self adjoint.

So $(T_1T_2)^* = T_1T_2 \Rightarrow T_2^*T_1^* = T_1T_2$

$\Rightarrow T_2T_1 = T_1T_2$ since T_1 and T_2 are self adjoint operators $T_1^* = T_1$, $T_2^* = T_2$

Thus when T_1T_2 is self adjoint the $T_1T_2 = T_2T_1$ i.e. they commute.

Case ii) Converse : Let T_1 and T_2 commute i.e. $T_1T_2 = T_2T_1$

So $(T_1T_2)^* = T_2^*T_1^* = T_2T_1$ (Since $T_1^* = T_1$ and $T_2^* = T_2$)

$= T_1T_2$ since $T_1T_2 = T_2T_1$

Thus $(T_1T_2)^* = T_1T_2$. So T_1T_2 is self adjoint.

Hence the theorem.

15.13.4 If T_1 and T_2 are self adjoint linear operators on an inner product space, then show that $T_1 + T_2$ is self adjoint.

Solution: T_1 is a self adjoint operator. So $T_1^* = T_1$

and T_2 is a self adjoint operator. So $T_2^* = T_2$

Now $(T_1 + T_2)^* = T_1^* + T_2^* = T_1 + T_2$

So $(T_1 + T_2)^* = T_1 + T_2$, Hence $T_1 + T_2$ is self adjoint.

15.13.5 If T is a self adjoint linear transformation on an inner product space then $T = \hat{0} \Leftrightarrow < T(u), u >= 0$ for all $u \in V$.

Proof: Let $T = \hat{0}$ (Zero operator)

$< T(u), u >=< \hat{0}(u), u >=< 0, u >= 0$ for all $u \in V$.

Conversely Let $< T(u), u >= 0$, $u \in V$, then to show that $T = \hat{0}$.

Consider $< T(u+v), u+v >= 0$ by given condition

$\Rightarrow < T(u) + T(v), u+v >= 0$

$\Rightarrow < T(u), u > + < T(u), v > + < T(v), u > + < T(v), v >= 0$

$\Rightarrow < T(u), v > + < T(u), v >= 0$ by given condition

$\Rightarrow < T(u), v > + < v, T^*(u) >= 0$ ...... (1) Since $< T(v), u >=< v, T^*(u) >$

$\Rightarrow < T(u), v > + < v, T(u) >= 0$ .... (2) Since $T^* = T$

Case i) Let V be a complex inner product space then by (1) replacing v by iv we get

$< T(u), iv > + < T(iv), u >= 0$

$\Rightarrow -i < T(u), v > + i < T(v), u >= 0$ since conjugate of i is -i.

$\Rightarrow - < T(u), v > + < T(v), u >= 0$ .... (3)

Adding (1) and (3) we get

$2 < T(u), v >= 0$ for all $u, v \in V$

So $< T(u), T(v) >= 0$ putting $v = T(u)$

$\Rightarrow T(u) = 0$

$\Rightarrow T = \hat{0}$

Case ii) In case if V is real inner product space we have $< v, T(u) >=< T(u), v >$ since $< u, v >=< v, u >$.

From (2) $< T(u), v > + < v, T(u) >= 0$

$\Rightarrow 2 < T(u), v >= 0$ using above, $\forall u, v \in V$

$\Rightarrow < T(u), T(u) >= 0$ putting $v = T(u)$

$\Rightarrow < T(u) = 0, \forall u, v \in V$ So $T = \hat{0}$

Hence the theorem.

15.13.6 If T is a linear transformation on a complex inner product space, then T is self adjoint $\Leftrightarrow < T(u), v >$ is real $\forall u, v \in V$.

Proof: Case i) Let T is self adjoint. So $T^* = T$.

then $\forall u, \in V$, $< T(u), u >=< u, T^*(u) >=< u, T(u) >$

$\therefore < T(u), u >= \overline{< T(u), u >} \Rightarrow < T(u), u >$ is real.

Case ii) Converse :

Let $< T(u), u >$ is real for each $u \in V$. Then to prove that T is a self adjoint transformation i.e. to show $< T(u), v >=< u, T(v) > \forall u, u \in V$.

Now $< T(u+v), u+v >=< T(u) + T(v), u+v >$

$=< T(u), u > + < T(u), v > + < T(v), u > + < T(v), v >$

Since by hypothesis $< T(u+v), u+v >, < T(v), u >$ and $< T(v), v >$ are all real and so $< T(u), v > + < T(u), u >$ is also real. So equating it to its complex conjugate.

We get $< T(u), v > + < T(v), u >= \overline{(< T(u), v > + < T(v), u >)}$

$= \overline{< T(u), v >} + \overline{< T(v), u >)}$

$=< v, T(u) > + < u, T(v) >$

Thus for all $u, v \in V$.

$$< T(u), v > + < T(v), u > = < v, T(u) > + < u, T(v) > \text{ .... (1)}$$

replacing v by iv we get

$$< T(u), iv > + < T(iv), u > = < iv, T(u) > + < u, T(iv) >$$

$$\Rightarrow -i < T(u), v > + < iT(v), u > = i < v, T(u) > + < u, iT(v) >$$

$$\Rightarrow -i < T(u), v > + i < T(v), u > = i < v, T(u) > - i < u, T(v) > \text{ .... (2)}$$

Multiplying (2) by i and adding to (1) we get

$$2 < T(u), v > = 2 < u, T(v) >$$

$\Rightarrow T$ is self adjoint operator.

Hence the theorem.

15.13.7 Theorem:

If T is a self adjoint linear operator on an inner product space V, then if $T \neq \hat{0}$ and $a \neq 0$, then aT is self adjoint if and only if a is real.

Proof: Let a be real. Then we have $(aT)^* = \overline{a}T^*$

$= aT$ since a is real and T is self adjoint.

Thus $(aT)^* = aT$. So it follows aT is self adjoint.

Converse: Let aT is self adjoint So $(aT)^* = aT$

$$\Rightarrow \overline{a}T^* = aT$$

$\Rightarrow \overline{a}T = aT$ since T is self adjoint.

$$\Rightarrow (\overline{a} - a)T = \hat{0}$$

as $T \neq \hat{0}$, so $\overline{a} - a = 0 \Rightarrow \overline{a} = a$

Hence a is real.

Hence the theorem.

15.13.8 Theorem:

Let T be a linear operator on a finite dimensional inner product space V. Then T is self adjoint if and only if its matrix in every orthonormal basis is a self adjoint matrix.

Proof: Let B be any orthonormal basis of V. Then $\left[T^*\right]_B = \left[T\right]_B^*$ .... (1)

If T is self adjoint, then $T^{*} = T$.

So from (1) we get $[T]_B = [T]_B^{*}$

So $[T]_B$ is a self adjoint matrix.

Conversely Let $[T]_B$ be a self adjoint matrix then $[T]_B = [T]_B^{*} = [T^{*}]_B$ from (1)

So $T = T^*$. Hence T is self adjoint.

15.13.9 Theorem:

If T is self adjoint linear operator on a finite dimensional inner product space, then prove that det (T) is real.

Proof: Let B be any orthonormal basis for V.

Then we have $[T^{*}]_B = [T]_B^{*}$ and T is self adjoint. So $T^{*} = T$ i.e. $[T]_B = [T]_B^{*}$

Suppose $[T]_B = A$ and then $[T]_B^{*} = A^*$

Then $A = A^* \Rightarrow \det(A) = \det(A^*)$

$\Rightarrow \overline{(\det A)} = \det A$

So det A is real. Hence $\det T$ is real. Since $\det T = \det[T]_B = \det A$

Hence the theorem.

15.13.10 Theorem:

Let T be a self adjoint linear operator on a finite dimensional inner product space.

Prove that the range of T is orthogonal complement of the null space of T i.e. $R(T) = \{N(T)\}^{\perp}$

Proof: Let u be any element in $R(T)$. Then there exists a vector v in V such that $u = T(v)$

Now if w$\in N(T)$ then $T(w) = O$

We have $< u, w > = < T(v), w > = < v, T^*(w) >$

$= < v, T(w) >$ since $T^* = T$

$$=< v, O >= 0$$

it follows that $< u, w >= 0$ for all $w \in N(T)$

Thus $u \in [N(T)]^{\perp}$

Thus $u \in R(T) \Rightarrow u \in [N(T)]^{\perp}$ and hence

$$R(T) \subseteq N(T)^{\perp} \text{ .... (1)}$$

Again $\dim[R(T)] + \dim[N(T)] = \dim V$

and $V = N(T) \oplus [N(T)]^{\perp}$

$$\Rightarrow \dim(N(T)) + \dim[N(T)]^{\perp} = \dim V$$

$$\therefore \dim R(T) = \dim[N(T)]^{\perp}$$

Now $R(T) \subseteq [N(T)]^{\perp}$ and $\dim[R(T)] = [N(T)]^{\perp}$

So $R(T) = [N(T)]^{\perp}$

15.13.11 Let T be a linear operator on a finite dimensional inner product space V. If T has an eigen vector then show that $T*$ has an eigen vector.

Proof: Let u be an eigen vector of T; with eigen value λ. Then for any $v \in V$, we have

$$0 =< O, v >=< (T - \lambda I)(u), v >$$

$$=< u, (T - \lambda I)*(u) >$$

$$=< u, (T* - \bar{\lambda} I)*(v) >$$

Hence u is orthogonal to the range of $T* - \bar{\lambda} I$.

So $T* - \bar{\lambda} I$ is not onto and hence is not one to one. Thus $T* - \lambda I$ has a non zero null space, and any non zero vector in this null space is an eigen vector of $T*$ with corresponding eigen value $\bar{\lambda}$.

15.14 Some Basic Concepts which are already discussed:

15.14.1 T Invariance:

W is a subspace of a vector space V and $T : V \to V$ is linear. W is said to be T - invariant if $T(w) \in W$ for every $w \in W$ that is $T(W) \subseteq W$. If W is T - invariant. We define the restriction of T on W to be the function $T_W : W \to W$ defined by $T_W(w) = T(w)$ for all $w \in W$.

15.14.2 Definition: Polynomial split:

A polynomial $f(t)$ in $P(F)$ is said to split over F, if there exists scalars $c, a_1, a_2, ... a_n$ not necessarily distinct in F such that $f(t) = C(t-a_1)(t-a_2),...(t-a_n)$.

Thus a polynomial is said to split if it factors into linear factors.

15.14.3 Theorem:

Let V be a finite dimensional inner product space. Let T be a linear operator on V. suppose W is a subspace of V, which is invariant under T. Then show that the orthogonal compliment is invariant under $T*$.

Solution: We are given that W is invariant under T. We have to prove that $W^{\perp}$ is invariant under $T*$. Let v be an vector in $W^{\perp}$. Then to prove that $T*(v)$ is in $W^{\perp}$. i.e. $T*(v)$ is orthogonal to every vector in W. Let u be any vector in W. then $< u, T*(v) > = < T(u), v > = 0$

Since $u \in W \Rightarrow T(u) \in W$ as W is T invariant.

Also v is orthogonal to every vector in W.

$T*(v)$ is orthogonal to every vector u in W.

$\therefore T*(v)$ is in $W^{\perp}$.

So $W^{\perp}$ is invariant under $T*$.

15.15.1 Schur Theorem: Let T be a linear operator on a finite dimensional inner product space V. Suppose that the characteristic polynomial of T splits. Then show that there exists an orthonormal basis B for V such that the matrix $[T]_B$ is upper triangular.

Proof: The proof is by mathematical induction on the dimension n of V. The result is immediate if $n = 1$. so suppose that the result is true for linear operators on $(n-1)$ dimensional inner product spaces whose characteristic polynomial splits. We know that if T is a linear operator on a finite dimensional inner product space V, and if T has an eigen vector, then $T*$ will have an eigen vector. so we can assume that $T*$ has a unit eigen vector w. Suppose that $T*(W) = \lambda W$ and that W = span $(\{w\})$ we will now show that $W^{\perp}$ is is T invariant.

If $v \in W^{\perp}$ and $u = cw \in W$ then

$$< T(v), u > = < T(v), cw > = < v, T*(cw) >$$

$$=< v, cT^*(w) >=< v, c\lambda w >$$

$$= \overline{c\lambda} < v, w >= \overline{c\lambda}(0) = 0$$

So $T(u) \in W^{\perp}$

We know that if T is a linear operator on a finite dimensional vector space V and W is a T-invariant subspace of V, then the characteristic polynomial of T_W divides the characteristic polynomial of T.

By this theorem, the characteristic polynomial of T_W divides the characteristic polynomials of T and hence splits.

We know that if W is any subspace of a finite dimensional inner product space V, then $\dim(V) = \dim(W) + \dim(W^{\perp})$

So $\dim(w^{\perp}) = n-1$ so we may apply the induction hypothesis to $T_{w^{\perp}}$ and obtain an orthonormal basis B' of $W^{\perp}$ such that $\left(T_{w^{\perp}}\right)_{B'}$ is upper triangular.

Clearly $B = B' \cup \{w\}$ is an orthonormal basis for V such that $[T]_B$ is upper triangular.

Hence the theorem.

15.15.2 Theorem: Let T be a self adjoint operator on a finite dimensional real inner product space V. Then show that the characteristic polynomial of T splits.

Proof: Let $\dim(V) = n$. Let B be an orthonormal basis for V and $A = [T]_B$. Then A is self adjoint. Let T_A be a linear operator on C^n. defined by T_A (u) = Au for all $u \in C^n$. T_A is self adjoint because $[T_A]_D = A$ where D is the standard ordered orthonormal basis for C^n. As T_A a self adjoint operator, the eigen values of T_A are real.

By fundamental theorem of Algebra, the characteristic poly nomial splits into factors of the form $t-\lambda$. Since each λ is real the characteristic polynomial splits over R. But T_A has the same characteristic polynomial as A; which has the same polynomial as T. So the characteristic polynomial of T splits.

Hence the theorem.

Note: Fundamental theorem of Algebra.

Suppose the $P(z) = a_n z^n + a_{n-1} z^{n-1} + ... + a_1 z + a_0$ is a polynomial in $P(C)$ of degree $n \geq 1$, then $P(z)$ has a zero and then exists complex numbers $c_1, c_2, ...c_n$ not necessarily distinct such that $P(z) = a_n(z-c_1)(z-c_2)...(z-c_n)$.

15.15.3 Theorem: Let V be a finite dimensional inner product space and let T be a self adjoint linear operator on V. Then there is an orthonormal basis for V, each vector of which is a characteristic vector for T and consequently the matrix of T with respect to B is a diagonal matrix.

Proof: As T is a self adjoint linear operator on a finite dimensional inner product space V, so T must have a characteristic value and T must have a characterisitc vector.

Let $O \neq u$ be a characteristic vector for T. Let $u_1 = \frac{u}{\|u\|}$. Then u_1 is a characterisitic vector for T and $\|u\| = 1$. If dim $V = 1$, then $\{u_1\}$ is an orthonormal basis for V, and u_1 is a characteristic vector for T. Thus the theorem is true if dim V = 1. Now we proceed by induction on the dimension of V. Suppose the theroem is true for inner product spaces of dimension less than dimension of V. Then we shall prove that it is true for n and the proof of will be complete by induction.

Let W be the one dimensional subspace of V spanned by the characteristic vector u_1 for T. Let u_1 be the characteristic vector corresponding to the characteristic value C. Then $T(u_1) = C(u_1)$. If v is any vector in W. Then $v = ku_1$, where k is a scalar. We have $T(v) = T(ku_1) = kT(u_1) = k(cu_1) = kc(u_1)$. So $T(u) \in W$. W is invariant under T. So $W^{\perp}$ is invariant under $T*$. But T is self adjoint means $T = T*$. So $W^{\perp}$ is invariant under T. If dim $V = n$, then dim $W^{\perp} = \dim V - \dim W = n-1$

So $W^{\perp}$ with the the inner product from V is an inner product space of dimension one less than the dimension of V.

suppose S is the linear operator induced by T on $W^{\perp}$ i.e. S is the restriction of T to $W^{\perp}$. Then $S(w) = T(w) \forall w \in W^{\perp}$. Then restriction of $T*$ to $W^{\perp}$ will be the adjoint of $S*$ of S. Thus S is a self adjoint linear joperator on $W^{\perp}$, because if w is any vector in $W^{\perp}$ then $S*(w) = T*(w) = T(w) = S(w)$

$$\therefore S* = S.$$

Thus S is a self adjoint linear operator on $W^{\perp}$; Whose dimension is less than dimension of V. so by our induction hypothesis $W^{\perp}$ has an orthonormal basis $\{u_1, u_2, ...u_n\}$. Consisting of characteristic vectors for S. Suppose u_i is the characteristic vector for S corresponding to the characteristic value c_i. Then $S(u_i) = C_i u_i$

$$\Rightarrow T(u_i) = C_i u_i$$

So u_i is also a characteristic vector of T. Thus $u_1, u_2, ...u_n$ are also characteristic vectors for T. since $V = W \oplus W^{\perp}$. So $B = \{u_1, u_2, ...u_n\}$ is an orthonormal basis for V each vector of which is a characteristic vector of T. The matrix T relative to B will be a diagonal matrix.

15.16 Normal Operator:

Definition: Let T be a linear operator on an inner product space V. Then T is said to be normal if it commutes with its adjoint i.e. if $TT^* = T^*T$.

Note i) If V is finite dimensional then T^* will definitely exist. If V is not finite dimensional, then the above definition will make sense if and only if T possess adjoint.

ii) Every self adjoint operator is normal.

Suppose T is self adnoint operator, then $T^* = T$, so obviously $T^*T = TT^*$.

So T is normal.

W.E.12 Examples: If $T: R^2 \to R^2$ be rotation by θ, where $0 < \theta < \pi$, and if, for standard basis B.

$[T]_B = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ show that T is normal.

Solution: $[T]_B = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}; [T^*]_B = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$

Now $[T]_B \cdot [T^*]_B = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

and $[T^*]_B \cdot [T]_B = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

As $[T]_B [T^*]_B = [T^*]_B [T]_B \Rightarrow [TT^*]_B [T^*T]_B$

So T is normal.

Normal Matrix : Definition:

A real or complex n x n matrix A is normal if and only if it commutes with its conjugate transpose. i.e. $AA^* = A^*A$.

Example : $A = \begin{bmatrix} 1 & 1 \\ i & 3+2i \end{bmatrix}$ then $A^* = \begin{bmatrix} 1 & -i \\ 1 & 3-2i \end{bmatrix}$

and $AA^* = \begin{bmatrix} 1 & 1 \\ i & 3+2i \end{bmatrix} \begin{bmatrix} 1 & -i \\ 1 & 3-2i \end{bmatrix}$

$$= \begin{bmatrix} 2 & 3-3i \\ 3+3i & 14 \end{bmatrix}$$

$$A^*A = \begin{bmatrix} 1 & -i \\ 1 & 3-2i \end{bmatrix} \begin{bmatrix} 1 & 1 \\ i & 3+2i \end{bmatrix} = \begin{bmatrix} 2 & 3-3i \\ 3+3i & 14 \end{bmatrix}$$

Thus $AA^* = A^*A$. So A is normal.

15.17 Theorems on Normal Operators:

15.17.1 An operator T on an inner product space is normal

$$\Leftrightarrow \|T^*(u)\| = \|T(u)\| \ \forall u \in V$$

Proof: For all $u \in V$, we have

$$\|T^*(u)\| = \|T(u)\| \Leftrightarrow \|T^*(u)\|^2 = \|T(u)\|^2$$

$$\Leftrightarrow < T^*(u), T^*(u) > = < T(u), T(u) >$$

$$\Leftrightarrow < TT^*(u), u > = < T^*T(u), u >$$

$$\Leftrightarrow < (TT^* - T^*T)u, u > = 0$$

$\Leftrightarrow TT^* - T^*T = \hat{0}$ (zero operator)

$$\Leftrightarrow TT^* = T^*T$$

$\Leftrightarrow T$ is normal.

15.17.2 Theorem:

If T is a normal operator on an inner product space $V(F)$, then $T - CI$ is normal for every $C \in F$.

Proof: $(T - CI)^* = T^* - (CI)^*$

$$= T^* - \bar{C}I^*$$

$= T^* - \bar{C}I$ .... (1) since $I^* = I$

Again $(T - CI)^*(T - CI)$

$= (T^* - \bar{C}I)(T - CI)$ using (1)

$= T^*T - CT^* - \bar{C}T + C\bar{C}I \quad (\therefore IT = TI = T)$

$= T^*T - \bar{C}T + -CT^* + C\bar{C}I \quad (\therefore TT^* = T^*T)$

$= (T - CI)(T - CI)^*$

As $(T - CI)^*(T - CI) = (T - CI)(T - CI)^*$ it follows that $T - CI$ is normal.

15.17.3 Theorem:

Let T be a normal operator on an inner product space V. Then a necessary and sufficient condition that u be a characteristic vector of T is that it be a characteristic vector of T^*.

Proof: T is a normal operator on an inner product space V. So $TT^* = T^*T$

$\Rightarrow < (TT^*)(u), u > = < (T^*T)(u), u >$ for any $u \in V$

$\Rightarrow < T(T^*)(u), u > = < T^*(T(u), u >$

$\Rightarrow < T^*(u), T^*(u) > = < T(u), (T^*)^*(u) >$

$\Rightarrow \|T^*(u)\|^2 = < T(u), T(u) >$ since $T^{**} = T$.

$\Rightarrow \|T^*(u)\|^2 = \|T(u)\|^2$

$\Rightarrow \|T^*(u)\| = \|T(u)\|$ ie $\|T(u)\| = \|T^*(u)\|$ .... (1)

We know that if T is normal, C is a scalar then $T - CI$ is normal.

Hence from (1) for all $u \in V$.

$\|(T - CI)(u)\| = \|(T - CI)^*(u)\|$

$\Rightarrow \|(T - CI)u\| = \|(T^* - \bar{C}I)u\|$

In particular $(T - CI)u = 0$ if and only if $(T^* - \bar{C}I)(u) = 0$ i.e. $T(u) = Cu$, if and only if $T^*(u) - \bar{C}u$. Hence u is the characteristic vector for T with characteristic value C if and only if u is a characteristic value for T^* with characteristic value $\bar{C}$.

Result: If u is an eigen vector of T, then u is also an eigen vector of T^*. Infact if $T(u) = Cu$; then $T^*(u) = \bar{C}u$.

15.17.4 Theorem:

Let T be a normal operator on an inner product space. If $u \in V$, then $T(u) = O \Leftrightarrow T^*(u) = O$

Proof: T is normal $\Leftrightarrow TT^* = T^*T$

Also $\|T(u)\|^2 = < T(u), T(u) >$

$= < u, T^*T(u) >$

$= < u, (T^*T)u >$

$= < u, (TT^*)u >$

$< T^*(u), T^*(u) >$

$= \|T^*(u)\|^2$

So $\|T(u)\| = \|T^*(u)\|$

Now $T(u) = O \Leftrightarrow \|T(u)\| = 0 \Leftrightarrow \|T^*(u)\| = 0$

$\Leftrightarrow T^*(u) = O$

Thus $T(u) = O \Leftrightarrow T^*(u) = O$

15.17.5 Let V be an inner product space. Let T be a normal operator on V. If λ_1, λ_2 are distinct eigen values of T with corresponding eigen vectors, u_1 and u_2 then show that u_1 and u_2 are orthogonal.

Proof: Let u_1, u_2 are the characteristic vectors of T corresponding to the characteristic values $\lambda_1, \lambda_2 (\lambda_1 \neq \lambda_2)$. Now $Tu_1 = \lambda_1 u_1, Tu_2 = \lambda_2 u_2$ then $T^* u_2 = \bar{\lambda} u_2$.

Again $\lambda_1 < u_1, u_2 > = < \lambda_1 u_1, u_2 >$

$= < Tu_1, u_2 >$

$= < u_1, T^* u_2 >$

$= < u_1, \bar{\lambda}_2 u_2 >$

or $\lambda_1 < u_1, u_2 > = \lambda_2 < u_1, u_2 >$

$\Rightarrow (\lambda_1 - \lambda_2) < u_1, u_2 >= 0$

$\Rightarrow < u_1, u_2 >= 0$ since $\lambda_1 \neq \lambda_2$

$\Rightarrow u_1, u_2$ are orthogonal.

Hence the theorem.

15.17.6 Let V be a finite dimensional complex inner product space and let T be a normal operator on V. Then V has an orthonormal basis B, each vector of which is a characteristic Vector for T, and conrequently the matrix of T with respect to B is a diagonal matrix.

Proof: As T is a linear operator on a finite dimensional complex inner product space V, so T must have a characteristic value and so T must value a characteristic vector.

Let $O \neq u$ be a characteristic vector of T.

Let $u_1 = \frac{u}{\|u\|}$. Then u_1 is also a characteristic, vector for T since $\|u\| = 1$. If $\dim V = 1$, then $\{u_1\}$ is an orthonormal basis for V. and u_1 is a characteristic vector for T.

Thus the theorem is true if $\dim V = 1$.

Now we proceed by induction on the dimension of V. We suppose that the theorem is true for inner product spaces of dimension les than $\dim V$. Then we shall prove that it is true for V and the proof is complete by induction.

Let W be the one dimensional subspace of V spanned by the characteristic vector u_1 for T. Let u_1 be the characteristic vector corresponding to the characteristic value C. Then $T(u_1) = Cu_1$. If v is any vector in W then $v = Ku_1$ where K is some scalar.

We have $T(v) = T(Ku_1) = KT(u_1)$

$= K(Cu_1) = (KC)u_1$

So $T(v) \in W$. Thus W is invariant under T and So $W^{\perp}$ is invariant under $T*$.

Now T is normal. So if u_1 is a characterisitc vector of T, then u_1 is also a characteristic vector of $T*$.

So by the same argument as above, W is also invariant under $T*$. So $W^{\perp}$ is invariant under $(T*)*$ i.e. $W^{\perp}$ is invariant under T. If dim $V = n$, then dim $W^{\perp} = \dim V - \dim W = n - 1$.

So $W^{\perp}$ with the inner product from V is a complex inner product space of dimension less than dimenstion of V.

Suppose S is the linear operator induced by T on $W^{\perp}$ i.e. S is the restriction of T on $W^{\perp}$.

Then $S(v) = T(v) \forall v \in W^{\perp}$

This restriction of $T*$ to $W^{\perp}$ will be the adjoint of $S*$ of S. Now S is a normal operator on $W^{\perp}$. For v is any vector in $W^{\perp}$

then $(SS*)(v) = S\left(S*(v)\right) = S\left(T*(v)\right)$

$$= T\left(T*(v)\right)$$

$$= (TT*)(v)$$

$$= (T*T)v = T*\left[T(v)\right]$$

$$= T*\left[S(v)\right]$$

$$= S*\left[S(v)\right]$$

$$= (S*S)(v)$$

So $(SS*) = (S*S)$ and thus S is a normal operator on $W^{\perp}$; whose dimension is less than the dimension of V.

So by our induction hypothesis, $W^{\perp}$ has an orthonormal basis $\{u_1, u_2, ..., u_n\}$ consisting of characteristic vectors of S. Suppose u_i is the characteristic vector for S corresponding to the characteristic value C_i.

Then $S(u_i) = Cu_i \Rightarrow T(u_i) = Cu_i$

So u_i is also a characteristic vector for T.

Thus $u_2, u_3, ...u_n$ are also characteristic vectors for T. Since $V = W \oplus W^{\perp}$, so $B = \{u_1, u_2, ...u_n\}$ is also an orthornormal basis for T each vector of which is a characteristic vector for T. The matrix of T relative to B will be the diagonal matrix. Hence the theorem.

15.15.7 Theorem:

Suppose T is a linear operator on a finite dimensional inner product space V and suppose that there exists an orthonormal basis $B=\{u_1,u_2,...u_n\}$ for V such that each vector in B is a characteristic vector for T. Then prove that T is normal.

Proof: If $u_i \in B$, then it is given that u_i is a characteristic root of T. So

Let $T(u_i)=C_iu_i$ for $i=1,2,..n$

Then $[T^*]_B$ is a diagonal matrix with diagonal elements $c_1,c_2,...c_n$. Since

$[T^*]_B=[T]_B$ so $[T^*]_B$ is also a diagonal matrix with diagonal elements $\bar{c}_1,\bar{c}_2,...\bar{c}_n$. Now as two diagonal matrices commute it follows

$$[T]_B=[T^*]_B=[T^*]_B[T]_B$$

$$\Rightarrow [TT^*]_B=[T^*T]_B$$

$\Rightarrow TT^*=T^*T \Rightarrow T$ is normal.

Hence the theorem.

Note: The above two theorems can be clubbed together and can be restated as.

Let T be a linear operator on a finite dimentional complex inner product space V. Then T is normal if and only if there exists an orthonormal basis of V consisting of eigen vectors of T.

15.18 Positive Definite and Positive Semidefinite Transformations:

15.18.1 1) Positive definite operator: Definition:

A linear operator T on a finite dimensional inner product space V is called positive definite in symbol $T>0,$ if T is self adjoint and $<T(u),u>0$ for all $O\neq u\in V$.

ii) Positive semi definite operator : A linear operator on an inner product space V is called postive semi definite (or non negative) in symbol $T\geq 0$ if it is self adjoint and if $<T(u),u> \ \geq 0 \forall u$ in V.

Note i) An n x n matrix A with entries from R or C is called positive definite if L_A is positive definite.

ii) An n x n matrix A with entries from R or C is called postive semidefinite if L_A is positive semi definite.

15.8.3 Theorem: Let T be a self adjoint operator on an inner product space V. If T is positive or non negative then every characteristic value of T is positive or non negative respectively.

Proof: Let C be a characteristic value of T.

Then $T(u)=Cu$ for some non zero vector u.

We have $<T(u),u>=<Cu,u>=C<u,u>$

$$= C\|u\|^2$$

$$\Rightarrow C = \frac{<T(u), u>}{\|u\|^2}$$

If T is postive then $<T(u), u> > 0$, so $C > 0$ i.e. C is postive.

If T is non negative then $<T(u), u> \geq 0$

So $C \geq 0$ i.e. C is non negative.

15.18.4 Theorem:

If T is a self adjoint operator on a finite dimensional inner product space V, such that the characteristic values of T are non-negative show that T is non-negative.

Proof: T is a self adjoint operator on a finite dimensional inner product space V. Let T has all charcterisitc values non-negative.

As T is self adjoint, we can find an orthonormal basis $B = \{v_1, v_2, v_3, ..., v_n\}$ consisting of characteristic vectors of T.

For each v_1 we have $T(v_i) = C_i V_i$ where $C_i \geq 0$. Let w be any vector in V

Let $w = a_1 u_1 + a_2 u_2 + ... + a_n u_n$

Then $T(w) = T(a_1 v_1 + a_2 v_2 + ... + a_n v_n)$

$$= a_1 T(v_1) + a_2 T(v_2) + ... + a_n T(v_n)$$

$$= a_1 c_1 v_1 + a_2 c_2 v_2 + ... + a_n c_n v_n$$

We have $<T(w), w> =< a_1 c_1 v_1 + a_2 c_2 v_2 + ... + a_n c_n v_n, a_1 v_1 + a_2 v_2 + ... + a_n v_n >$

$= a_1 \overline{a}_1 c_1 + a_2 \overline{a}_2 c_2 + ... + a_n \overline{a}_n c_n$ $(\because \{v_1, v_2, v_3, ..., v_n\}$ is an orthonormal)

$$= |a_1|^2 c_1 + |a_2|^2 c_2 + ... + |a_n|^2 c_n. \geq 0$$

since $c_i \geq 0$ and $|a_i| \geq 0$

Thus $(T(w), w> \geq 0 \forall w \in V$

Hence $T \geq O$ i.e. T is non negative.

15.18.5 Theory:

Let T be a linear operator on a finite dimensional inner product space V. Let $A=\left[a_{ij}\right]_{n\times n}$ be the matrix of T relative to an ordered orthonormal basis $B=\{u_1,u_2,...,u_n\}$. Then T is positive if and only if the matrix A satisfies the following conditions.

i) $A = A*$ i.e. A is self adjoint

ii) $\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}\overline{x_i}x_j > 0$ where $x_1, x_2, ..., x_n$ and n scalars not all zero.

Proof: Let v be any vector in V. Then

$$v = x_1u_1 + x_2u_2 + ...x_nu_n$$

then $< T(v), v >=< \left(T\sum_{j=1}^{n} x_ju_j\right), \sum_{i=1}^{n} x_iu_i >$

$$=< \sum_{j=1}^{n} x_jT(u_j), \sum_{i=1}^{n} x_iu_i >$$

$$= \sum_{j=1}^{n}\sum_{i=1}^{n} x_j\overline{x_i} < T(u_j), u_i >$$

We know if $A=\left[a_{ij}\right]_{n\times n}$ be the matrix of T with respect to the ordered basis B then $a_{ij} =< T(u_j), u_i >$ using in the above we get $< T(v), v >= \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}\overline{x_i}x_j$

Now suppose T is positive. Then $T = T*$

So $A = A*$

If $x_1, x_2, ...x_n$ are any n scalars not all zero, then $v = x_1u_1 + x_2u_2, ... + x_nu_n$ is a non zero vector in V. Since T is positive $< T(v), v > \; > 0$ Hence $\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}\overline{x_i}x_j > 0$

conversely suppose that the conditions (1) and (ii) of the theorem hold. $A = A* \Rightarrow T = T*$

Also (ii) implies $< T(v), v > 0$. If non zero $v \in V$, then we can write $v = x_1u_1 + x_2u_2, ... + x_nu_n$ where $x_1, x_2, ...x_n$ are scalars not all zero. Hence T is positive.

15.18.6 Working procedure to verify the positiveness of a square matrix:

Let $A=\left[a_{ij}\right]_{n\times n}$ be a square matrix of order n; over the field F. Then the principal minors of A are the following n scalars

$$\det\begin{bmatrix} a_{11} & a_{12} & .... & a_{1k} \\ a_{k1} & .... & .... & a_{kk} \end{bmatrix} = X_k \text{ say } (k=1,2,..n)$$

Then the matrix A is positive if and only if the principal minors are all positive and $A = A^*$. Where as the matrix A is not positive if det A is not positive.

15.19 Worked out Examples:

W.E.13 : T is a linear operator on an inner product space $V = R^2$ defined by $T(a,b)=(2a-2b,-2a+5b)$. Determine whether T is normal, self adjoint or neither. If possible, find an orthonormal basis of eigen vectors of T for V and list the corresponding eigen values.

Solution: $V = R^2(F=R)$ is an inner product space of dimension 2.

So $V(F)=R^2(R)$ has an orthonormal basis $B=\{u_1=(1,0),u_2=(0,1)\}$. Here T is a linear operator on $V(F)$ such that $T(a,b)=(2a-2b,-2a+5b)$ for $u=(a,b)$.

$$T(1,0)=\big(2(1)-2(0),-2(1)+5(0)\big)=(2,-2)$$

and $T(0,1)=\big(2(0)-2(1),-2(0)+5(1)\big)=(-2,5)$

So $[T]_B = \begin{bmatrix} 2 & -2 \\ -2 & 5 \end{bmatrix}$ and $[T^*]_B = \begin{bmatrix} 2 & -2 \\ -2 & 5 \end{bmatrix}$

So $T = T^* \Rightarrow T$ is self adjoint.

Also $[T]_B[T^*]_B = \begin{bmatrix} 2 & -2 \\ -2 & 5 \end{bmatrix}\begin{bmatrix} 2 & -2 \\ -2 & 5 \end{bmatrix} = \begin{bmatrix} 8 & -14 \\ -14 & 29 \end{bmatrix}$

$$[T^*]_B[T]_B = \begin{bmatrix} 2 & -2 \\ -2 & 5 \end{bmatrix}\begin{bmatrix} 2 & -2 \\ -2 & 5 \end{bmatrix} = \begin{bmatrix} 8 & -14 \\ -14 & 29 \end{bmatrix}$$

So $[T]_B[T^*]_B = [T^*]_B[T]_B$

$$\Rightarrow [TT^*]_B = [T^*T]_B \Rightarrow TT^* = T^*T$$

$\Rightarrow T$ is normal.

Let $A=[T]_B$. Characteristic equation is $|A-\lambda I|=0 \Rightarrow \left\|\begin{bmatrix} 2 & -2 \\ -2 & 5 \end{bmatrix}-\lambda\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right\|=0$

$$\begin{vmatrix} 2-\lambda & -2 \\ -2 & 5-\lambda \end{vmatrix}=0 \Rightarrow (2-\lambda)(5-\lambda)-4=0$$

$$\Rightarrow \lambda^2-7\lambda+6=0 \Rightarrow (\lambda-6)(\lambda-1)=0$$

$$\Rightarrow \lambda=6, \lambda=1$$

To find eige vector corresponding $\lambda=6$.

$$(A-\lambda I)X=O$$

$$\begin{bmatrix} 2-6 & -2 \\ -2 & 5-6 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}=O \Rightarrow \begin{bmatrix} -4 & -2 \\ -2 & -1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}=O$$

$\Rightarrow -4x_1-2x_2=0 \Rightarrow x_2=-2x_1$

put $x_1=1$, then $x_2=-2$

So $X=\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}=\begin{bmatrix} 1 \\ -2 \end{bmatrix}$ and every scalar multiple of it is an eigen vector.

To find eigen vector corresponding to $\lambda=1$.

$$(A-\lambda I)X=O$$

$$\Rightarrow \begin{bmatrix} 2-1 & -2 \\ -2 & 5-1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}=O \Rightarrow \begin{bmatrix} 1 & -2 \\ -2 & 4 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}=0$$

$\Rightarrow x_1-2x_2=0$ and $-2x_1+4x_2=0$

$\Rightarrow x_1=2x_2$ so put $x_2=1$, then $x_1=2$

So $X=\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}=\begin{bmatrix} 2 \\ 1 \end{bmatrix}$ and every scalar multiple of it is an eigen vector.

An orthonormal basis of eigen vectors is $\left\{\frac{1}{\sqrt{5}}(1,-2), \frac{1}{\sqrt{5}}(2,1)\right\}$ with corresponding eigen values 6 and 1.

W.E. 14 : Let V be a finite dimensional inner product vector space and T be an idempotent operator on V i.e. $T^2 = T$ then T is self adjoint if and only if $TT^* = T^*T$.

Solution: Let T be self adjoint then $T^* = T$ and $T^2 = T$.

To prove $TT^* = T^*T$

Now $<T(u), v> = <u, T^*(v)> = <u, T(v)>$

$$<u, TT(v)> = <u, T^*T(v)>$$

$$= <u, TT^*(v)> \text{ since } T^* = T.$$

$$\Rightarrow T^*T = TT^*$$

Converse : Given $T^2 = T, TT^* = T^*T$ to prove that T is self adjoint i.e. $T^* = T$.

Now $<T(u), T(u)> = <u, T^*T(u)> = <u, TT^*(u)>$

$$= <T^*u, T^*u>$$

$$\Rightarrow \|T(u)\|^2 = \|T^*(u)\|^2$$

$\therefore T(u) = O$ if and only if $T^*(u) = O$ .... (1)

Choose any vector in V such that $u = v - T(v)$.

$$T(u) = T(v - T(v) = T(v) - T^2(v) = T(v) - T(v) = O$$

Since $T(u) = 0$ it follows that $T^*(u) = O$

But $T^*(u) = O \Rightarrow T^*(v - T(v) = O$

or $T^*(v) - T^*T(v) = O$

or $T^*(v) = T^*T(v) \forall v \in V$

So $T^* = T^*T$ .... (2)

Now $T = (T^*)^* = (T^*T)^* = T^*T^{**} = T^*T = T^*$

As $T = T^*$, therefore, T is self adjoint.

W.E.15 : Let V be the space of polynomials over the field of complex numbers with inner product defined as $< f,g >= \int_0^1 f(t)\overline{g(t)}, f,g \in V$. If D is the differetial operator, find out if D is itself adjoint or not.

Solution: Let dash denotes the differentiation i.e $Df = f^{'}$

$$< Df,g >=< f^{'},g >= \int_0^1 f^{'}(t)\overline{g(t)}dt$$

intergrating by parts we get

$$< Df,g >= [f \quad \overline{g}]_0^1 - \int_0^1 f(t)\overline{g^{'}(t)}dt \text{ .... (1)}$$

Also $< f,Dg >=< f,g^{'} > -\int_0^1 f(t)\overline{g^{'}(t)}dt$ by definition .... (2)

If D were to be self adjoint $< Df,g >=< f,Dg >$

But since (1) and (2) are not the same. D is not selfadjoint.

W.E.16 : If T_1, T_2 are positive linear operators on an inner product vector space then prove that $T_1 + T_2$ is also positive.

Solution: Given $T_1^* = T_1, < T_1(u), u > \ > 0; \ \ T_2^* = T_1$

and $< T_2(u), u > \ > 0$

Now $(T_1 + T_2)^* = T_1^* + T_2^* = T_1 + T_2$

$\therefore T_1 + T_2$ is self adjoint.

Again $< (T_1 + T_2)(u), u >=< T_1(u) + T_2(u), u >$

$=< T_1(u), u > + < T_2(u), u >> 0$ by given conditions.

Hence $T_1 + T_2$ is also +ve.

15.20 Summary:

In this lesson we discussed about linear operators. Adjoint operator. Properties of adjoint operators. Normal and self adjoint operators their properties-polynomial split-schur theorem, positive, semi positive square matrices.

15.21 Technical Terms:

The technical terms we come across in this lesson are adjoint operator, self adjoint operator, polynomial split, Normal operator, positive matrix, semi positive matrix.

15.22 Model Questions:

1. Define adjoint of a linear operator on an inner product space V. If S^*, T^* are adjoint operators of S and V, then prove that

(i) $(S+T)^* = S^* + T^*$

(ii) $(ST)^* = T^* S^*$

2. Define self adjoint operator. Let T be a self adjoint linear operator on a finite dimensional inner product space. Then prove that $R(T) = \{N(T)\}^{\perp}$

3. State and prove schur theorem in Linear operators.

4. Define normal operator. If T is a normal operator on an inner product space then show that $T - CI$ is normal for every $C \in F$.

5. Define positive linear operator. If T_1, T_2 are positive linear operators on an inner product space then prove that $T_1 + T_2$ is above positive.

15.23 Exercise:

1. For each linear operator T on an inner product space V, determine whether T is normal, self adjoint or neither. If possible produce an orthonormal basis of eigen vectors of T for V and list the corresponding eigen values.

1) $V = C^2$, and T defined by $T(a,b) = (2a + ib, a + 2b)$

Ans : T is normal, but not self adjoint. An orthonormal basis of eigen vectors is $\left\{\frac{1}{2}\left((1+i), \sqrt{2}\right), \frac{1}{2}(1+i, -\sqrt{2}\right\}$ with corresponding eigen values $2 + \frac{1+i}{\sqrt{2}}, 2 - \frac{(1+i)}{\sqrt{2}}$.

Ans: T is self adjoint.

ii) $V = M_{2\times 2}(R)$ and T is defined by $T(A) = A^T$

Ans: T is self adjoint. An orthonormal basis of eigen vectors.

$$\left\{\frac{1}{\sqrt{2}}\begin{pmatrix}0 & 1\\1 & 0\end{pmatrix},\frac{1}{\sqrt{2}}\begin{pmatrix}1 & 0\\0 & 1\end{pmatrix},\frac{1}{\sqrt{2}}\begin{pmatrix}0 & -1\\1 & 0\end{pmatrix},\frac{1}{\sqrt{2}}\begin{pmatrix}-1 & 0\\0 & 1\end{pmatrix}\right\}$$ corresponding eigen values are $1,1,-1,-1$.

2. Let V be a complex inner product space, and let T be a linear operator on V.

Prove that T is normal if and only if $T_1T_2 = T_2T_1$.

3. Prove that every entry on the main diagonal of a positive matrix is positive.

4. Which of the following matrices are positive.

i) $A = \begin{bmatrix} 0 & i \\ -i & 0 \end{bmatrix}$ ii) $B = \begin{bmatrix} 1 & 1+i \\ 1-i & 3 \end{bmatrix}$

Ans: (i) A is not positive

(ii) B is positive

5. If T is a linear operator on an inner product space $V(F)$ and a,b are scalars such that $|a| = |b|$, then show that $aT + bT^*$ is normal.

15.24 Reference Books:

1. Linear Algebra 4th edition; Stephen H. Friedberg. Arnold J. Insel, Lawrence E. Spence.

2. Sehaum's outlines, Beginning Linear Algebra, Seymour Lipschutz.

3. Linear Algebra Dr. S.N. Goel

4. Linear Algebra K.P. Gupta

- A. Mallikharjana Sarma

LESSON - 16

UNITARY AND ORTHOGONAL OPERATORS

16.1 Objective of the Lesson:

In this chapter we study the analogy between complex numbers and linear operators. In the previous lessons, we observed that the adjoint of a linear operator acts similarly to the conjugate of a complex number. A complex number has length 1, if $z\bar{z} = 1$.

In this lesson we study those linear operators T on an inner product space V, such that $TT^* = T^*T = I$. We see that these are precisely the linear operators that preserves length; in the sense that $\|T(u)\| = \|u\|$ for all $u \in V$. As another characterisation we prove that on a finite dimensional complex inner product space, these are the normal operators whose eigen values all have absolute value 1.

16.2 Structure of the lesson :

This lesson contains the following items:

16.3 Introduction :

Linear operators preserve the operations of vector addition and scalar multiplication and isomorphisms preserve all the vector space structure. We now consider those linear operators T on an inner product space that preserve the length. We see that this condition guarantees, that T preserves the inner product.

16.4 Some Basic Definitions:

16.4.1 Definition:

Let U and V be two inner product spaces over F. Let $T : U \to V$ be a linear transformation. Then we say that

i) T preserves inner products if $< T(u), T(v) > = < u, v >$ for all $u, v \in U$.

ii) T Preserves norms if $\|T(u)\| = \|u\| \forall u \in U$.

iii) T preserves isometry if T preserves distances i.e. if $\|T(u) - T(v)\| = \|u - v\| \forall u, v \in U$.

Note that the distance $T(u)$ to $T(v)$ is d$\left(T(u), T(v)\right)$ and is equal to $\|T(u) - T(v)\|$.

16.4.2 Equivalent Conditions:

Let $U(F)$ and $V(F)$ be two inner product spaces. Let $T : U(F) \to V(F)$ be a linear transformation. Then the following three conditions are equivalent.

i) T preserves inner product ii) T preserves norms iii) T is an isometry.

16.4.3 Inner Product Isomorphism:

16.4.4 Definition: Let T be a linear transformation from an inner product space $V(F)$ to an inner product space $V(F)$. Then T is said to be an inner product space isomorphism if

i) T is invertible i.e. T is one one onto

ii) T preserves inner products

Here $U(F)$ and $V(F)$ are said to be isomorphic and we write $U \cong V$.

T preserves inner products $\Rightarrow$ T is non singular

$\Rightarrow$ T is one one.

Hence an inner product space isomorphism from U onto V can also be defined as a linear transformation from U onto V, which preserves inner products.

16.5 Definition:

i) Unitary Operator:

Let T be a linear operator on a finite dimensional inner product space over the field of complex numbers and $\|T(u)\| = \|u\|$ for all $u \in U$ then T is called a unitary operator.

ii) Orthogonal Operator: Let T be a linear operator on a finite dimensional inner product space V, over the field of real numbers R and if $\|T(u)\| = \|u\|$ for all $u \in V$, then T is said to be orthogonal opeator.

iii) Isometry: Let T be a linear operator on an infinite dimensional inner product space V over F and if $\|T(u)\| = \|u\|$ for all $u \in V$, then T is called an isometry.

If in addition, the operator is onto (the condition guarantees one to one), then the operator is called unitary if $F = C$ or orthogonal operator if $F = R$.

iv) Definition: U and V are two vector spaces over a field F. Then the zero transformation $T_0 : U \to V$ is defined by $T_0(u) = O \;\forall\; u \in U$. The zero transformation is also denoted by $\hat{0}$.

16.6.1 Theorem: Let T be a self adjoint operator on a finite dimensional inner product space V. If $< u, T(u) >= 0 \quad u \in V$; then $T = To$.

Proof: We know that if T is a linear operator on a finite dimensional inner product space V, then T is said to be self adjoint, if and only if there exists on orthonormal basis B for V consisting of eigen vectors of T.

By the above theorem we can choose an orthonormal basis B for V. consisting of eigen vectors of T. If $u \in B$ then $T(u) = \lambda u$ for some λ.

then 0 = $< u, T(u) >=< u, \lambda u >= \bar{\lambda} < u, u >$

$\Rightarrow \bar{\lambda} = 0$ Hence $T(u) = O$ for all $u \in B$

So $T = T_0$

16.6.2 Equivalent Statements:

Theorem: Let T be a linear operator on a finite dimensional inner product space V. Then the following statements are equivalent

i) $TT^* = T^*T = I$

ii) $< T(u), T(v) >=< u, v > \;\forall\; u, v \in V$

iii) If B is an orthonormal basis for V, then $T(B)$ is an orthonormal basis for V.

iv) There exists an orthonormal basis B for V, such that $T(B)$ is an orthonormal basis for V.

v) $\|T(u)\| = \|u\| \; \forall \; u \in V$

Proof:

1. We will now prove that (i) $\Rightarrow$ (ii)

given $TT^* = T^*T = I$ .... (1)

Let $u, v \in V \Rightarrow \langle u, v \rangle = \langle I(u), v \rangle$

$= \langle (T^*T)(u), v \rangle$ using (1)

$= \langle T^*(T(u), v \rangle$

$= \langle T(u), T(v) \rangle$

Thus $TT^* = T^*T = I \Rightarrow \langle T(u), T(v) \rangle = \langle u, v \rangle$ for all $u, v \in V$

Hence (i) $\Rightarrow$ (ii)

ii) To show (ii) $\Rightarrow$ (iii)

Given $\langle T(u), T(v) \rangle = \langle u, v \rangle \; \forall u, v \in V$ .... (2)

Let $B = \{u_1, u_2, ...u_n\}$ be an orthonormal basis for V. So $T(B) = \{T(u_1), T(u_2), ..., T(u_n)\}$

from (2) $\langle T(u_i), T(u_j) \rangle = \langle u_i, u_j \rangle$

$$\langle \delta_{ij} \rangle = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

$\Rightarrow$ $T(B) = \{T(u_1), T(u_2), ...T(u_n)\}$ is an orthonormal basis for V.

Hence (ii) $\Rightarrow$ (iii)

3. To show that (iii) $\Rightarrow$ (iv)

Given B is an orthonormal basis for V, then $T(B)$ is an orthonormal basis for V....... (3)

Let $B = \{u_1, u_2, ...u_n\}$ be an orthonormal basis for V; then by (3) $T(B) = \{T(u_1), T(u_2), ...T(u_n)\}$ is an orthonormal basis for V.

Hence (iii) $\Rightarrow$ (iv)

4. To show that (iv) $\Rightarrow$ (v)

Let $u \in V$ and $B = \{u_1, u_2, ...u_n\}$ is a basis for V. So $u = \sum_{i=1}^{n} a_i u_i$ for some scalar a_i

$$\|u\|^2 = < \sum_{i=1}^{n} a_i u_i, \sum_{j=1}^{n} a_j u_j >$$

$$= \sum_{i=1}^{n} a_i \sum_{j=1}^{n} \overline{a_j} < u_i, u_j >$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} a_i \overline{a_j} < u_i, u_j >$$

$= \sum_{i=1}^{n} a_i \bar{a}_i < u_i, u_i >$ summing of $j = 1, 2, ...n$ and remembering

$$< u_i, u_j > = \begin{cases} 0 \text{ if } i \neq j \\ 1 \text{ if } i = j \end{cases}$$ as B is orthonormal.

$$= \sum_{i=1}^{n} a_i \bar{a}_i \qquad (1)$$

$$= \sum_{i=1}^{n} |a_i|^2 \qquad ...... (A)$$

Applying the same manipulation to $T(u) = \sum_{i=1}^{n} a_i T(u_i)$ and using the fact $T(B)$ is also orthonormal we obtain $\|T(u)\|^2 = \sum_{i=1}^{n} |a_i|^2$ .... (B)

From (A) and (B) we get $\|T(u)\| = \|u\|$

So (iv) $\Rightarrow$ (v)

5. Finally we will prove that $V \to i$ (i)

Let $u \in V$ we have $< u,u >= \|u\|^2 = \|T(u)\|^2$

$$=< T(u),T(u) >$$

$$=< u,T*T(u) >$$

So $< u,u > - < u,T^*T(u) >= 0 \forall u \in V$

$\Rightarrow < u,(I - T*T)u >= 0 \forall u \in V$

Let $S = (I - T*T)$ then S is self - adjoint and $< u,S(u) >= 0 \forall u \in V$,

We know if T is a self adjoint operator on a finite dimensional inner product space and if $< u,T(u) >= 0 \forall u \in V$. Then $T = T_0$ where T_0 is the zero transformation.

So $T_0 = S = (I - T*T)$

So $T*T = I$

Since V is finite dimensional, $TT* = I$

Hence $\|T(u)\| = \|u\| \forall u \in V, \Rightarrow TT* = T*T = I$

Hence (v) $\Rightarrow$ (i)

Note: Among the equivalent conditions, any one can be taken as a definition in doing problems.

16.6.3 Theorem:

If T is a unitary operation then show that $T* = T^{-1}$

Proof: T is a unitary operation $\Rightarrow \|T(u)\| = \|u\|$

$$\Rightarrow \|T(u)\|^2 = \|u\|^2$$

$$\Rightarrow < T(u),T(u) >=< u,u >$$

$$\Rightarrow < (T*T)(u),u >=< u,u >$$

$$\Rightarrow < (T*T)(u),u >=< I(u),u >$$

$$\Rightarrow < (T*T - I)u,u >= 0$$

$\Rightarrow T*T - I = \hat{O}$ (null operator)

$\Rightarrow T*T = I \Rightarrow T*TT^{-1} = IT^{-1}$

$\Rightarrow T*I = T^{-1} \Rightarrow T* = T^{-1}$

Thus if T is a unitary operator then $T* = T^{-1}$

Remark: If T is unitary, then T is non singular.

16.6.4 Theorem:

Let T be a linear operator on a finite dimensional real inner product space V, then V has an orthonormal basis of eigen vectors of T with corresponding eigen values of obsolute value 1 if and only if T is both self adjoint and orthogonal.

Proof: Let V has then orthogonal basis $B = \{u_1, u_2, ..., u_n\}$

such that $T(u_i) = \lambda u_i$ and $|\lambda| = 1$ for all i,

As T is a linear operator and there exists an orthonormal basis B for V consisting of eigen vectors of T, and so T is self adjoint.

Thus $(TT*)(u_i) = T(\lambda_i u_i) = \lambda_i \lambda_i u_i = \lambda_i^2 uu_i = u_i$ for each i. So $TT* = I$ so T is orthogonal.

Hence T is both self adjoint and orthogonal.

Converse: Let T be both self adjoint and Orthogonal:

We know if T is a linear operator on a finite dimensional real inner product space V, then T is self adjoint if and only if there exists an orthonormal basis B for V consisting of eigen vectors of T.

So by this theorem V possess an orthonormal basis $B = \{u_1, u_2, ..., u_n\}$ such that $T(u_i) = \lambda_i u_i$ for all i. If T is also orthogonal. We have $|\lambda_i|.\|v_i\| = \|\lambda_i v_i\| = \|T(v_i)\| = \|v_i\|$

Thus $|\lambda_i|\|v_i\| = \|v_i\| \Rightarrow |\lambda_i| = 1$ for every i.

Hence from the above two cases the theorem follows.

16.6.5 Corollary: Let T be a linear operator on a finite dimensional complex inner product space V. Then V has an orthonormal basis of eigen vectors of T with corresponding eigen values of absolute value 1 if and only if T is unitary.

Proof: The proof is similar as above.

16.7 Reflection:

16.7.1 Definition: Let L be a one dimensional subspace of R^2 about a line L through the origin.

A linear operator T on R^2 is called a reflection of R^2 about L; if $T(u) = u \forall u \in L$ and $T(u) = -u. \forall u \in L^{\perp}$.

16.7.2 Example: Let T be a reflection of R^2 about a line through the origin. We shall show that T is an orthogonal operator. Select vectors $u_1 \in L$ and $u_2 \in L^{\perp}$ such that $\|u_1\| = \|u_2\| = 1$.

Then $T(u_1) = u_1, T(u_2) = -u_2$ thus u_1 and u_2 are eigen vectors of T with corresponding eigen values 1 and -1 respectively. Further more $\{u_1, u_2\}$ is an orthonormal basis for R^2. It follows that T is an orthogonal operator.

16.8 Some Basic Properties of Unitary Operators:

16.8.1 Show that every unitary operator is normal:

Proof: Let T be linear operatory on the inner product space V. Which is unitary.

So $TT^* = T^*T = I$. Hence T is invertible and $T^{-1} = T^*$.

So T is normal.

Hence every unitary operator is normal.

16.8.2 If S and T are unitary operators, then ST is unitary or the product of two unitary operators is unitary.

Proof: S, T are two unitary operators on V. Then

$$S^{-1} = S^*, T^{-1} = T^* \text{ .... (1)}$$

Now $(ST)^{-1} = T^{-1}S^{-1} = T^*S^*$ using (1)

$$= (ST)^*$$

As $(ST)^{-1} = (ST)^*$; ST is unitary. Hence the product of two unitary operators is unitary.

Aliter i) : S and T are two unitary operators on a finite dimensional inner product space V.

We have to show that ST is unitary.

Now $(ST)(ST)^* = ST(T^*S^*)$

$$= S(TT^*)S^*$$

$$= SIS^* = SS^* = I$$

So $(ST)(ST)^* = I$

Similarly $(ST)^*(ST) = I$

So ST is unitary.

Hence the theorem.

ii) Let S and T be the given unitary operators then S and T are invertible. So ST is invertible

Also $\|(ST)(u)\| = \|S(T(u)\|$

$= \|T(u)\| = \|u\|$ since S and T are unitary.

So $\|ST(u)\| = \|u\|$

So ST is unitary.

Hence the theorem.

16.8.3 Corollary: Prove that the composite of orthogonal operators is orthogonal.

Proof: Similar as above.

16.8.4 Theorem:

Show that the inverse of a unitary operator is unitary.

Proof: Let V be an inner product space. T is a unitary operator on V. We have to show that T^{-1} is unitary.

As T is unitary $\|T(u)\| = \|u\|$ for all $u \in V$

We put $v = T(u)$ ie $T^{-1}(v) = u$

We get $\|TT^{-1}(v)\| = \|T^{-1}(v)\|$

$\Rightarrow \|v\| = \|T^{-1}(v)\|$

Hence $\|T^{-1}(v)\| = \|v\| \forall v \in$ V

Which implies T^{-1} in unitary.

Aliter : Let T be unitary. T^{-1} is the inverse operator of T. Now $(T^{-1})^{-1} = (T^*)^{-1} = (T^{-1})^*$

Since T is unitary $T^{-1} = T^*$.

Thus as $\left(T^{-1}\right)^{-1} = \left(T^{-1}\right)^*$ it follows that the inverse of a unitary operator is unitary.

16.8.5 Show that the set of all unitary operators on an inner product space V is a group with respect to composite of operations.

Solution: Let G denote the set of all unitary operators on an inner product space $V(F)$.

Let T_1, T_2 be arbitary unitary operators belonging to G.

So $T_1 T_1 * = T_1 * T_1 = I, T_2, T_2 * = T_2 * T_2 = I$

First we will show that $T_1 T_2$ is a unitary operator

Now $(T_1, T_2)(T_1 T_2)* = (T_1 T_2)(T_2 * T_1 *) = T_1 (T_2 T_1 *) T_2$

$$= T_1 I T_1^* = T_1 T_2^* = I$$

Thus $(T_1 T_2)(T_1 T_2) = I$

$(T_1 T_2) * (T_1 T_2) = (T_2 * T_1 *)(T_1 T_2) = T_2^* (T_1 * T_1) T_2$

$$= T_2 * I T_2 = (T_1 * T_2) I$$

Hence $(T_1 T_2)(T_1 T_2)* = I = (T_1 T_2) * (T_1 T_2)$

So $T_1 T_2$ is a unitary operators.

We verify group axiones.

i) Closure Property: If $T_1 T_2$ are any two unitary operators belongging to G, $T_1 T_2$ is a unitary operator and hence belongs to G. Hence G is closed.

ii) Associativity: We know composite of operators is associative. Hence if T_1, T_2, T_2 are any three unitary operators in G. then $(T_1 T_2) T_3 = T_1 (T_2 T_3)$.

iii) Existence of Identity: Let I be the identity operator in G. So I is inversible and $\|I(u)\| = \|u\| \forall u \in V$. so I is unitary operator on V. Hence belongs to G and $IT = TI = T \forall T \in G$.

So I is the identity element in G.

iv) Existence of Inverse : T is unitary $\Rightarrow T$ is inversible $\Rightarrow T^{-1}$ is also invertible.

Let $T^{-1}(u) = v$ so that $T(v) = u$ for all $u, v \in V$

$\therefore \|T(u)\| = \|v\|$ .... (1)

But T is unitary $\Rightarrow \|(u)\| = \|T^{-1}(v)\|$

$\Rightarrow \left\| T^{-1}(u) \right\| = \|u\| \Rightarrow T^{-1}$ is also unitary.

$\therefore T^{-1} \in G \cdot$

As all the group axious are satisfied, the set G of all unitary operators on V is a group.

16.8.6 Let T be a unitary operator on an inner product space V, let W be a finite dimensional T-invariant subspace of V. Prove that $W^{\perp}$ is T - invariant.

Solution: W is a subspace V, W is T - invariant.

So any $w \in W \Rightarrow T(w) \in W$

To prove that $W^{\perp}$ is T - invariant, it enough to show that any $w \in W^{\perp} \Rightarrow T(w) \in W^{\perp}$

Let $w \in W$ and $u \in W^{\perp}$ be arbitrary, then $< u, w >= 0$ and $T(w) = w_1 \in W$

As T is unitary

$< T(u), w_1 >=< T(u), T(w) >=< u, w >= 0$

Thus $< T(u), w_1 >= 0 \forall w_1 \in W$

This implies $T(u)$ is $\perp$ W ($T(u)$ is perpendicular to W) $\Rightarrow T(u) \in W^{\perp}$

any $u \in W^{\perp} \Rightarrow T(u) \in W^{\perp}$

So $W^{\perp}$ is T-invariant.

16.8.7 Show that the determinant of a unitary operator has absolute value.

Solution: Let T be a unitary operator on a finite dimensional inner product space $V(F)$. Let B be an ordered orthonormal basis for V. Let A denote matrix of T relative to B. Then

$\det T = \det A = \det [T]_B$

T is unitary $\Rightarrow T^*T = I \Rightarrow [T^*T]_B = [I]_B$

$\Rightarrow [T^*]_B . [T]_B = [I]_B$

So $\det [T^*]_B . \det [T]_B = 1 \Rightarrow \det A^* . \det A = 1$

$\Rightarrow \overline{(\det A)} (\det A) = 1$

$\Rightarrow |(\det A)|^2 = 1 \Rightarrow \det A = 1$

So det A and hence det T has absolute value 1.

16.9 Matrices representing Unitary and Orthogonal Transformations:

16.9.1 Definition: A square matrix A is called an orthogonal matrix. If $A^T A = AA^T = I$ and unitary if $A^* A = AA^* = I$ ·

Note i) : For a real matrix A we have $A^* = A^T$

So a real unitary matrix is also orthogonal. In this case, we call orthogonal rather then unitary.

ii) The condition $AA^* = I$ is equivalent to the statement that the rows of A form an orthonormal basis for F^n because

$\delta_{ij} = I_{ij} = (AA^*)_{ij} = \sum_{k=1}^{n} A_{iK} (A^*)_{kj} = \sum_{k=1}^{n} A_{ik} \overline{A_{jk}}$ and the last term represents the inner product of ith row and jth column.

Here A_{ik} represents the ikth entry of the matrix A.

iii) The condition $A^* A = I$ is equivalent to the statement that the columns of A form an orthonormal basis of F^n.

iv) A linear operator T on an inner prouct space V is unitary (Orthogonal) if and only if $[T]_B$ is unitary (orthogonal) for some orthonormal basis B for V.

Ex: The matrix $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ is clearly orthogonal. One can easily see that the rows of the matrix form an orthonormal basis for R^2 . Similarly the columns of the matrix form an orthonormal basis for R^2.

16.9.2 Equivalent Matrices : Definition:

Let A and B are unitary (orthogonal) matrices of order n x n. There B is unitarily equivalent or (orthogonally equivalient) if and only if there exists an n x n unitary. (Orthogonal) matrix.

P such that $B = P^* AP$

Note: The relation unitarily equivalent to (orthogonally equivalent to) is an equivalence relation on $M_{n\times n}(C)\left[M_{n\times n}(R)\right]$.

W.E.1: Let T be a reflection of R^2 about a lin L through the origin, Let B be the standard ordered basis for R^2 and let $A = [T]_B$ then $T = L_A$.

Since T is an orthogonal operator and B is an orthonormal basis, A is an orthogonal matrix. Discribe A.

Solution: Suppose that α is the angle from the positive x axis to L. Let $v_1 = (\cos\alpha, \sin\alpha)$ and $v_2 = (-\sin\alpha, \cos\alpha)$. Then $\|v_1\| = \|v_2\| = 1$.

$v_1 \in L,\ v_2 \in L^{\perp}$. Hence $W = \{v_1, v_2\}$ is an orthonormal basis for R^2. Because $T(v_1) = v_1$ and $T(v_2) = -v_2$ we have

$$[T]_W = [L_A]_W = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

Let $Q = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}$

Thus $A = Q\left(L_A\right)_W Q^{-1}$

$$= \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix}$$

$$= \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix} \begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix}$$

$$= \begin{bmatrix} \cos^2\alpha - \sin^2\alpha & 2\sin\alpha\cos\alpha \\ 2\sin\alpha\cos\alpha & -(\cos^2\alpha - \sin^2\alpha) \end{bmatrix}$$

$$= \begin{bmatrix} \cos z\alpha & \sin 2\alpha \\ \sin 2\alpha & -\cos 2\alpha \end{bmatrix}$$

We know that for a complex normal (real symmetric) matrix A, there exists an orthogonal basis B for F^n. Consisting of eigen vectors of A. Hence A is similar to a diagonal matrix D. We know that $A \in M_{n\times m}(F)$ and W be an ordered basis for F^n.

Then $\left(L_A\right)_W = Q^{-1}AQ$. Where Q is the n x n matrix whose jth column is the vector of W. Hence by this theorem, the matrix Q whose columns are the vectors in B is such that $D = Q^{-1}$AQ, But since the columns of Q are the orthonormal basis for F^n, it follows that Q is unitary (orthogonal). Hence A is unitarily equivalent (orthogonally equivalent) to D.

16.10.1 Theorem:

Let V be a finite dimensional inner product space and T be the linear operator on V. Then T is unitary if and only if the matrix T in some (or every) ordered orthonormal basis for V is a unitary matrix.

Proof: V is a finite dimensional inner product space. T is a linear operator on V. Let $B = \{u_1, u_2, ...u_n\}$ be an ordered orthonormal basis for V. and let A be the matrix of T relative to B. i.e. $[T]_B = A$.

Case i) : Let T be unitary. Then T is invertible and hence $(T^*T) = I \Rightarrow [T^*T]_B = I_B$

$$\Rightarrow [T^*]_B . [T]_B = I \quad \text{where} \quad A^* = [T^*]_B$$

$A = [T]_B$ then $A^*A = I$

So $A = [T]_B$ is unitary.

Converse:

Suppose that the matrix A is unitary there we have $A^*A = I$.

$$\Rightarrow [T^*]_B [T]_B = [I]_B$$

$$\Rightarrow [T^*T]_B = [I]_B$$

$$\Rightarrow T^*T = I$$

So T is unitary.

From the above two cases the theorem follows.

16.10.2 Corollary: A linear operator T on an inner product space V is orthogonal if and only if $[T]_B$ is orthogonal for same orthonormal basis B for V.

Proof: Similar as above.

16.10.3 Let A be a complex n x n matrix, then A is normal if and only if A is unitary equivalent to a diagonal matrix.

Proof: Let C^n be the vector space V with standard inner product defined on it and let B be its ordered basis. If T is the linear operator on V such that it is represented in the standard ordered basis by the matrix A, then we have $[T]_B = A$ and $[T^*]_B = A^*$

Now $[T^*T]_B = [T]_B . [T^*]_B = AA^*$ and $[T^*T]_B = A^*A$.

Case i) If A is a normal matrix then $A^*A = AA^*$ and hence $[TT^*]_B = [T^*T]_B$. i.e. $TT^* = T^*T$ i.e. T is a normal operator. From the above, i.e. T being a linear operator on an inner product space V, it follows that there exists an orthonormal basis say B_1 for V each vector of which is a characteristic vector for T, and hence $[T]_{B_1}$ is a diagonal matrix. Further if P is a transition matrix from B to B_1, then it is a unitary matrix as both B and B_1 are orthonormal bases. When P is a unitary matrix we have $P^*P = I$ i.e. $P^* = P^{-1}$. Also we have $[T]_{B_1} = P^{-1}[T]_B P = P^*AP$ = diagonal matrix.

Converse: Suppose that A is unitarily equivalent to a diagonal matrix D.

Suppose that $A = P^*DP$ where P is a unitary matrix and D is a diagonal matrix

then $AA^* = (P^*DP)(P^*DP)^*$

$= (P^*DP)(P^*DP^{**})$

$= (P^*DP)(P^*D^*P)$

$P^*D(PP^*)D^*P$

P^*DID^*P

$P^*(DD^*)P$ .... (1)

Similarly $AA^* = P^*(DD^*)P$ .... (2)

Since D is a diagonal matrix, however

We have $DD^* = D^*D$ so $AA^* = A^*A$.

Hence A is normal.

From the above two cases, the theorem follows.

16.10.4 Corollary: Let A be a real n x n matrix: Then A is symmetric if and only if A is orthogonally equivalent to a real diagonal matrix.

Proof: The proof is similar as above.

16.10.5 Schur's Theorem for Matrices: By Schaur's theorem proved in normal and adjoint operators, in the matrix form, it can be stated as

Let A be a matrix $M_{n \times n}(R)$ whose characteristic polynomial splits over F then

i) If $F = C$ then A is unitarily equivalent to a complex upper triangular matrix.

ii) If $F = R$, then A is orthogonally equivalent to a real upper triangular matrix.

16.11 Worked out examples:

W.E.2: Let T be a linear operator on R^3 which rotates every vector in R^3 about Z axis by a constant angle θ. Prove that T is a orthogonal transformation:

Solution: T is invertible since there exists a linear transformation T^{-1} which rotates every vector in R^3 about Z axis by a constem angle θ in the direction opposite to T. Hence

$$T^{-1}(T(u) = (T^{-1}T)u = I(u) = u \forall u \in R^3$$

Also if $u = (x, y, z) \forall x, y, z \in R$ then

$$T(u) = T(x, y, z) = (x\cos\theta - y\sin\theta, x\sin\theta + y\cos\theta, z)$$

$$\|u\|^2 = x^2 + y^2 + z^2$$

$$\|T(u)\|^2 = (x\cos\theta - y\sin\theta)^2 + (x\sin\theta + y\cos\theta)^2 + z^2$$

$$= x^2 + y^2 + z^2$$

Hence $\|T(u)\| = \|u\|$

Hence the linear transformation T is orthogonal.

W.E.3: Show that the matrix $A = \begin{bmatrix} a+ic & -b+id \\ b+id & a-ic \end{bmatrix}$ is unitary if and only if $a^2 + b^2 + c^2 + d^2 = 1$.

Solution: $A^* = \begin{bmatrix} a-ic & b-id \\ -b-id & a+ic \end{bmatrix}$

Now $AA^* = \begin{bmatrix} a+ic & -b+id \\ b+id & a-ic \end{bmatrix}\begin{bmatrix} a-ic & b-id \\ -b-id & a+ic \end{bmatrix}$

$$= \begin{bmatrix} a^2+b^2+c^2+d^2 & 0 \\ 0 & a^2+b^2+c^2+d^2 \end{bmatrix}$$

So $AA^* = I$ and only if $a^2 + b^2 + c^2 + d^2 = 1$

Hence A is unitary if and only if $a^2 + b^2 + c^2 + d^2 = 1$.

W.E.3A: Show that the matrix $A=\begin{bmatrix}0 & 2m & n\\ l & m & -n\\ l & -m & n\end{bmatrix}$ where $l=\frac{1}{\sqrt{2}}, m=\frac{1}{\sqrt{6}}, n=\frac{1}{\sqrt{3}}$ is orthogonal.

Solution: Let C_1, C_2, C_3 be the column vectors of A. Then

$$C_1=\begin{bmatrix}0\\ l\\ l\end{bmatrix}; C_2=\begin{bmatrix}2m\\ m\\ -m\end{bmatrix}; C_3=\begin{bmatrix}n\\ n\\ -n\end{bmatrix}$$

We have $< C_1, C_1 >= 0+l^2+l^2 = 2l^2 = 2\times\frac{1}{2}=1$

$$< C_2, C_2 >= 4m^2+m^2+m^2 = 6m^2 = 6.\frac{1}{6}=1$$

$$< C_3, C_3 >= n^2+n^2+n^2 = 3n^2 = 3.\frac{1}{3}=1$$

$$< C_1, C_2 >= 0(2m)+l(m)-l(m)=0$$

$$< C_2, C_3 >= 2mn-mn-mn=0$$

$$< C_3, C_1 >= 0n+l(-n)+ln=0$$

Thus the columns of A form an orthogonal set of vectors. So A is orthogonal.

W.E.4 : Find an orthogonal matrix P whose first row is $\left(\frac{1}{3},\frac{2}{3},\frac{2}{3}\right)$.

Solution: Let $u_1=\left(\frac{1}{3},\frac{2}{3},\frac{2}{3}\right)$

First we find a non zero vector $w_2=(x,y,z)$ which is orthogonal to u_1, for which

$$< u_1, w_2 >= 0$$

$$< \left(\frac{1}{3},\frac{2}{3},\frac{2}{3}\right),(x,y,z) >= 0$$

$$\Rightarrow \frac{x}{3}+\frac{2y}{3}+\frac{2z}{3}=0 \Rightarrow x+2y+2z=0$$

Put $x = 0, \Rightarrow z = -y$ put $y=1$, then $z=-1$

So one such solution is $w_2 = (x,y,z) = (0,1,-1)$

Normalize w_2 to get the second row of P i.e. $u_2 = \left(0, \frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}\right)$

Now find a non zero vector $w_3 = (x,y,z)$

Which is orthogonal to both u_1 and u_2 for which

$$< u_1, w_3 >= 0, < u_2, w_3 >= 0$$

$$< u_1, w_3 >= 0 \Rightarrow < \left(\frac{1}{3}, \frac{2}{3}, \frac{2}{3}\right), (x,y,z) >= 0$$

$$\Rightarrow \frac{x}{3}+\frac{2y}{3}+\frac{2z}{3}=0 \Rightarrow x+2y+2z=0 \quad \text{........... (1)}$$

$$< u_2, w_3 >= 0 \Rightarrow < \left(0, \frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}\right), (x,y,z) >= 0$$

$$\Rightarrow 0x+\frac{y}{\sqrt{2}}-\frac{z}{\sqrt{2}}=0 \Rightarrow y-z=0 \quad \text{.............. (2)}$$

Put $z=-1$ then $y=-1 \Rightarrow x=4$ from (1)

So $w_3 = (4,-1,-1)$ normalize w_3, to

obtain the third row of P

i.e. $u_3 = \left(\frac{4}{\sqrt{18}}, \frac{-1}{\sqrt{18}}, \frac{-1}{\sqrt{18}}\right) = \left(\frac{4}{3\sqrt{2}}, \frac{-1}{3\sqrt{2}}, \frac{-1}{3\sqrt{2}}\right)$

Hence the required orthogonal matrix is $P=\begin{bmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ 0 & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ \frac{4}{3\sqrt{2}} & \frac{-1}{3\sqrt{2}} & \frac{-1}{3\sqrt{2}} \end{bmatrix}$

Caution: The above matrix is not unique.

W.E.5 : Let $A=\begin{bmatrix} 1 & 1 & -1 \\ 1 & 3 & 4 \\ 7 & -5 & 2 \end{bmatrix}$ determine whether or not i) the rows of A are orthogonal.

ii) A is an orthogonal matrix.

iii) The columns of A are orthogonal.

Solution: 1. The rows of A are orthogonal since

$$<(1,1,-1),(1,3,4)>=1(1)+1(3)+(-1)4=0$$

$$<(1,1,-1),(7,-5,2)>=(1)(7)+1(-5)+(-1)2=0$$

$$<(1,3,4),(7,-5,2)>=(7)+3(-5)+4(2)=0$$

2. A is not an orthogonal matrix since the rows of A are not unit vectors.

$$\|u_1\|^2=<(1,1,-1),(1,1,-1)>=(1)(1)+1(1)+(-1)(-1)=3$$

$\|u_1\|=\sqrt{3}$ not unity.

3. The columns of A are not orthogonal since for example $<(1,1,7),(1,3,-5)>=1(1)+1(3)+7(-5)=-31\neq 0$.

W.E.6: For which value of α is the following matrix is unitary $\begin{bmatrix} \alpha & \frac{1}{2} \\ \frac{-1}{2} & \alpha \end{bmatrix}$

Solution: Let $A = \begin{bmatrix} \alpha & \frac{1}{2} \\ \frac{-1}{2} & \alpha \end{bmatrix}$ so $A^* = \begin{bmatrix} \bar{\alpha} & \frac{-1}{2} \\ \frac{1}{2} & \bar{\alpha} \end{bmatrix}$

Where A^* is the cojugate transpose of the matrix A.

The matrix A is unitary if $AA^* = I$

i.e. $\begin{bmatrix} \alpha & \frac{1}{2} \\ \frac{-1}{2} & \alpha \end{bmatrix}\begin{bmatrix} \bar{\alpha} & \frac{-1}{2} \\ \frac{1}{2} & \bar{\alpha} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

i.e. $\begin{bmatrix} \alpha\bar{\alpha} + \frac{1}{4} & \frac{-1}{2}\alpha + \frac{1}{2}\bar{\alpha} \\ \frac{-1}{2}\bar{\alpha} + \frac{1}{2}\alpha & \frac{1}{4} + \alpha\bar{\alpha} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

i.e. $\alpha\bar{\alpha} + \frac{1}{4} = 1 \quad \frac{-1}{2}\alpha + \frac{1}{2}\bar{\alpha} = 0$

$\frac{-1}{2}\bar{\alpha} + \frac{1}{2}\alpha = 0 \quad \frac{1}{4} + \alpha\bar{\alpha} = 1$

Solving we get $\frac{1}{2}(\bar{\alpha} - \alpha) = 0 \Rightarrow$ α is real say a

(Since if $\alpha = x + iy$ then $\bar{\alpha} = x - iy$ so

$$\bar{\alpha} - \alpha = (x - iy) - (x + iy) = 0 \Rightarrow y = 0)$$

Again for α real we get $\frac{1}{4} + \bar{\alpha}\alpha = 1$

$$\Rightarrow \alpha^2 = 1 - \frac{1}{4} \Rightarrow \alpha^2 = \frac{3}{4} \text{ so } \alpha = \pm\sqrt{\frac{3}{4}}$$

Hence the matrix A is unitary if $\alpha = \pm\sqrt{\frac{3}{4}}$

W.E. 7: If $V(F)$ is a finite dimensional unitary space and T be a linear transformation on $V(F)$, then show that T is self adjoint $\Leftrightarrow < T(u), u >$ is real for each $u \in V$.

Solution: Case i) Let T be self adjoint.

So $T^* = T$ .... (1)

Now $< T(u), u > = < u, T^*(u) >$ for all $u \in V$

$= < u, T(u) >$` by (1)

$= \overline{< T(u), u >}$`

Thus $< T(u), u > = \overline{< T(u), u >}$

So $< T(u), u >$ is real.

(Since in complex numbers $z = \overline{z} \Rightarrow z$ is real)

Case ii) Converse : Let $< T(u), u >$ be real $\forall u \in V$

then $< T(u), u > = \overline{< T(u), u >} = < u, T(u) >$

$\Rightarrow < u, T^*(u) > = < u, T(u) >$

by definition of adjoint $\Rightarrow < T(u), u > = < u, T^*(u) >$

$\Rightarrow T^* = T$.

So T is self adjoint.

Hence from the two cases the result follows.

W.E.8 : If B and B' are two orthonormal bases for a finite dimensional complex inner product space V. Prove that for each linear transformation T on V, the matrix $[T]_{B'}$ is unitarily equivalent to the matrix $[T]_B$.

Solution: Let P is a transition matrix from B to B' and as B, B' are orthonormal bases , P is a unitary matrix. Thus $P^*P = I \Rightarrow P^* = P^{-1}$

Now $[T]_{B'} = P^{-1}[T]_B P = P^*[T]_B P$

So $[T]_{B^1}$ is unitarily equivalent to the matrix $[T]_B$.

16.12 Working Procedure:

To find an orthogonal or unitary matrix P and a diagonal matrix D for a given matrix A such that $P*AP = D$.

1) Find the characteristic polynomial $f(\lambda)$ and all eigen values of A.

2. Find a maximal set S of non zero orthogonal eigen vectors of A.

3. Then find an orthonormal set $S' = \{u_1, u_2, ..., u_n\}$ of non zero vectors of A.

4. Let P be the matrix whose columns are $u_1, u_2, ..., u_n$.

5. Let D be the diagonal matrix whose diagonal elements are the characteristic roots of A.

6. Then we got the orthogonal matrix P, and diagonal matrix D such that $P*AP = D$.

W.E.9 : If $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$ find an orthogonal matrix P and a diagonal matrix D such that $P^T AP = D$

Solution: The characteristic equation is $|A - \lambda I| = 0$

$$\Rightarrow \begin{vmatrix} 1-\lambda & 2 \\ 2 & 1-\lambda \end{vmatrix} = 0 \Rightarrow (1-\lambda)^2 - 4 = 0$$

$$\Rightarrow (1-\lambda+2)(1-\lambda-2) = 0$$

$$\Rightarrow (3-\lambda)(-\lambda-1) = 0$$

$$\Rightarrow \lambda = 3, \lambda = -1$$

For $\lambda = 3,$ To find a basis for eigen space of A:

$$[A - \lambda I]X = O$$

$$\Rightarrow \begin{vmatrix} 3 & 2 \\ 2 & 1-3 \end{vmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = O \Rightarrow \begin{bmatrix} -2 & 2 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = O$$

$R_2 + R_1$ gives $\begin{bmatrix} -2 & 2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 0 \Rightarrow -2x + 2y = 0$

$$\Rightarrow x = y$$

if $x = 1,$ then $y = 1$

So $v_1 = (1,1)$

To find a basis for eigen space of A for $\lambda = -1$

$$[A - \lambda I]\begin{bmatrix} x \\ y \end{bmatrix} = O$$

$$\Rightarrow \begin{bmatrix} 1+1 & 2 \\ 1 & 1+1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = O \Rightarrow 2x + 2y = 0$$

$$\Rightarrow y = -x$$

if $x = 1$, then $y = -1$

So $v_2 = (1,-1)$

Evidently v_2 is orthogonal to v_1.

Hence the orthogonal basis of A is $\{v_1, v_2\} = \{(1,1),(1,-1)\}$

The corresponding orthonormal basis is $\{u_1, u_2\}$.

Where $u_1 = \dfrac{v_1}{\|v_1\|} = \dfrac{1}{\sqrt{2}}(1,1) = \left(\dfrac{1}{\sqrt{2}}, \dfrac{1}{\sqrt{2}}\right)$

$$u_2 = \frac{v_2}{\|v_2\|} = \frac{1}{\sqrt{2}}(1,-1) = \left(\frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}\right)$$

Thus one possible choice for $P = \begin{bmatrix} \dfrac{1}{\sqrt{2}} & \dfrac{1}{\sqrt{2}} \\ \dfrac{1}{\sqrt{2}} & \dfrac{-1}{\sqrt{2}} \end{bmatrix}$

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

and $D = \begin{bmatrix} 3 & 0 \\ 0 & -1 \end{bmatrix}$

Note : We can apply Gram - Schmidt orthogonalisation process to find orthonormal basis.

W.E.10 : If $A=\begin{bmatrix}4&2&2\\2&4&2\\2&2&4\end{bmatrix}$ find an orthogonal matrix P and a diagonal matrix D such that $P^T AP = D$.

Solution: A is a symmetric matrix. A is orthogonally equivalent to a diagonal matrix.

We will now find the orthogonal matrix P and a diagonal matrix D such that $P^T AP = D$.

To find P:

To find P we first find an orthonormal basis of eigen vectors.

The characteristic equation is $|A-\lambda I| = 0$.

$\Rightarrow \begin{vmatrix}4-\lambda&2&2\\2&4-\lambda&2\\2&2&4-\lambda\end{vmatrix}=0$ on expansion we get the characteristic equation.

Or

trance of $A = 4+4+4=12, |A| = 4(16-4)-2(8-4)+2(4-8)$ i.e. $|A| = 48-8-8=32$

$M_{11} = \begin{vmatrix}4&2\\2&4\end{vmatrix} = 16-4=12$; $M_{22} = \begin{vmatrix}4&2\\2&4\end{vmatrix} = 16-4=12$; $M_{33} = \begin{vmatrix}4&2\\2&4\end{vmatrix} = 16-4=12$

So $M_{11}+M_{22}+M_{33} = 12+12+12=36$

Hence the characteristic equation is $\lambda^3 - (A$ trace of A) $\lambda^2 + (M_{11}+M_{22}+M_{33})\lambda - |A| = 0$

i.e. $f(\lambda) = \lambda^3 - 12\lambda^2 + 36\lambda - 32 = 0$

Here $f(2) = 2^3 - 12(4) + 36(2) - 32 = 8-48+72-32 = 0$

Hence one characteristic value is $\lambda = 2$.

Hence the other factor is

$\lambda^2 - 10\lambda + 16$

$\lambda = 2$	1	−12	36	−32
		2	−20	32
	1	-10	16	0

So $f(\lambda) = (\lambda - 2)(\lambda^2 - 10 + 16) = 0$

$\Rightarrow (\lambda - 2)(\lambda - 2)(\lambda - 8) = 0$

So $\lambda = 2, 2, 8$

To find a basis for the eigen space of A, corresponding to $\lambda = 2$. Then $[A - \lambda I] X = O$

$$\Rightarrow \begin{bmatrix} 4-2 & 2 & 2 \\ 2 & 4-2 & 2 \\ 2 & 2 & 4-2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = O$$

By $R_2 - R_1, R_3 - R_1$ we get

$$\begin{bmatrix} 2 & 2 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = O$$

$\Rightarrow 2x + 2y + 2z = 0$ ie $x + y + z = 0$ .... (1)

This system has two independent solutions

put $y = 1, z = 0$ i.e. $x = -y = -1$

$v_1 = (-1, 1, 0)$

We seek a second solution which is orthogonal to v_1. Let this solution be $v_2 = (a, b, c)$.

So $-a + b = 0$; and from (1) $a + b + c = 0$

using $a = b$ we get $2b + c = 0$ i.e. $c = -2b$

if $b = 1$, then $c = -2$ $a = 1$

So $v_2 = (a, b, c) = (1, 1, -2)$

Thus $v_1 = (-1, 1, 0)$ and $v_2 = (1, 1, -2)$ form an orthogonal basis for eigen space of $\lambda = 2$.

for $\lambda = 8, (A - \lambda I)(X) = O$

$$\Rightarrow \begin{bmatrix} 4-8 & 2 & 2 \\ 2 & 4-8 & 2 \\ 2 & 2 & 4-8 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = O$$

$$\Rightarrow \begin{bmatrix} -4 & 2 & 2 \\ 2 & -4 & 2 \\ 2 & 2 & -4 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = O$$

$2R_2 + R_1, 2R_3 + R_1$ gives

$$\begin{bmatrix} -4 & 2 & 2 \\ 0 & -6 & 6 \\ 0 & 6 & -6 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = O$$

$R_3 + R_2$ gives

$$\begin{bmatrix} -4 & 2 & 2 \\ 0 & -6 & 6 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = O$$

$$\Rightarrow -4x + 2y + 2z = 0$$

$$-6y + 6z = 0$$

or $-2x + y + z = 0$ and $-y + z = 0 \quad y = z$

So $2x = 2z \Rightarrow x = z$

put $z = 1,$ then $x = 1, y = 1$

This system gives the non zero solution.

$v_3 = (1,1,1)$ which is orthogonal to both v_1 and v_2.

Thus v_1, v_2, v_3 form a maximal set of non zero orthogonal vectors of A.

Normalize v_1, v_2, v_3 by dividing each with their corresponding lengths to obtain an orthonormal basis is u_1, u_2, u_3

$$u_1 = \frac{v_1}{\|v_1\|} = \frac{1}{\sqrt{2}}(-1,1,0)$$

$$u_2 = \frac{v_2}{\|v_2\|} = \frac{1}{\sqrt{6}}(1,1,-2)$$

$$u_3 = \frac{v_3}{\|v_3\|} = \frac{1}{\sqrt{3}}(1,1,1)$$

Aliter: we can apply Gram - Schmidth orthogonalisation process to find an orthonormal basis.

Let P be the matrix whose columns are u_1, u_2, u_3, Then $P = \begin{bmatrix} \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{bmatrix}$

and the diagonal matrix D which is formed with the characteristic roots given by

$$D = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 8 \end{bmatrix}$$

Note: In finding the basis for the eigen space corresponding $\lambda = 2,$ we get $x + y + z = 0$ ...... (1) by putting $y = 0,$ we get $x + z = 0$ when $z = 1, x = -1$

So $v_2 = (-1, 0, 1)$ $v_1 = (-1, 1, 0)$

So $\{(-1,1,0), (-1,0,1)\}$ is a basis of eigen space for $\lambda = 2,$ this set is not orthogonal. So we can apploy Gram-Sdmidt orthogonalisation process to obtain the orthogonal basis $\left\{(-1,1,0), \frac{-1}{2}(1,1,-2)\right\}$.

Find a orthogonal basis for f th eigen space for $\lambda = 8$, the union of the there two bases is the orthonormal basis. Normalising the vectors, we get the orthonormal basis.

16.14 Summary:

In this lesson we discussed about unitary operators, orthogonal operators, equivalent statements. Inverse of unitary operator, Matrices representing unitary and orthogonal transformations, Equivalent matrices.

16.15 Technical Terms:

The technical terms. we come across in this lesson are unitary operator, orthogonal operator, Isometry - Orthogonal matrix, unitary matrix, equivalent matrices, unitarily equivalent.

16.16 Model Questions:

1. Define unitary operator. If T is a unitary operator then show that $T^* = T^{-1}$.

2. If S and T are unitary operators then show that ST is unitary.

3. Let T be a unitary operator on an inner product space V, Let W be a finite dimensional T invariant subspace of V. Then prove that $W^{\perp}$ is T-invariant.

4. Let A be a complex n x n matrix, then A is normal if and only if A is unitarily equivalent to a diagonal matrix.

16.17 Exercise:

1. For each of the following matrices A; find an orthogonal or unitary matrix P and a diagonal matrix D. Such that $P^* AP = D$

i) $\begin{bmatrix} 0 & 2 & 2 \\ 2 & 0 & 2 \\ 2 & 2 & 0 \end{bmatrix}$ Ans: $P = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{bmatrix}$ and $D = \begin{bmatrix} -2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 4 \end{bmatrix}$

ii) $\begin{bmatrix} 5 & 4 \\ 1 & 2 \end{bmatrix}$ Ans : $P = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{5}{2} \\ \frac{-1}{\sqrt{2}} & \frac{5}{2} \end{bmatrix}$ and $D = \begin{bmatrix} 1 & 0 \\ 0 & 6 \end{bmatrix}$

2. Show that the matrices $\begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ and $\begin{bmatrix} 1 & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & -i \end{bmatrix}$ are unitarily equivalent.

3. Show that the matrices $\begin{bmatrix} 1 & 2 \\ 2 & i \end{bmatrix}$ and $\begin{bmatrix} i & 4 \\ 1 & 1 \end{bmatrix}$ are not unitarily equivalent.

4. Find the number and exhibit all 2 x 2 orthogonal matrices of the form $\begin{bmatrix} \frac{1}{3} & x \\ y & z \end{bmatrix}$

Ans: 4, $\begin{bmatrix} \frac{1}{3} & \sqrt{8}/3 \\ \sqrt{8}/3 & \frac{-1}{3} \end{bmatrix}, \begin{bmatrix} \frac{1}{3} & \sqrt{8}/3 \\ -\sqrt{8}/3 & \frac{1}{3} \end{bmatrix}, \begin{bmatrix} \frac{1}{3} & -\sqrt{8}/3 \\ \sqrt{8}/3 & \frac{1}{3} \end{bmatrix}$ and $\begin{bmatrix} \frac{1}{3} & -\sqrt{8}/3 \\ -\sqrt{8}/3 & \frac{-1}{3} \end{bmatrix}$

5. Prove that the following matrices are unitary.

i) $\begin{bmatrix} \frac{1}{2}(1+i) & \frac{i}{\sqrt{3}} & \frac{3+i}{2\sqrt{15}} \\ \frac{-1}{2} & \frac{1}{\sqrt{3}} & \frac{4+3i}{2\sqrt{15}} \\ \frac{1}{2} & \frac{-i}{\sqrt{3}} & \frac{5i}{2\sqrt{15}} \end{bmatrix}$

ii) $\begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}} \\ \frac{-i}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \end{bmatrix}$

iii) $\begin{bmatrix} \frac{1}{\sqrt{3}} & i\sqrt{\frac{2}{3}} \\ \frac{-i}{\sqrt{2/3}} & \frac{-1}{\sqrt{3}} \end{bmatrix}$

iv) $\frac{1}{2}\begin{bmatrix} (1+i) & -(1-i) \\ (1+i) & (1-i) \end{bmatrix}$

6. Prove the following matrices are orthogonal

i) $\begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{bmatrix}$

ii) $\begin{bmatrix} \frac{1}{\sqrt{14}} & \frac{2}{\sqrt{14}} & \frac{3}{\sqrt{14}} \\ \frac{3}{\sqrt{10}} & 0 & \frac{-1}{\sqrt{10}} \\ \frac{-1}{\sqrt{35}} & \frac{5}{\sqrt{35}} & \frac{-3}{\sqrt{35}} \end{bmatrix}$

7. Find an orthogonal matrix whose first row is

i) $\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}$ ii) a multiple of $(1,1,1)$

Ans : i) $\begin{bmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{-1}{\sqrt{5}} \end{bmatrix}$ ii) $\begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ \frac{2}{\sqrt{6}} & \frac{-1}{\sqrt{6}} & \frac{-1}{\sqrt{6}} \end{bmatrix}$

8. Find a unitary matrix whose first row is

i) a multiple of $(1,1-i)$ ii) $\left(\frac{1}{2}, \frac{1}{2}i, \frac{1}{2}-\frac{1}{2}i\right)$

Ans: i) $\begin{bmatrix} \frac{1}{\sqrt{3}} & (1-i)/\sqrt{3} \\ (1+i)/\sqrt{3} & \frac{-1}{\sqrt{3}} \end{bmatrix}$ ii) $\begin{bmatrix} \frac{1}{2} & \frac{1}{2}i & \frac{1}{2}-\frac{1}{2}i \\ i/\sqrt{2} & \frac{-1}{\sqrt{2}} & 0 \\ \frac{1}{2} & \frac{-1}{2}i & \frac{-1}{2}+\frac{1}{2}i \end{bmatrix}$

9. Find a 3×3 orthogonal matrix P whose first two rows are multiples of $u=(1,1,1)$ and $v=(1,-2,3)$ respectively.

Ans : $P = \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{14}} & \frac{-2}{\sqrt{14}} & \frac{3}{\sqrt{14}} \\ \frac{5}{\sqrt{38}}, & \frac{-2}{\sqrt{38}}, & \frac{-3}{\sqrt{38}} \end{bmatrix}$

10. Real matrices A and B are said to be orthogonally equivalent if there exists an orthogonal matrix P such that $B = P^T AP$. Show that this relation is an equivalence relation.

11. Prove that if A and B are unitarily equivalent matrices, then A is positive definite if and only if B is positive definite.

12. Let U be a unitary operator on an inner product space V, and let W be a finite dimensional U - Invariant subspace of V. Prove that $U(W) = W$.

13. Let A and B are n x n matrices that are unitarily equivalent then prove that $tr(A^*A) = tr(B^*B)$

16.18 Reference Books:

1. Linear Algebra 4th edition Stephen H. Friedberg, Arnol J. Insel, Lawrence E. Spence.

2. Schaum's Out lines Beginning Linear Algebra, Seymour Lipschutz.

3. Topics in Algebra I.N. Herstein

4. Linear Algebra P.P. Gupta Ph.D. and S.K. Sharma Ph.D.

- A. Mallikharjana Sarma