R88

DIRECTORATE OF DISTANCE EDUCATION

MATHEMATICS

B.Sc., Third Year

PAPER - VII

LINEAR ALGEBRA AND NUMBER SYSTEM

Madurai Kamaraj University
Madurai - 625 021

Lesson printed by

Jagan Computers

4/ 107 A, R.R. Compound,
Viraganoor, Teppakkulam
Madurai.

Copies 2000 Year : 2007-2008

SYLLABUS

III YEAR – MAJOR – PAPER VII

LINEAR ALGEBRA AND NUMBER SYSTEM

LINEAR ALGEBRA

Unit - 1 Abstract vector spaces - Elementary properties - Linear span - Linear independence and dependence - Basis and dimension.

Unit - 2 Subspace - Linear transformation.

Unit - 3 Matrices - Algebra of matrices - transpose, ad joint, inverse, hermitian, symmetric, skew symmetric, singular or orthogonal and conjugate matrices.

Unit - 4 Inner product - orthonormal basis - orthogonalisation - Gram schmidt's method.

Unit - 5 Cayley - Hamilton theorem - eigen value and eigen vectors, column rank-row rank and rank.

Unit - 6 Reduction to normal forms - similar and congruent matrices - solution of a system of linear equation using matrices and determinant.

NUMBER SYSTEM

Unit - 7 Theory of numbers - Prime and composite number - The sieve of Eratosthenes - Divisors of a given number - simple problems.

Unit - 8 Euler's function - Integral part of a real number - simple problems - numbers in Arithmetic progression - Fermat's theorem - simple problems - product of r consecutive integers is divisible by r!

Unit - 9 Congruence - Criteria of divisibility of a number - simple problems - numbers in Arithmetic progression - Fermat's theorem - simple problems.

Unit - 10 Generalization of Fermat's theorem - Wilson's theorem - Lagrange's theorem - simple problems.

Books Recommended

1. Modern Analysis : Arumugam and others
2. Algebra Vol. I & II : T.K.Manicavachagam Pillai and Narayanan

SCHEME

LINEAR ALGEBRA AND NUMBER SYSTEM

LINEAR ALGEBRA

Lesson Compiled by

K.S. RAJA RAJESWARI, M.Sc., M.Phil., B.Ed., M.A.(JMC),
Lecturer in Mathematics,
E.M.G. Yadava Women's College,
Madurai - 14.

MODEL QUESTION PAPER - I

LINEAR ALGEBRA AND NUMBER SYSTEM

Time : 3 Hours **Max : 100 marks**

SECTION A

Answer any **8 questions.** 5×8 = 40 marks

1. Prove that R×R is not a vector space over R.
2. Define Inner Product Space. Give an example.
3. If S is any subset of V, prove that $S^{\perp}$ is a subspace of V.
4. Define Elementary Transformations of matrices. Given an example.
5. If $S = \{V_1, V_2, \ldots\ldots, V_n\}$ is an Orthogonal set of non-zero vectors in an inner product space V, $v \in V$ and if $v = \alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n$ prove that $\alpha_k = \dfrac{\langle v, v_k \rangle}{\|v_k\|^2}$.
6. Verify Cayley-Hamilton theorem for the matrix $\begin{bmatrix} 2 & -1 & 1 \\ 1 & 0 & 2 \\ 3 & -1 & 3 \end{bmatrix}$
7. If A and B are similar matrices prove that their determinants are same.
8. Find the rank of the matrix $\begin{bmatrix} 1 & 2 & 3 & 1 \\ 2 & 4 & 6 & 2 \\ 1 & 2 & 3 & 2 \end{bmatrix}$
9. Find the number and the sum of all the divisorsof 1458.
10. Prove that the sum of integers less than N and prime to it $\frac{1}{2} N\phi(N)$.
11. show that $3^{4n+2}+5^{2n+1}$ is divisible by 14.
12. Show that $x^3 - x$ is divisible by 6.

SECTION B

Answer any **6 questions.** 6×10 = 60 marks

13. If W is a subspace of a finite dimensional vector space V over a field F, prove that dim(V/W) = dim V – dim W.

14. Explain Gram-Schmidt Orthogonalisation process. Using it find the orthonormal basis of $V_3(R)$ with the basis $\{(1, -1, 0), (2, -1, -2), (1, -1, -2)\}$.

15.a) Prove that every square matrix satisfies its characteristic equation.

b) Find the eigen roots and eigen vectors of the matrix A given below.

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

16. Prove the below the two inequalities in an inner product spaces.

i) $|\langle x, y\rangle| \leq \|x\| \ \|y\|$

ii) $\|x+y\| \leq \|x\| + \|y\|$

17. Find the value of $\phi(N)$

18. State and Prove Wilson's theorem.

19. If m and n are prime numbers show that $m^{n-1}+n^{m-1}-1 \equiv 0 \pmod{mn}$

20. Prove that $712!+1 \equiv 0 \pmod{719}$

MODEL QUESTION PAPER - II

LINEAR ALGEBRA AND NUMBER SYSTEM

Time : 3 Hours **Max : 100 marks**

SECTION A

Answer any **8 questions.** 5×8 = 40 marks

1. Prove that the vectors (1, 2, –3), (2, 5, 1) and (–1, 1, 4) form a basis for $V_3(R)$.

2. Prove that any subset of linearly independent set is linearly independent.

3. Let V be a vector space over F. Prove that

i) $\alpha(0) = 0 \quad \forall \alpha \in F$

ii) $0v = 0 \quad \forall v \in V$

iii) $(-\alpha)v = \alpha(-v) = -(\alpha v) \quad \forall \alpha \in F, v \in V$

4. If A & B are Orthogonal matrices of same order prove that A^T and AB are Orthogonal matrics.

5. If V is inner product space, prove that

$||u+v||^2+||u-v||^2 = 2(||u||^2+||v||^2)$, u, v, $\in V$

6. Find the rank of the matrix A given below

$$A=\begin{pmatrix}4 & 2 & 1 & 3\\ 6 & 3 & 4 & 7\\ 2 & 1 & 0 & 7\end{pmatrix}$$

7. Show that the map $T:R^2\to R^2$ defined by $T(x, y) = (x, -y)$ is a linear transformation.

8. Define similar matrices. If $A=\begin{pmatrix}2 & -3\\ 3 & 1\end{pmatrix}$ find a matrix similar to A.

9. Find the highest power of 5 contained in 1000!

10. If a, b, c,..... are the different prime factors of N, then show that the sum of all the numbers less than N and prime to N is $\frac{N^2}{2}\left(1-\frac{1}{a}\right)\left(1-\frac{1}{b}\right)\left(1-\frac{1}{c}\right).....$

11. Show that x^5-x is divisible by 30.

12. Show that $7^{2n}+16n-1 \equiv 0 \pmod{64}$

SECTION B

Answer any **6 questions.** $6\times10 = 60$ **marks**

13. If S is a non empty subset of a vector space V over a field F, prove that

    i) L(S) is a subspace of V

    ii) $S\subseteq L(S)$

    iii) L(S) is the smallest subspace of V containing S.

14. Prove that any vector space of dimension n over a field F is isomorphic to $V_n(F)$.

15. Using Cayley-Hemilton theorem, find the inverse of the matrix B given below.

$$B=\begin{pmatrix}3 & 3 & 4\\ 2 & -3 & 4\\ 0 & -1 & 1\end{pmatrix}$$

16. If A and B are two m × n matrices prove that

i) $(A^T)^T = A$

ii) $(A+B)^T = A^T + B^T$

17. Reduce the matrix $\begin{pmatrix} 2 & -2 & 0 & 6 \\ 4 & 2 & 0 & 2 \\ 1 & -1 & 0 & 3 \\ 1 & -2 & 1 & 2 \end{pmatrix}$ to its normal form.

18.a) State and prove a necessary and sufficient condition for a non-empty subset w of u vector space V to be a subspace of V over F.

b) If A and B are two subspaces of a vector spaces V over F, prove that $A \cap B$ is a subspace of V.

19. Prove that $\{(1, -1, 0), (2, -1, -2), (1, -1, -2)\}$ is a basis of R^3. Also find the orthonormal basis from this basis.

20. Show that $28!+233 \equiv 0 \pmod{899}$

21. State and prove Fermat's theorem.

22. Show that every integer which is a perfect cube is of the form 7P or $7P \pm 1$.

1.1. VECTOR SPACES

Introduction :

In this chapter we introduce another algebraic system known as vector spaces. The idea of a vector arises in the study of various physical applications. Many physical entities like mass, temperature etc. are characterised in terms of a real number and are called scalars. Other physical entities such as the velocity of a particle or force acting at a point are determined only when both magnitude and direction are specified. Such entities are called **vectors.**

Definition and Examples :

Definition :

A non empty set V is said to be a vector space over a field F if

(i) V is an abelian group under an operation called addition which we denote by +.

(ii) For every $\alpha \in F$ and $v \in V$, there is defined an element αv in V subject to the following conditions.

a) $\alpha(u+v) = \alpha u+\alpha v$ for all $u, v \in V$ and $\alpha \in F$.

b) $(\alpha+\beta)u = \alpha u+\beta u$ for all $u \in V$ and $\alpha, \beta \in F$

c) $\alpha(\beta u) = (\alpha\beta)u$ for all $u \in V$ and $\alpha, \beta \in F$.

d) $1.u = u$ for all $u \in V$.

Remarks :

1. The elements of F are called **scalars** and the elements of V are called **vectors.**

2. The rule which associates with each scalar $\alpha \in F$ and a vector $v \in V$, a vector αv is called the **scalar multiplication.** Thus a scalar multiplication gives rise to a function from $F \times V \to V$ defined by $(\alpha, v) \to \alpha v$.

Examples :

1. $R \times R$ is a vector space over R under addition and scalar multiplication defined by

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

and

$$\alpha(x_1, x_2) = (\alpha x_1, \alpha x_2)$$

Proof :

Clearly the binary operation + is commutative and associative and (0, 0) is the zero element.

The inverse of (x_1, x_2) is $(-x_1, -x_2)$.

Hence $(R\times R, +)$ isan abelian group.

Now, let $u = (x_1, x_2)$ and $v = (y_1, y_2)$ and let $\alpha, \beta \in R$.

Then
$$\begin{aligned}\alpha(u+v) &= \alpha[(x_1, x_2)+(y_1, y_2)]\\ &= \alpha[x_1+y_1, x_2+y_2]\\ &= (\alpha x_1+\alpha y_1, \alpha x_2+\alpha y_2)\\ &= (\alpha x_1, \alpha x_2)+(\alpha y_1, \alpha y_2)\\ &= \alpha(x_1, x_2)+\alpha(y_1, y_2)\\ &= \alpha u+\alpha v\end{aligned}$$

Now,
$$\begin{aligned}(\alpha+\beta)u &= (\alpha+\beta)(x_1, x_2)\\ &= ((\alpha+\beta)x_1, (\alpha+\beta)x_2)\\ &= (\alpha x_1+\beta x_1, \alpha x_2+\beta x_2)\\ &= (\alpha x_1, \alpha x_2)+(\beta x_1, \beta x_2)\\ &= \alpha(x_1, x_2)+\beta(x_1, x_2)\\ &= \alpha u+\beta u\end{aligned}$$

Also,
$$\begin{aligned}\alpha(\beta u) &= \alpha(\beta(x_1, x_2))\\ &= \alpha(\beta x_1, \beta x_2)\\ &= (\alpha\beta x_1, \alpha\beta x_2)\\ &= (\alpha\beta)(x_1, x_2)\\ &= (\alpha\beta)u\end{aligned}$$

Obviously, $1.u = u$

$\therefore$ $R\times R$ is a vector space over R.

2. $R^n = \{(x_1, x_2,.....x_n)/x_i \in R,\ 1 \le i \le n\}$. Then R^n is a vector space over R under addition and scalar multiplication defined by

$$(x_1,x_2,...,x_n)+(y_1,y_2,...,y_n) = (x_1+y_1, x_2+y_2,...,x_n+y_n)$$

and $$\alpha(x_1,x_2,.....,x_n) = (\alpha x_1,\alpha x_2,.....,\alpha x_n)$$

Proof :

Clearly the binary operation + is commutative and associative.

(0, 0,.....,0) is a zero element.

The inverse of $(x_1,x_2,.....,x_n)$ is $(-x_1,-x_2,.....,-x_n)$

Hence $(R^n, +)$ is an abelian group.

Now, let $u = (x_1, x_2,.....x_n)$

and $v = (y_1, y_2,....,y_n)$ and let $\alpha, \beta \in R$

Then
$$\begin{aligned}\alpha(u+v) &= \alpha[(x_1,x_2,.....,x_n)+(y_1,y_2,.....,y_n)]\\ &= \alpha(x_1+y_1,x_2+y_2,.....,x_n+y_n)\\ &= (\alpha x_1+\alpha y_1, \alpha x_2+\alpha y_2,....,\alpha x_n+\alpha y_n)\\ &= (\alpha x_1,\alpha x_2,.....,\alpha x_n)+(\alpha y_1,\alpha y_2,.....,\alpha y_n)\\ &= \alpha(x_1,x_2,....,x_n)+\alpha(y_1,y_2,.....,y_n)\\ &= \alpha u+\alpha v\end{aligned}$$

Similarly $(\alpha+\beta)u = \alpha u+\beta u$

and $\alpha(\beta u) = (\alpha\beta)u$

Obviously, $1.u = u$

$\therefore$ R^n is vector space over R.

3. C is a vector space over the field R.

Here addition is the usual addition in C and the scalar multiplication is the usual multiplication of a real number and a complex number.

i.e., $(x_1+ix_2)+(y_1+iy_2) = (x_1+y_1)+i(x_2+y_2)$

and $\alpha(x_1+ix_2) = \alpha x_1+i\alpha x_2$

Proof :

Clearly (C, +) is an abelian group. Also the remaining axioms of a vector space are true since the scalars and vectors involved are complex numbers and further the operations are usual addition and multiplication. Hence C is a vector space over R.

4. Let $V = \{a + b\sqrt{2}/a, b \in Q\}$. Then V is a vector space over Q under addition and multiplication.

Proof :

Obviously V is an abelian group under usual addition.

The remaining axioms of a vector space are true since the scalars and vectors are real numbers and the operations are usual addition and multiplication. Hence V is a vector space over Q.

5. The set $M_2(R)$ of all 2×2 matrices is a vector space over R under matrix addition and scalar multiplication defined by

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

6. Let V={0}. V is a vector space over any field F under the obvious operations of addition and scalar multiplication.

7. R is not a vector space over C.

Clearly (R, +) is an abelian group.

But the scalar multiplications is not defined, for if $\alpha = a+ib \in C$ and $u \in R$, then $\alpha u = au+ibu \notin R$.

∴ R is not vector space over C.

8. Consider R×R with usual addition. We define scalar multiplication by $\alpha(x, y) = (\alpha x, \alpha^2 y)$.

Then R×R is not a vector space over R.

Clearly R×R with usual addition is an abelian group.

$$(\alpha+\beta)(x, y) = ((\alpha+\beta)x, (\alpha+\beta)^2y)$$
$$= (\alpha x+\beta x, \alpha^2y+\beta^2y+2\alpha\beta y)$$

Also, $\alpha(x, y)+\beta(x,y) = (\alpha x, \alpha^2y)+(\beta x, \beta^2y)$
$$= (\alpha x+\beta x, \alpha^2y+\beta^2y)$$

Hence $(\alpha+\beta)(x, y) \neq \alpha(x, y)+\beta(x, y)$

$\therefore$ $R\times R$ is not a vector space over R.

1.2. ELEMENTARY PROPERTIES

Theorem 1.1 :

Let V be a vector space over a field F. Then

(i) $\alpha 0 = 0$ for all $\alpha \in F$

(ii) $0v = 0$ for all $v \in V$

(iii) $(-\alpha)v = \alpha(-v) = -(\alpha v)$ for all $\alpha \in F$ and $v \in V$.

(iv) $\alpha v = 0 \Rightarrow \alpha = 0$ (or) $v = 0$

Proof :

(i) $\alpha 0 = \alpha(0+0) = \alpha 0+\alpha 0$

Hence $\alpha 0 = 0$

(ii) $0v = (0+0)v = 0v+0v$

Hence $0v = 0$

(iii) $0 = [\alpha+(-\alpha)]v = \alpha v+(-\alpha)v$

Hence $(-\alpha)v = -(\alpha v)$

Similarly $\alpha(-v) = -(\alpha v)$

Hence $(-\alpha)v = \alpha(-v) = -(\alpha v)$

(iv) Let $\alpha v = 0$. If $\alpha = 0$, there is nothing to prove.

$\therefore$ Let $\alpha \neq 0$.

Then $\alpha^{-1} \in F$

Now, $v = 1v = (\alpha^{-1}\alpha)v = \alpha^{-1}(\alpha v) = \alpha^{-1}0 = 0$

Theorem 1.2 :

Let V be a vector space over a field F. Then

(i) $\alpha(v_1-v_2) = \alpha v_1-\alpha v_2 \quad \forall \alpha \in F, v_1,v_2 \in V$

(ii) $\alpha v_1 = \alpha v_2$ and $\alpha \neq 0 \Rightarrow v_1=v_2$

(iii) $\alpha v = \beta v$ and $v \neq 0 \Rightarrow \alpha=\beta$

Proof :

(i) $\alpha(v_1-v_2) = \alpha[v_1+(-v_2)]$

$= \alpha v_1+\alpha(-v_2)$

$= \alpha v_1-\alpha v_2$ $\quad (\because (-\alpha)v = -(\alpha v))$

(ii) Since $\alpha v_1 = \alpha v_2$

$\alpha v_1-\alpha v_2 = 0$

$\alpha(v_1-v_2) = 0$

Also $\alpha \neq 0$

$\therefore$ $v_1-v_2 = 0$ $\quad (\because \alpha v=0 \Rightarrow \alpha=0$ or $v=0)$

$v_1 = v_2$

(iii) since $\alpha v = \beta v$

$\alpha v-\beta v = 0$

i.e., $(\alpha-\beta)v = 0$

Also $v \neq 0$

$\therefore$ $\alpha-\beta = 0$ $\quad (\because \alpha v=0 \Rightarrow \alpha=0$ (or) $v=0)$

i.e., $\alpha = \beta$

$\therefore$ $\alpha v = \beta v$

and $v \neq 0 \Rightarrow \alpha=\beta.$

1.3. LINEAR SPAN

Definition :

Let V be a vector space over a field F. Let $v_1,v_2,.....v_n \in V$. Then an element of the form $\alpha_1v_1+\alpha_2v_2+.....+\alpha_nv_n$ where $\alpha_i \in F$ is called a **linear combination** of the vectors $v_1,v_2,.....,v_n$.

Definition :

Let S be a non-empty subset of a vector space V. Then the set of all linear combinations of finite sets of elements of S is called the linear span of S and is denoted by L(S).

Note :

Any element of L(S) is of the form.

$\alpha_1 v_1+\alpha_2 v_2+.....+\alpha_n v_n$ where $\alpha_1,\alpha_2,.....,\alpha_n \in F$.

Theorem 1.3 :

Let V be a vector space over a field F and S be a non-empty subset of V. Then

(i) L(S) is a subspace of V

(ii) $S \subseteq L(S)$

(iii) If W is any subspace of V such that $S \subseteq W$, then $L(S) \subseteq W$. (i.e.,) L(S) is the smallest subspace of V containing S.

Proof :

(i) Let $v,w \in L(S)$ and $\alpha,\beta \in F$

Then $v = \alpha_1 v_1+\alpha_2 v_2+.....+\alpha_n v_n$ where $v_i \in S$ and $\alpha_i \in F$.

Also, $w = \beta_1 w_1+\beta_2 w_2+.....+\beta_m w_m$ where $w_j \in S$ and $\beta_j \in F$.

Now, $\alpha v+\beta w = \alpha(\alpha_1 v_1+\alpha_2 v_2+.....+\alpha_n v_n)+\beta(\beta_1 w_1+\beta_2 w_2+...+\beta_m w_m)$

$= (\alpha\alpha_1)v_1+.....+(\alpha\alpha_n)v_n+(\beta\beta_1)w_1+....+(\beta\beta_m)w_m$

$\therefore$ $\alpha v+\beta w$ is also a linear combination of a finite number of elements of S.

Hence $\alpha v+\beta w \in L(S)$.

$\therefore$ L(S) is a subspace of V.

(ii) Let $u \in S$

Then $u=1u \in L(S)$

Hence $S \subseteq L(S)$

(iii) Let W be any subspace of V such that $S \subseteq W$. We claim that $L(S) \subseteq W$.

Let $u \in L(S)$

Then $u = \alpha_1 u_1 + \alpha_2 u_2 + \ldots\ldots + \alpha_n u_n$ where $u_i \in S$ and $\alpha_i \in F$

Since $S \subseteq W$ we have $u_1, u_2, \ldots, u_n \in W$

$\therefore \alpha_1 u_1 + \alpha_2 u_2 + \ldots\ldots + \alpha_n u_n \in W$ (since W is a subspace of V)

$\therefore u \in W$

Hence $L(S) \subseteq W$.

Note :

L(S) is called the subspace **spanned (generated)** by the set S.

Examples :

1. In $V_3(R)$ let $e_1 = (1,0,0)$, $e_2 = (0, 1, 0)$ and $e_3 = (0, 0, 1)$

(a) Let $S = \{e_1, e_2\}$

Then $L(S) = \{\alpha e_1 + \beta e_2 / \alpha, \beta \in R\} = \{(\alpha, \beta, 0) / \alpha, \beta \in R\}$

(b) Let $S = \{e_1, e_2, e_3\}$. Then

$$L(S) = \{\alpha e_1 + \beta e_2 + \gamma e_3 / \alpha, \beta, \gamma \in R\}$$
$$= \{(\alpha, \beta, \gamma) / \alpha, \beta, \gamma \in R\}$$
$$= V_3(R)$$

Thus $V_3(R)$ is spanned by $\{e_1, e_2, e_3\}$

2. In $V_n(R)$, let $e_1 = (1,0,0,\ldots\ldots 0)$.

$e_2 = (0,1,0,\ldots,0), \ldots\ldots,\ e_n = (0,0,\ldots\ldots,1)$

Let $S = \{e_1, e_2, \ldots\ldots, e_n\}$

Then $L(S) = \{\alpha_1 e_1 + \alpha_2 e_2 + \ldots\ldots + \alpha_n e_n / \alpha_i \in R\}$

$$= \{(\alpha_1, \alpha_2, \ldots\ldots, \alpha_n) / \alpha_i \in R\}$$
$$= V_n(R)$$

Thus $V_n(R)$ is spanned by $\{e_1, e_2, \ldots, e_n\}$

Theorem 1.4 :

Let V be a vector space over a field F.

Let $S,T \subseteq V$. Then

(a) $S \subseteq T \Rightarrow L(S) \subseteq L(T)$

(b) $L(S \cup T) = L(S)+L(T)$

(c) $L(S) = S$ iff S is a subspace of V.

Proof :

(a) Let $S \subseteq T$. Let $s \in L(S)$

Then $s = \alpha_1 s_1 +_2 s_2 + \ldots\ldots + \alpha_n s_n$ where $s_i \in S$ and $\alpha_i \in F$.

Now, since $S \subseteq T$, $s_i \in T$

Hence $\alpha_1 s_1 + \alpha_2 s_2 + \ldots\ldots + \alpha_n s_n \in L(T)$

Thus $L(S) \subseteq L(T)$

(b) Let $s \in L(S \cup T)$

Then $s = \alpha_1 s_1 + \alpha_2 s_2 + \ldots + \alpha_n s_n$ where $s_i \in S \cup T$ and $\alpha_i \in F$

Without loss of gnerality we can assume that

$s_1, s_2, \ldots, s_m \in S$ and $s_{m+1}, \ldots\ldots, s_n \in T$

Hence $\alpha_1 s_1 + \alpha_2 s_2 + \ldots\ldots\ldots + \alpha_m s_m \in L(S)$

and $\alpha_{m+1} s_{m+1} + \ldots\ldots + \alpha_n s_n \in L(T)$

$\therefore$ $s = (\alpha_1 s_1 + \ldots\ldots\ldots + \alpha_m s_m) + (\alpha_{m+1} s_{m+1} + \ldots\ldots + \alpha_n s_n) \in L(S)+L(T)$

Hence $L(S \cup T) \subseteq L(S)+L(T)$

Also by (a) $L(S) \subseteq L(S \cup T)$

and $L(T) \subseteq L(S \cup T)$

Hence $L(S)+L(T) \subseteq L(S \cup T)$

Hence $L(S \cup T) = L(S)+L(T)$

(c) Let L(S) = S. By theorem 1.3. L(S) = S is a subspace of V.

Conversely, Let S be a subspace V. Then the smallest subspace containing S is S itself.

Hence L(S) = S.

Corollary :

L(L(S)) = L(S).

Examples :

1. C is spanned by {1, i} where $i = \sqrt{-1}$.

For, if S = {1, i}, then L(S) = {a1+bi/a,b∈R}=C

2. The vectors $v_1 = (1, 2, 3)$, $v_2 = (0, 1, 2)$ and $v_3 = (0, 0, 1)$ generate R^3.

For, let $S = \{v_1, v_2, v_3\}$

To prove $L(S) = R^3$

We know $L(S) \subseteq R^3$

Therefore to prove $R^3 \subseteq L(S)$

Let v = (a, b, c) be any element of R^3.

To prove $v \in L(S)$

(i.e.,) v is a linear combination of v_1, v_2, v_3.

Put $(a, b, c) = xv_1+yv_2+zv_3$ where x, y, z are suitable scalars to be found.

Then $(a, b, c) = x(1, 2, 3) + y(0, 1, 2) + z(0, 0, 1)$

$= (x, 2x+y, 3x+2y+z)$

∴ x = a, 2x+y = b, 3x+2y+z = c.

These equations are consistent and so have a solution, namely, x=a, y=b–2a, z=c–2b+a.

Thus $v=(a, b, c) = av_1+(b-2a)v_2+(c-2b+a)v_3 \in L(S)$

∴ $R^3 \subseteq L(S)$ and hence $L(S)=R^3$.

Exercises :

1. Find L(S) in the following cases.

   (a) $S = \{(1, 0), (0,1)\}$ in $V_2(R)$

   (b) $S = \{(1,0,0), (2,0,0), (3,0,0)\}$ in $V_3(R)$

   (c) $S = \{(1,2,3), (2,3,1), (3,1,2)\}$ in $V_3(R)$

   (d) $S = \{(2,0)\}$ in $V_2(R)$

   (e) $S = \{1, x, x^2, \ldots\ldots, x^n\}$ in $R[x]$

   (f) $S = \{(1,2,3)\}$ in $Z_5 \times Z_5 \times Z_5$ over Z_5

   (g) $S = \{(0,1,2), (1,2,0)\}$ in $Z_3 \times Z_3 \times Z_3$ over Z_3

   (h) $S = \left\{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}\right\}$ in $M_2(R)$

2. Let V be a vector space over a field F and $S=\{v_1, v_2, \ldots v_n\} \subseteq V$

   Let $S_1 = \{\alpha_1 v_1, \alpha_2 v_2, \ldots\ldots, \alpha_n v_n\}$ where $\alpha_i \in F - \{0\}$

   Let $S_2 = \{v_1 + \alpha v_2, v_2, v_3, \ldots\ldots v_n\}$ where $\alpha \in F$

   Show that $L(S) = L(S_1) = L(S_2)$

3. Show that in $V_2(R)$

   $(3, 7) \in L(\{(1,2), (0,1)\})$

Answers :

1. (a) $V_2(R)$ (b) $\{(x,0,0) / x \in R\}$

   (c) $V_3(R)$ (d) $\{(x,0) / x \in R\}$

   (e) The set of all polynomials of degree $\leq n$ and zero polynomial.

   (f) $\{(0,0,0), (1,2,3), (2,4,1), (3,1,4), (4,3,2)\}$

   (g) $\{(0,0,0), (0,1,2), (1,2,0), (1,0,2), (0,2,1,), (2,1,0), (2,0,1), (1,1,1), (2,2,2)\}$

   (h) $\left\{\begin{bmatrix} a & 0 \\ b & 0 \end{bmatrix} / a, b \in R\right\}$

1.4. LINEAR INDEPENDENCE AND DEPENDENCE

In $V_3(R)$, let $S = \{e_1, e_2, e_3\}$.

We have seen that $L(S) = V_3(R)$

Thus S is a subset of $V_3(R)$ which spans the whole space $V_3(R)$.

Definition :

Let V be a vector space over a field F. V is said to be **finite dimensional** if there exists a finite subset S of V such that $L(S) = V$.

Examples :

1. $V_3(R)$ is a fintie dimensional vector space.

2. $V_n(R)$ is a finite dimensional vector space.

   Since $S=\{e_1, e_2,.....,e_n\}$ is a finite subset of $V_n(R)$ such that $L(S)=V_n(R)$. In general if F is any field $V_n(F)$ is a finite dimensional vector space over F.

3. Let V be the set of all polynomials in F[x] of degree $\leq n$. Let $S = \{1,x,x^2,....,x^n\}$

   Then $L(S) = V$ and hence V is finite dimensional.

4. C is a finite dimensional vector space over R, Since $L(\{1, i\}) = C$.

5. In $M_2(R)$ consider the set S consisting of the matrices

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

$$C = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}; \qquad D = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\text{Then } \begin{bmatrix} a & b \\ c & d \end{bmatrix} = aA+bB+cC+dD$$

Hence $L(S) = M_2(R)$. So that $M_2(R)$ is finite dimensional.

Note :

All the vector spaces we have considered above are finite dimensional. However there are vector spaces which cannot be spanned by a finite number of vectors. For

example, consider R[x]. Let S be any finite subset of R[x]. Let f be a polynomial of maximum degree in S. Let deg f = n. Then any element of L(S) is a polynomial of degree $\leq$ n and hence L(S) $\neq$ R[x]. Thus R[x] is not finite-dimensional.

In this chapter, we consider all the vector spaces are finite dimensional.

Although we have defined what is meant by a finite dimensional space we have not yet defined what is meant by the dimension of a vector space. We now proceed to introduce the concepts necessary to define the dimension of a finite dimensional vector space.

Consider the vectors $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$, $e_3 = (0, 0, 1)$ in $V_3(R)$

Suppose that $\alpha_1 e_1+\alpha_2 e_2+\alpha_3 e_3 = 0$

Then $(\alpha_1, 0, 0)+(0, \alpha_2, 0)+(0, 0, \alpha_3) = (0, 0, 0)$

$\therefore (\alpha_1, \alpha_2, \alpha_3) = (0, 0, 0)$

$\therefore \alpha_1 = \alpha_2 = \alpha_3 = 0$

(i.e.,) $\alpha_1 e_1+\alpha_2 e_2+\alpha_3 e_3 = 0$ iff $\alpha_1 = \alpha_2 = \alpha_3 = 0$

Thus a linear combination of the vectors, e_1, e_2 and e_3 will yield the zero vector iff all the coefficients are zero.

Definition :

Let V be a vector space over a field F. A finite set of vectors $v_1, v_2,....,v_n$ in V is said to be **linearly independent** if

$$\alpha_1 v_1+\alpha_2 v_2+.....+\alpha_n v_n = 0$$

$$\Rightarrow \alpha_1 = \alpha_2 = ... = \alpha_n = 0$$

If $v_1, v_2,.....,v_n$ are not linearly independent, then they are said to be **linearly dependent.**

Note :

If $v_1, v_2,.....,v_n$ are linearly dependent, then there exists scalars $\alpha_1,\alpha_2,....,\alpha_n$ not all zero, such that $\alpha_1 v_1+\alpha_2 v_2+.....+\alpha_n v_n = 0$.

Examples :

1. In $V_n(F)$, $\{e_1, e_2, ...e_n\}$ is a linearly independent set of vectors, for,

$$\alpha_1e_1+\alpha_2e_2+.....+\alpha_ne_n = 0$$

$$\Rightarrow \alpha_1(1,0,....,0)+\alpha_2(0,1,.....,0)+....+\alpha_n(0,0,.....1) = (0, 0,.....,0)$$

$$\Rightarrow (\alpha_1,\alpha_2,.....\alpha_n) = (0, 0,.....,0)$$

$$\Rightarrow \alpha_1 = \alpha_2 = ... = \alpha_n = 0$$

2. In $V_3(R)$ the vectors (1,2,1), (2,1,0) and (1,–1,2) are linearly independent. For, let $\alpha_1(1,2,1)+\alpha_2(2,1,0)+\alpha_3(1,-1,2) = (0,0,0)$

$\therefore (\alpha_1+2\alpha_2+\alpha_3, 2\alpha_1+\alpha_2-\alpha_3, \alpha_1+2\alpha_3) = (0, 0, 0)$

$\therefore$ $\alpha_1+2\alpha_2+\alpha_3 = 0$ --------(1)

$2\alpha_1+\alpha_2-\alpha_3 = 0$ --------(2)

$\alpha_1+2\alpha_3 = 0$ --------(3)

Solving equations (1), (2) and (3) we get $\alpha_1 = \alpha_2 = \alpha_3 = 0$

$\therefore$ The given vectors are linearly independent.

3. In $V_3(R)$ the vectors $v_1 = (-1, 2, 1)$ and $v_2 = (3, 1, -2)$ are linearly independent.

For $\alpha_1v_1+\alpha_2v_2 = 0$

$\alpha_1(-1, 2, 1) + \alpha_2(3, 1, -2) = (0, 0, 0)$

$(-\alpha_1+3\alpha_2, 2\alpha_1+\alpha_2, \alpha_1-2\alpha_2) = (0, 0, 0)$

$\therefore$ $-\alpha_1+3\alpha_2 = 0$ --------(1)

$2\alpha_1+\alpha_2 = 0$ --------(2)

$\alpha_1-2\alpha_2 = 0$ --------(3)

Solving (1), (2) and (3), we get $\alpha_1 = \alpha_2 = 0$

$\therefore$ $\{v_1, v_2\}$ is linearly independent.

4. In $V_3(R)$ the vectors (1, 4, –2), (–2, 1, 3) and (–4, 11, 5) are linearly dependent. For, let $\alpha_1(1, 4, -2)+\alpha_2(-2,1,3)+\alpha_3(-4, 11, 5) = (0, 0, 0)$

$\therefore$ $\alpha_1-2\alpha_2-4\alpha_3 = 0$ --------(1)

$4\alpha_1+\alpha_2+11\alpha_3 = 0$ --------(2)

$-2\alpha_1+3\alpha_2+5\alpha_3 = 0$ --------(3)

From (1) & (2)

$$\frac{\alpha_1}{-18}=\frac{\alpha_2}{-27}=\frac{\alpha_3}{9}=K \text{ (say)}$$

$\therefore$ $\alpha_1 = -18K$, $\alpha_2 = -27K$, $\alpha_3 = 9K$.

These values of α_1, α_2 and α_3 for any K satisfy (3) also

Taking K = 1 we get

$\alpha_1 = -18$, $\alpha_2 = -27$, $\alpha_3 = 9$ as a non-trivial solution.

Hence the three vectors are linearly dependent.

5. The vectors $v_1 = (1, -2, 1)$, $v_2 = (2, 1, -1)$ and $v_3 = (7, -4, 1)$ are linearly dependent in R^3.

For let $\alpha_1v_1+\alpha_2v_2+\alpha_3v_3 = 0$

Then $\alpha_1(1, -2, 1)+\alpha_2(2, 1, -1)+\alpha_3(7, -4, 1) = (0, 0, 0)$

(i.e.,) $(\alpha_1+2\alpha_2+7\alpha_3, -2\alpha_1+\alpha_2-4\alpha_3, \alpha_1-\alpha_2+\alpha_3) = (0, 0, 0)$

$\therefore$ $\alpha_1+2\alpha_2+7\alpha_3 = 0$ ---------(1)

$-2\alpha_1+\alpha_2-4\alpha_3 = 0$ ---------(2)

$\alpha_1-\alpha_2+\alpha_3 = 0$ ---------(3)

From (1) and (2), by the rule of cross-multiplication, $\frac{\alpha_1}{-15}=\frac{\alpha_2}{-10}=\frac{\alpha_3}{5}=K$ (say).

$\therefore$ $\alpha_1 = -15K$, $\alpha_2 = -10K$, $\alpha_3 = 5K$.

These values of α_1, α_2, α_3 clearly satisfy (3) for any value of K. Take K=1.

Then $\alpha_1 = -15$, $\alpha_2 = -10$, $\alpha_3 = 5$

Thus there exist scalars α_1, α_2, α_3 not all zero such that $\alpha_1v_1+\alpha_2v_2+\alpha_3v_3 = 0$

$\therefore$ $\{v_1, v_2, v_3\}$ is linearly dependent.

6. Let V be a vector space over a field F. Then any subset S of V containing the zero vector is linearly dependent.

Proof :

Let $S = \{0, v_1, \ldots, v_n\}$

Clearly $\alpha 0 + 0v_1 + 0v_2 + \ldots + 0v_n = 0$ where α is any element of F. Hence for any $\alpha \neq 0$, we get a non-trivial linear combination of vectors in S giving the zero vector. Hence S is linearly dependent.

Exercises :

1. Determine whether the following sets of vectors are linearly independent or linearly dependent in $V_3(R)$

   a) $\{(1,0,0), (0,1,0), (1,1,0)\}$

   b) $\{(1,2,3), (2,3,1)\}$

   c) $\{(1,2,3), (4,1,5), (-4,6,2)\}$

   d) $\{(0,0,0), (2,5,3), (-1, 0, 6)\}$

   e) $\{(1,0,0), (1,1,0), (1,1,1), (0,1,0)\}$

2. Determine whether the following sets of vectors are linearly independent or not.

   a) $\{(1,1,0,0), (0,0,1,1), (1,0,0,4), (0,0,0,2)\}$ in $V_4(R)$

   b) $\{(2i,1,0), (2,-i, 1), (0,1,i,-i)\}$ in $V_3(C)$

   c) $\{(\pi,0,0), (0,e,0), (0,0,\sqrt{5})\}$ in $V_3(R)$

   d) V = the set of all polynomials of degree $\leq n$ in R[x] and $S=\{1,x,x^2,\ldots.x^n\}$

   e) $\left\{\begin{bmatrix}2 & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}0 & 0\\ 2 & 0\end{bmatrix}\begin{bmatrix}0 & -1\\ 3 & 0\end{bmatrix}\right\}$ in $M_2(R)$

3. In $V_3(z_5)$ determine whether the following sets of vectors are linearly dependent.

   a) $\{(1,3,2), (2,1,3)\}$

   b) $\{(1,1,2), (2,1,0), (0,4,1)\}$

4. In $V_2(R)$ prove that the vectors (a,b) and (c,d) are linearly dependent iff ad–bc=0.

5. Let $\{v_1,v_2,v_3\}$ be a linearly independent set of vectors in $V_3(R)$.

Show that

(a) $\{v_1+v_2, v_2+v_3, v_3+v_1\}$ is linearly independent.

(b) $\{2v_1+v_2, v_1+v_2, v_1-v_3\}$ is linearly independent.

6. If the vectors (0,1,a), (1,a,1) and (a,1,0) of $V_3(R)$ are linearly dependent then find the value of a.

Answers :

1. (b) is linearly independent

2. (a), (b), (c), (d) and (e) are linearly independent.

3. (a) is linearly independent.

4. $a = 0, \pm\sqrt{2}$

Theorem 1.5 :

Any subset of a linearly independent set is linearly independent.

Proof :

Let V be a vector space over a field F.

Let $S = \{v_1,v_2,.....,v_n\}$ be a linearly independent set.

Let S^1 be a subset of S. Without loss of generality we take $S^1 = \{v_1,v_2,....,v_k\}$ where $k \leq n$.

Suppose S^1 is a linearly dependent set. Then there exist $\alpha_1,\alpha_2,.....,\alpha_k$ in F not all zero, such that $\alpha_1v_1+\alpha_2v_2+...+\alpha_kv_k = 0$.

Hence $\alpha_1v_1+\alpha_2v_2+.....+\alpha_kv_k+0v_{k+1}+.....+0v_n = 0$ is a non-trivial linear combination giving the zero vector.

Here S is a linearly dependent set which is a contradiction.

Hence S^1 is linearly independent.

Theroem 1.6 :

Any set containing a linearly dependent set is also linearly dependent.

Proof :

Let V be a vector space. Let S be a linearly dependent set. Let $S' \supset S$.

If S' is linearly independent S is also linearly independent (by previous theorem) which is a contradiction.

Hence S' is linearly dependent.

Theorem 1.7 :

Let $S = \{v_1, v_2, \ldots\ldots, v_n\}$ be a linearly independent set of vectors in a vector space V over a field F. Then every element of L(S) can be uniquely written in the form $\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n$, where $\alpha_i \in F$.

Proof :

By definition every elements of L(S) is of the form $\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n$.

Now, let $\alpha_1 v_1 + \alpha_2 v_2 + \ldots\ldots + \alpha_n v_n = \beta_1 v_1 + \beta_2 v_2 + \ldots\ldots + \beta_n v_n$

Hence $(\alpha_1 - \beta_1)v_1 + (\alpha_2 - \beta_2)v_2 + \ldots\ldots + (\alpha_n - \beta_n)v_n = 0$

Since S is a linearly independent set, $\alpha_i - \beta_i = 0$ for all i.

$\therefore \alpha_i = \beta_i$ for all i.

Hence the theorem.

Theorem 1.8 :

$S = \{v_1, v_2, \ldots\ldots, v_n\}$ is a linearly dependent set of vectors in V iff there exists a vector $v_k \in S$ such that v_k is a linear combination of the preceding vectors $v_1, v_2, \ldots, v_{k-1}$.

Proof :

Suppose $v_1, v_2, \ldots\ldots, v_n$ are linearly dependent. Then there exists $\alpha_1, \alpha_2, \ldots, \alpha_n \in F$, not all zero, such that $\alpha_1 v_1 + \alpha_2 v_2 + \ldots\ldots + \alpha_n v_n = 0$.

Let k be the largest integer for which $\alpha_k \neq 0$.

Then $\alpha_1 v_1 + \ldots\ldots + \alpha_k v_k = 0$

$\therefore\ \alpha_k v_k = -\alpha_1 v_1 - \alpha_2 v_2 - \ldots\ldots\ldots - \alpha_{k-1} v_{k-1}$

$\therefore\ v_k = (-\alpha_k^{-1}\alpha_1)v_1 + \ldots\ldots + (-\alpha_k^{-1}\alpha_{k-1})v_{k-1}$

$\therefore$ v_k is a linear combination ofthe prceding vectors.

Conversely, suppose there exists a vector v_k such that $v_k = \alpha_1 v_1 + \ldots\ldots + \alpha_{k-1} v_{k-1}$

Hence $-\alpha_1 v_1 - \ldots\ldots - \alpha_{k-1} v_{k-1} + v_k + 0v_{k+1} + \ldots + 0v_n = 0$

Since the coefficient of $v_k = 1$, we have $S = \{v_1, \ldots, v_n\}$ is linearly dependent.

Examples :

In $V_3(R)$, let $S = \{(1,0,0), (0,1,0), (0,0,1), (1,1,1)\}$

Here $(1,1,1) = (1,0,0) + (0,1,0) + (0,0,1)$

Thus (1,1,1) is a linear combination of the preceding vectors. Hence S is a linearly dependent set.

Theorem 1.9 :

Let V be a vector space over F. Let $S = \{v_1, v_2, \ldots v_n\}$ and $L(S) = W$. Then there exists a linearly independent subset S^1 of S such that $L(S') = W$.

Proof :

Let $S = \{v_1, v_2, \ldots, v_n\}$

If S is linearly independent there is nothing to prove.

If not, let v_K be the first vector in S wich is a linear combination of the preceding vectors.

Let $S_1 = \{v_1, v_2, \ldots\ldots, v_{k-1}, v_{k+1}, \ldots\ldots, v_n\}$

(i.e.,) S_1 is obtained by deleting the vector v_k from S.

We claim that $L(S_1) = L(S) = W$.

Since $S_1 \subseteq S$, $L(S_1) \subseteq L(S)$ (refer theorem 1.4)

Now, let $v \in L(S)$

Then $v = \alpha_1 v_1 + ... + \alpha_k v_k + ... + \alpha_n v_n$

Now, v_k is a linear combination of the preceding vectors.

Let $v_k = \beta_1 v_1 + ... + \beta_{k-1} v_{k-1}$

Hence $v = \alpha_1 v_1 + ... \alpha_{k-1} v_{k-1}$

$+\alpha_k(\beta_1 v_1 + ... + \beta_{k-1} v_{k-1}) + \alpha_{k+1} v_{k+1} + ... + \alpha_n v_n$

$\therefore$ v can be expressed as a linear combination of the vectors of S_1 so that $v \in L(S_1)$.

Hence $L(S) \subseteq L(S_1)$

Thus $L(S) = L(S_1) = W$

Now, if S_1 is linearly independent, the proof is complete.

If not, we continue the above process of removing a vector from S_1, which is a linear combination of the preceeding vectors until we arrive at a linearly independent subset S' of S such that $L(S') = W$.

Solved Problems :

Problem 1 :

If the vectors $v_1, ..., v_n$ and the scalars $\alpha_2, ... \alpha_n$ are such that $A = \{v_2 + \alpha_2 v_1, v_3 + \alpha_3 v_1, ... v_n + \alpha_n v_1\}$ is linearly dependent, show that $B = \{v_1, ..., v_n\}$ is also linearly dependent.

Since A is linearly dependent, there exists scalars $\beta_2, ... \beta_n$ not all zero, such that

$\beta_2(v_2 + \alpha_2 v_1) + \beta_3(v_3 + \alpha_3 v_1) + ... + \beta_n(v_n + \alpha_n v_1) = 0$ -------(1)

Say $\beta_2 \neq 0$.

From (1), $(\beta_2\alpha_2 + \beta_3\alpha_3 + ... + \beta_n\alpha_n)v_1 + \beta_2 v_2 + \beta_3 v_3 + ... + \beta_n v_n = 0$

Since the coefficient of v_2 is $\beta_2 \neq 0$, it follows that $\{v_1, ... v_n\}$ is linearly dependent.

Problem 2 :

If the set {u, v, w} is linearly independent in V(C). Show that {u+v, u–v, u–2v+w} is also linearly independent.

Let $\alpha(u+v)+\beta(u-v)-\gamma(u-2v+w) = 0$

Then $(\alpha+\beta+\gamma)u+(\alpha-\beta-2\gamma)v+\gamma w = 0$

But {u, v, w} is linearly independent.

$\therefore$ $\alpha+\beta+\gamma = 0$

$\alpha-\beta-2\gamma = 0$

$\gamma = 0$

These equations have the only solution $\alpha=0$, $\beta=0$, $\gamma=0$.

$\therefore$ u+v, u–v, u–2v+w are linearly independent.

3. Find the linearly independent subset A of $S = \{v_1,v_2,v_3,v_4\}$ in $R^3(R)$ such that $L(A)=L(S)$ where $v_1=(1,2,-1)$, $v_2=(-3, -6, 3)$, $v_3=(2,1,3)$ and $v_4=(8,7,7)$.

$S = \{v_1,v_2,v_3,v_4\}$

Now $v_2 = (-3, -6, 3) = -3(1, 2, -1) = -3v_1$ = a linear combination of v_1.

$\therefore$ remove v_2 from v_1.

Let $T = \{v_1, v_3, v_4\}$. Clearly v_3 is not a linear combination of v_1. Is v_4 a linear combination of v_1, v_3?

Suppose $v_4 = av_1+bv_3$

Then $(8,7,7) = a(1, 2, -1)+b(2,1,3)$

$= (a+2b, 2a+b, -a+3b)$

$\therefore$ $a+2b = 8$

$2a+b = 7$

$-a+3b = 7$

Solving, we get $a = 2$, $b = 3$.

$\therefore$ $v_4 = 2v_1+3v_3$ = a linear combination of v_1 & v_3.

$\therefore$ remove v_4 from T. Thus we get $A = \{v_1, v_3\}$

Since v_3 is not a linear combination of v_1, A is linearly independent. Since $A \subset S$, $L(A) \subseteq L(S)$.

Let $v \in L(S)$. Then

$$\begin{aligned} v &= av_1+bv_2+cv_3+dv_4 \\ &= av_1-3bv_1+cv_3+d(2v_1+3v_3) \\ &= (a-3b+2d)v_1+(c+3d)v_3 \in L(A) \end{aligned}$$

$\therefore L(S) \subseteq L(A)$

Hence $L(A) = L(S)$.

Thus A is the required subset of S.

4. Let V be the vector space of functions form R into R. Show that {f, g, h} is a linearly independent subset in V, where $f(x)=e^{2x}$, $g(x)=x^2$ and $h(x) = x$.

Let a, b, c be scalars and let $af+bg+ch = 0$, where the function 0 is defined by $0(x) = 0\ \forall x \in R$

$\therefore (af+bg+ch)(x) = 0 = 0(x)\ \forall x \in R$

$\therefore af(x)+bg(x)+ch(x) = 0$

(i.e.,) $ae^{2x}+bx^2+cx = 0\ \forall x \in R$

Put $x = 0$ $\therefore a = 0$ ---------(1)

Put $x = 1$ $\therefore ae^2+b+c = 0$ ---------(2)

Put $x = 2$ $\therefore ae^4+4b+2c = 0$ ---------(3)

Solving (1), (2) & (3) we get

$a = 0, b = 0, c = 0$

$\therefore af+bg+ch = 0 \Rightarrow a = 0, b = 0, c = 0$

$\therefore$ {f, g, h} is linearly independent.

5. Show that the vectors (a, b) and (c, d) in c^2 are linearly dependent $\Leftrightarrow ad = bc$.

Let $x, y, \in c$

Then $x(a, b)+y(c, d) = 0$

$(xa+yc, xb+yd) = (0, 0)$

$\therefore$ $ax+cy = 0$ and

$bx+dy = 0$

We know that these equations have a solution other than

$$x=y=0 \Leftrightarrow \Delta = \begin{vmatrix} a & c \\ b & d \end{vmatrix} = 0$$

$$\Leftrightarrow ad-bc = 0$$

$\therefore$ The vectors (a, b) and (c, d) are linearly dependent $\Leftrightarrow$ $ad = bc$.

6. Let $V = F[x]$. Show that the infinite set $S = \{1,x,x^2,.....\}$ is linearly independent.

Let $A = \{x^{\alpha_1}, x^{\alpha_2},.....x^{\alpha_n}\}$ be any finite subset of S, where $\alpha_1, \alpha_2,.....,\alpha_n$ are non-negative integers.

Let $a_1, a_2,.....a_n$ be scalars such that $a_1x^{\alpha_1}+a_2x^{\alpha_2}+.....+a_nx^{\alpha_n} = 0$ (zero polynomial).

Then by the definition of equality of two polynomials, we get $a_1=0$, $a_2=0,....,a_n=0$.

$\therefore$ A is linearly independent and hence S is also linearly independent.

1.5. BASIS AND DIMENSION

Definition :

A Linearly Independent subset S of a vector space V which spans the whole space V is called a **basis** of the vector space.

Theorem 1.10 :

Any finite - dimensional vector space V contains a finite number of Linearly Independent vectors which span V. (i.e.,) A finite dimensional vector space has a basis consisting of a finite number of vectors.

Proof :

Since V is finite dimensional there exists a finite subset S of V such that $L(S)=V$.

By the theorem 1.9 this set S contains a linearly Independent subset $S^1 = \{v_1, v_2, \ldots v_n\}$ such that $L(S') = L(S) = V$

Hence S^1 is a basis for V.

Theorem 1.11 :

Let V be a vector space over a field F. Then $S = \{v_1, v_2, \ldots, v_n\}$ in a basis for V iff every element of V can be uniquely expressed as a linear combination of elements of S.

Proof :

Let S be a basis for V. Then by definition S is Linearly Independent and L(S)=V. Hence by theorem 1.7 every element of V can be uniquely expressed as a linear combination of elements of S.

Conversely :

Suppose every element of V can be uniquely expressed as a linear combination of elements of S.

Clearly $L(S) = V$

Now, Let $\alpha_1 v_1 + \alpha_2 v_2 + \ldots\ldots + \alpha_n v_n = 0$

Also $Ov_1 + Ov_2 + \ldots\ldots + Ov_n = 0$

Thus we have expressed O as a linear combination of vectors of S in two ways.

∴ By hypothesis $\alpha_1 = \alpha_2 = \ldots\ldots = \alpha_n = 0$.

Hence S is Linearly Independent

Hence S is a basis.

Examples :

1) $S = \{(1,0,0), (0,1,0), (0,0,1)\}$ is a basis for $V_3(R)$ for, $(a,b,c) = a(1,0,0)+b(0,1,0)+c(0,0,1)$.

∴ Any vector (a, b, c) of $V_3(R)$ has been expressed uniquely as a linear combination of the elements of S and hence S is a basis for $V_3(R)$.

2) $S = \{e_1, e_2, \ldots\ldots e_n\}$ is a basis for $V_n(F)$. This is known as the **standard basis** for $V_n(F)$.

3) $S = \{(1,0,0), (0,1,0), (1,1,1)\}$ is a basis for $V_3(R)$.

Proof :

We shall show that any element (a, b, c) of $V_3(R)$ can be uniquely expressed as a linear combination of the vectors of S.

Let $(a, b, c) = \alpha(1, 0, 0)+\beta(0,1,0)+\gamma(1,1,1)$

Then $\alpha+\gamma = a$, $\beta+\gamma = b$, $\gamma = c$

Hence $\alpha = a-c$ and $\beta = b-c$

Thus $(a, b, c) = (a-c)(1,0,0) + (b-c)(0,1,0) + c(1,1,1)$

$\therefore$ S is a basis for $V_3(R)$.

4) $S = \{1\}$ is a basis for the vector space R over R.

5) $S = \left\{\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}, \begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}, \begin{pmatrix}0 & 0\\ 1 & 0\end{pmatrix}, \begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}\right\}$ is a basis for $M_2(R)$, since any matrix $\begin{pmatrix}a & b\\ c & d\end{pmatrix}$ can be uniquely written as

$$\begin{pmatrix}a & b\\ c & d\end{pmatrix} = a\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}+b\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}+c\begin{pmatrix}0 & 0\\ 1 & 0\end{pmatrix}+d\begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}$$

6) $\{1, i\}$ is a basis for the vector space C over R.

7) Let V be the set of all polynomials of degree $\leq n$ in R[x]. Then $\{1,x,x^2,\ldots\ldots,x^n\}$ is a basis for V.

8) $\{(1,0), (i,0), (0,1), (0,i)\}$ is a basis, for the vector space $C\times C$ over R, for

$(a+ib, c+id) = a(1,0) + b(i, 0)+c(0,1)+d(0,i)$

9) $S = \{(1,0,0), (0,1,0), (1,1,1), (1,1,0)\}$ spans the vector space $V_3(R)$ but is not a basis.

Proof :

Let S' = {(1,0,0), (0,1,0), (1,1,1)}

Then L(S') = $V_3(R)$ (refer examples 3)

Now since $S \subseteq S'$. We get L(S) = $V_3(R)$

Thus S spans $V_3(R)$

But S is linearly dependent since

$$(1,1,0) = (1,0,0) + (0,1,0)$$

Hence S is not a basis.

10) S = {(1,0,0), (1,1,0)} is Linearly Independent but not a basis of $V_3(R)$.

Proof :

Let $\alpha(1,0,0)+\beta(1,1,0) = (0,0,0)$

Then $\alpha+\beta = 0$ and $\beta = 0$

$\therefore \alpha = \beta = 0$. Hence S is linearly independent.

Also L(S) = $\{(a,b,0) / a,b \in R\} \neq V_3(R)$

$\therefore$ S is not a basis.

Solved Problems :

1. Show that B = {(0,1,0), (1,0,1), (1,1,0)} is a basis for $V_3(R)$ i.e., R^3.

Solution :

For $\alpha(0,1,0)+\beta(1,0,1)+\gamma(1,1,0) = 0$

$\Rightarrow (\beta+\gamma, \alpha+\gamma, \beta) = (0,0,0)$

$\Rightarrow \beta+\gamma = 0, \alpha+\gamma = 0, \beta = 0$

$\Rightarrow \alpha = \beta = \gamma = 0$

$\therefore$ B is linearly independent.

To prove L(B) = R^3, we have to show that every vector in R^3 can be expressed as a linear combination of elements of B.

Let (a, b, c) be any element of R^3. We want to write (a, b, c) = x(0,1,0) + y(1,0,1) + z(1,1,0) where x,y,z are scalars to found.

Hence we require (a, b, c) = (y+z, x+z, y)

$\therefore$ y+z = a, x+z = b, y = c

Clearly, these equations have the solution x = b+c–a, y = c, z = a–c.

Thus, (a,b,c) = (b+c–a)(0,1,0) + c(1,0,1) + (a–c) (1,1,0)

$\therefore$ L(B) = R^3 and hence B is a basis for R^3.

Note that {(1,0,0), (0,1,0), (0,0,1)} is the standard basis for R^3. Thus, we find that a vector space may have more than one basis.

2. Show that the set B = {(1,2,1), (2,1,0), (1,–1,2)} is a basis for $V_3(R) = R^3$.

Solution :

Now $\alpha_1(1,2,1)+\alpha_2(2,1,0)+\alpha_3(1,-1,2) = 0$

$\Rightarrow (\alpha_1+2\alpha_2+\alpha_3, 2\alpha_1+\alpha_2-\alpha_3, \alpha_1+2\alpha_3) = (0,0,0)$

$\Rightarrow$ $\alpha_1+2\alpha_2+\alpha_3 = 0$

$2\alpha_1+\alpha_2-\alpha_3 = 0$

$\alpha_1+2\alpha_3 = 0$

$\Rightarrow \alpha_1 = \alpha_2 = \alpha_3 = 0$

$\therefore$ B is linearly independent.

Let (a,b,c) $\in R^3$ be arbitrary.

Then (a,b,c) = a(1,0,0)+b(0,1,0)+c(0,0,1) -------(1)

We write (1,0,0) as a linear combination of elements of B as follows.

Let (1,0,0) = x(1,2,1)+y(2,1,0)+z(1,–1,2)

= (x+2y+z, 2x+y–z, x+2z)

$\therefore$ x+2y+z = 1

2x+y–z = 0

x+2z = 0

On solving these equations, we get $x = -\frac{2}{9}$, $y = \frac{5}{9}$, $z = \frac{1}{9}$.

$\therefore$ $(1,0,0) = -\frac{2}{9}(1,2,1)+\frac{5}{9}(2,1,0)+\frac{1}{9}(1,-1,2)$

Similarly, $(0,1,0) = \frac{4}{9}(1,2,1)-\frac{1}{9}(2,1,0)-\frac{2}{9}(1,-1,2)$

and $(0,0,1) = \frac{1}{3}(1,2,1)-\frac{1}{3}(2,1,0)+\frac{1}{3}(1,-1,2)$

Using these in (1), we find that (a,b,c) is a linear combination of (1,2,1), (2,1,0) and (1,–1, 2)

$\therefore$ L(B) = R^3 and hence B is a basis for R^3.

3. Show that S = {(1,0,0), (0,1,0)} is linearly independent but is not a basis for R^3.

$$\alpha(1,0,0)+\beta(0,1,0) = (0,0,0)$$
$$(\alpha,\beta,0) = (0,0,0)$$
$$\Rightarrow \quad \alpha = \beta = 0$$

$\therefore$ S is linearly independent.

$$L(S) = \{a(1,0,0)+b(0,1,0)/a,b\in R\}$$
$$= \{(a,b,0)/a,b\in R\}$$
$$\neq R^3$$

$\therefore$ S is not a basis for R^3.

Exercise :

1) Show that the following three vectors from a basis for $V_3(R)$.

   a) (1,2,–3), (2,5,1), (–1,1,4)

   b) (1,1,0), (0,1,1), (1,0,1)

   c) (2,–3,1), (0,1,2), (1,1,2)

2) Show that the following sets of vectors do not from a basis for $V_3(R)$.

   a) {(1,0,0), (1,1,0)}

   b) {(1,2,1), (1,3,5), (–1,0,1), (1,–1,2)}

c) $\{(0,0,0), (1,0,0), (0,1,0), (0,0,1)\}$

d) $\{(3,2,1), (3,1,5), (3,4,-7)\}$

e) $\{(1,2,3), (2,3,4), (3,4,5)\}$

3) Show that $(1,i,0), (2i,1,1), (0, 1+i, 1-i)$ from a basis for $V_3(c)$.

4) Find a basis for the vector space consisting of all matrices of the form.

a) $\begin{bmatrix} a & b \\ 0 & d \end{bmatrix}$ b) $\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$

5) If $\{v_1,v_2,v_3\}$ is a basisfor $V_3(R)$, show that $\{v_1+v_2, v_2+v_3, v_3+v_1\}$ is also a basis. Is this true in (a) $V_3(z_2)$ (b) $V_3(z_3)$?

Answers :

4) a) $\left\{\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right\}$ b) $\left\{\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right\}$

Theorem 1.12 :

Let V be a vector space over a field F. Let $S = \{v_1,v_2,.....,v_n\}$ span V. Let $S = \{w_1,w_2,....w_m\}$ be a linearly independent set of vectors in V. Then $m \leq n$.

Proof :

Since $L(S) = V$, every vector in V and in particular w_1, is a linear combination of $v_1,v_2,\ldots,v_n$.

Hence $S_1 = \{w_1,v_1,v_2,....,v_n\}$ is a linearly dependent set of vectors. Hence there exists a vector $v_k \neq w_1$ in S_1. Which is a linear combination of the preceding vectors.

Let $S_2 = \{w_1,v_1,.....,v_{k-1},v_{k+1},.....v_n\}$

Clearly $L(S_2) = V$

Hence w_2 is a linear combination of the vectors in S_2.

Hence $S_3 = \{w_2,w_1,v_1,.....,v_{k-1},v_{k+1},.....v_n\}$ is linearly dependent. Hence there exists a vector in S_3 which is a linear combination of the preceding vectors.

Since the w_i's are linearly independent, this vector cannotbe w_2 or w_1 and hence must be some v_j. where $j \neq k$ (say, with $j>k$). Deletion of v_j from the set S_3 gives the set

$S_4 = \{w_2,w_1,v_1,v_2,.....,v_{k-1}, v_{k+1},.....,v_{j-1},v_{j+1},.....v_n\}$ of n vectors spanning V.

In this process, at each step we insert one vector from $\{w_1,w_2,.....w_m\}$ and delete one vector from $\{v_1,v_2,.....,v_n\}$.

If m>n after repeating this process n times, we arrive at the set $\{w_n,w_{n-1},.....w_1\}$ which spans V.

Hence w_{n+1} is a linear combination of $w_1,w_2.....,w_n$.

Hence $\{w_1,w_2,.....w_n,w_{n+1},.....w_m\}$ is linearly dependent.

Which in a contradiction.

Hence m≤n.

Theorem 1.13 :

Any two bases of a finite dimensional vector space V have the same number of elements.

Proof :

Since V is finite dimenional. It has a basis say $S = \{v_1,v_2,....v_n\}$

Let $S' = \{w_1,w_2,.....,w_m\}$ be any other basis for V.

Now L(S) = V and S^1 is a set of m linearly independent vectors. Hence by theorem 1.12, m≤n.

Also since $L(S^1) = V$ and S is a set of n linearly independent vectors, n≤m.

Hence m=n.

Definition :

Let V be a finite dimensional vector space over a field F. The number of elements in any basis of V is called the **dimension** of V and is denoted by dim V.

Examples :

1. dim $V_n(R) = n$. since $\{e_1,e_2,.....e_n\}$ is a basis of $V_n(R)$.

2. $M_2(R)$ is a vector space of dimension 4 over R.

since $\left\{\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix},\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix},\begin{pmatrix}0 & 0\\ 1 & 0\end{pmatrix},\begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}\right\}$ is a basis for $M_2(R)$.

3. C is a vector space of dimension 2 over R. Since {1, i} is a basis for C.

4. Let V be the set of all polynomials of degree $\leq$ n in R[x].V is a vector space over R having dimension n+1, since $\{1,x,x^2.....x^n\}$ is a basis for V.

Theorem 1.14 :

Let V be a vector space of dimension n Then,

(i) any set of m vectors where m>n is linearly dependent

(ii) any set of m vectors where m<n cannot span V.

Proof :

(i) Let $S = \{v_1,v_2,....,v_n\}$ be a basis for V.

Hence L(S) = V.

Let S' be any set consisting of m vectors where m > n. Suppose S' is linearly Independent. Since S spans V by theorem 1.12, m$\leq$n.

Which is a contradiction

Hence S^1 is linearly dependent.

(ii) Let S^1 be a set consisting of m vectors where m<n. Suppose L(S') = V.

Now $S = \{v_1,v_2,....,v_n\}$ is a basis for V and hence linearly Independent. Hence by theorem 1.12 n$\leq$m. Which is a contradiction.

Hence S' cannot span V.

Theorem 1.15 :

Let V be a finitie dimensional vector space over a field F. Any linearly Independent set of vectros in V is part of a basis.

Proof :

Let $S = \{v_1,v_2,....,v_r\}$ be a linearly Independent set of vectors.

If L(S) = V then S itself is a basis

If L(S) $\neq$ V, choose an element $v_{r+1} \in V-L(S)$.

Now consider, $S_1 = \{v_1,v_2,....,v_r,v_{r+1}\}$

We shall prove that S_1 is linearly Independent by showing that no vector in S_1 is a linear combination of the preceeding vectors. (refer theorem 1.8)

Since $\{v_1,v_2,.....,v_r\}$ is linearly Independent, v_i where $1\leq i\leq r$ is not a linear combination of the preceeding vectors.

Also $v_{r+1}\notin L(S)$ and Hence v_{r+1} is not a linear combination of $v_1,v_2.....v_r$.

Hence S_1 is linearly Independent.

If $L(S_1) = V$, then S_1 is a basis for V. If not we take an element $v_{r+2}\in V-L(S_1)$ and proceed as before. Since the dimension of V is finite, this process must stop at a certain stage giving the required basis containing S.

Theorem 1.16 :

Let V be a finite dimensional vector space over a field F. Let A be a subspace of V. Then there exists a subspace B of V such that $V=A\oplus B$.

Proof :

Let $S = \{v_1,v_2,.....v_r\}$ be a basis of A. By theorem 1.15. We can find $w_1,w_2,.....,w_s\in V$ such that $S' = \{v_1,v_2,.....v_r, w_1, w_2.....w_s\}$ is a basis of V.

Now let $B = L(\{w_1,w_2,.....,w_s\})$

We claim that $A\cap B=\{0\}$ and $V = A+B$

Now, let $v\in A\cap B$. Then $v\in A$ and $v\in B$

Hence $$v = \alpha_1v_1+\alpha_2v_2+.....+\alpha_rv_r$$
$$= \beta_1w_1+\beta_2w_2+.....+\beta_sw_s$$

$\therefore \alpha_1v_1+\alpha_2v_2+.....+\alpha_rv_r-\beta_1w_1-\beta_2w_2-.....-\beta_sw_s = 0$

Now since S' is linearly Independent $\alpha_i = 0 = \beta_j$ for all i and j.

Hence $v = 0$. Thus $A\cap B = \{0\}$

Now, let $v\in V$

Then $$v = (\alpha_1v_1+\alpha_2v_2+.....+\alpha_rv_r)$$
$$+ (\beta_1w_1+\beta_2w_2+....+\beta_sw_s)\in A+B$$

Hence $A+B = V$ so that $V = A\oplus B$.

Exercise :

1. Let V be a finite dimensional vector space. Let A and B be subspaces of V such that $V=A\oplus B$. Then show that dim V = dim A + dim B.

2. Construct 3 subspaces W_1, W_2, W_3 of a vector space V such that $V = W_1\oplus W_2 = W_1\oplus W_3$ but $W_2\neq W_3$.

3. For each of the following subspaces A of $V_3(R)$ find another subspace B such that $A\oplus B=V_3(R)$

   (i) $A = L\{(1,1,0), (0,1,1)\}$

   (ii) $A = L\{(1,1,1)\}$

   (iii) $A = L(\{e_1,e_2,e_3\})$

Definition :

Let V be a vector space and $S = \{v_1,v_2,.....,v_n\}$ be a set of independent vectors in V. The S is called a **maximal linearly Independent set** if for every $v\in V-S$, the set $\{v,v_1,v_2,.....,v_n\}$ is linearly dependent.

Definition :

Let $S = \{v_1,v_2,.....,v_n\}$ be a set of vectors in V and Let L(S) = V. Then S is called a **minimal generating set** if for any $v_i\in S$, $L(S-\{v_i\})\neq V$.

Theorem 1.17 :

Let V be a vector space over field F Let $S = \{v_1,v_2,.....v_n\}\subseteq V$. Then the following are equivalent.

(i) S is a basis for V

(ii) S is a Maximal linearly Independent set

(iii) S is a minimal generating set

Proof :

(i) $\Rightarrow$ (ii)

Let $S = \{v_1,v_2,.....v_n\}$ be basis for V. Then by theorem 1.14 any n+1 vectors in V are linearly dependent and hence S is a Maximal linearly Independent set.

(ii) $\Rightarrow$ (i)

Let $S = \{v_1, v_2, \ldots\ldots v_n\}$ be a maximal linearly Independent set. Now to prove that S is a basis for V we shall show that $L(S) = V$.

Obviously $L(S) \subseteq V$

Now, Let $v \in V$

If $v \in S$, then $v \in L(S)$. (Since $S \subseteq L(S)$)

If $v \notin S$, $S' = \{v_1, v_2, \ldots\ldots, v_n, v\}$ is a linearly dependent set (since S is a Maximal linearly Independent set).

$\therefore$ There exists a vector in S'. Which is a linear combination of the preceeding vectors.

Since $v_1, v_2, \ldots\ldots v_n$ are linearly Independent, this vector must be v. Thus V is a linear combination of $v_1, v_2, \ldots. v_n$. Therefore $v \in L(S)$.

Hence $V \subseteq L(S)$

Thus $V = L(S)$.

(i) $\Rightarrow$ (iii)

Let $S = \{v_1, v_2, \ldots\ldots v_n\}$ be a basis. Then $L(S) = V$.

If S is not minimal, there exists $v_i \in S$ such that $L(S - \{v_i\}) = V$.

Since S is linearly Independent, $S - \{v_i\}$ is also linearly independent. Thus $S - \{v_i\}$ is a basis consisting of n–1 elements.

Which is a contradiction.

Hence S is a minimal generating set.

(iii) $\Rightarrow$ (i)

Let $S = \{v_1, v_2, \ldots\ldots v_n\}$ be a minimal generating set. To prove that S is a basis.

We have to show that S is linearly Independent.

If S is linearly dependent, there exists a vector v_k which is a linear combination of the preceeding vectors.

Clearly $L(S - \{v_k\}) = V$ contradicting the minimality of S.

Thus S is linearly Independent and since $L(S) = V$, S is a basis for V.

Theorem 1.18 :

Any vector space of dimension n over a field F is isomorphic to $V_n(F)$.

Proof :

Let V be a vector space of dimension n. Let $\{v_1,v_2,....v_n\}$ be a basis for V.

Then we known that if $v \in V$, v can be written uniquely as $v=\alpha_1v_1+\alpha_2v_2+....+\alpha_nv_n$ where $\alpha_i \in F$.

Now, consider the map $f:V \to V_n(F)$ given by $f(\alpha_1v_1+.....+\alpha_nv_n) = (\alpha_1,\alpha_2,.....,\alpha_n)$

Clearly f is 1–1 and onto.

Let $v,w, \in V$

Then $v = \alpha_1v_1+.....+\alpha_nv_n$ and

$w = \beta_1v_1+.....+\beta_nv_n$

$f(v+w) = [(\alpha_1+\beta_1)v_1+.....+(\alpha_n+\beta_n)v_n]$

$= [(\alpha_1+\beta_1),(\alpha_2+\beta_2).....(\alpha_n+\beta_n)]$

$= (\alpha_1,\alpha_2,.....\alpha_n) + (\beta_1,\beta_2,.....,\beta_n)$

$= f(v)+f(w)$

Also $f(\alpha v) = f(\alpha\alpha_1v_1+.....+\alpha\alpha_nv_n)$

$= (\alpha\alpha_1,\alpha\alpha_2,.....,\alpha\alpha_n)$

$= \alpha(\alpha_1,\alpha_2,.....,\alpha_n)$

$= \alpha f(v)$

Hence f is an isomorphism of V to $V_n(F)$.

Corollary :

Any two vector spaces of the same dimension over a field F are isomorphic, for, if the vector spaces are of dimension n, each is isomorphic to $V_n(F)$ and hence they are isomorphic.

Theorem 1.19 :

Let V and W be vector spaces over a field F. Let $T:V \to W$ be an isomorphism. Then T maps a basis of V onto a basis of W.

Proof :

Let $\{v_1,v_2,.....v_n\}$ be a basis for V. We shall prove that $T(v_1)$, $T(v_2).....T(v_n)$ are linearly Independent and that they span W.

Now $\alpha_1T(v_1)+\alpha_2T(v_2)+.....+\alpha_nT(v_n) = 0$

$\Rightarrow T(\alpha_1v_1)+T(\alpha_2v_2)+....+T(\alpha_nv_n) = 0$

$\Rightarrow T(\alpha_1v_1+\alpha_2v_2+.....+\alpha_nv_n) = 0$

$\Rightarrow \alpha_1v_1+\alpha_2v_2+.....+\alpha_nv_n = 0$ (since T is 1–1)

$\Rightarrow \alpha_1 = \alpha_2 = ... = \alpha_n = 0$

(Since $v_1,v_2.....v_n$ are linearly Independent).

$\therefore$ $T(v_1)$, $T(v_2).....T(v_n)$ are linearly Independent.

Now, let $w \in W$. Then since T is onto, there exists a vector $v \in V$ such that $T(v)=w$.

Let $v = \alpha_1v_1+.....+\alpha_nv_n$

Then $w = T(v)$

$= T(\alpha_1v_1+.....+\alpha_nv_n)$

$= \alpha_1T(v_1)+.....+\alpha_nT(v_n)$

Thus w is a linear combination of the vectors, $T(v_1),T(v_2),.....T(v_n)$.

$\therefore$ $T(v_1)$, $T(v_2),.....,T(v_n)$ span W and hence is a basis for W.

Corollary :

Two finite dimensional vector spaces V and W over a field F are isomorphic iff they have the same dimension.

Theorem 1.20 :

Let V and W be finit dimensional vector spaces over a field F. Let $\{v_1,v_2,.....v_n\}$ be a basis for V and let $w_1,w_2,.....w_n$ be any n vectors in W (not necessrily distinct). Then there exists a unique linear transformation $T:V\rightarrow W$ such that $T(v_i)=w_i$, i=1,2,....n.

Proof :

Let $v = \alpha_1 v_1+\alpha_2 v_2+.....+\alpha_n v_n \in V$

We define $T(v) = \alpha_1 w_1+\alpha_2 w_2+....+\alpha_n w_n$

Now, let $x, y \in V$

Let $x = \alpha_1 v_1+.....+\alpha_n v_n$

and $y = \beta_1 v_1+.....+\beta_n v_n$

$\therefore x+y = (\alpha_1+\beta_1)v_1+.....+(\alpha_n+\beta_n)v_n$

$\therefore T(x+y) = (\alpha_1+\beta_1)w_1+.....+(\alpha_n+\beta_n)w_n$

$= (\alpha_1 w_1+.....+\alpha_n w_n)+(\beta_1 w_1+.....+\beta_n w_n)$

$= T(x)+T(y)$

Similarly $T(\alpha x) = \alpha T(x)$

Hence T is a linear transformation.

Also $v_1 = 1v_1+0v_2+.....0v_n$

Hence $T(v_1) = 1w_1+0w_2+.....+0w_n = w_1$

Similarly $T(v_i) = w_i$ for all $i = 1,2,.....n.$

Now to prove the uniqueness, Let $T' : V \to W$ be any other linear transformation such that $T'(v_i) = w_i$

Let $v = \alpha_1 v_1+\alpha_2 v_2+.....+\alpha_n v_n \in V$

$T^1(v) = \alpha_1 T^1(v_1)+\alpha_2 T^1(v_2)+.....+\alpha_n T^1(v_n)$

$= \alpha_1 w_1+\alpha_2 w_2+.....+\alpha_n w_n$

$= T(v)$

Hence $T = T^1$

Remark :

The above thorem shows that a linear transformation is completely determined by its values on the elements of a basis.

Theorem 1.21 :

Let V be a finite dimensional vector space over a field F. Let W be a subspace of V.

Then (i) $\dim W \leq \dim V$

(ii) $\dim \frac{V}{W} = \dim V - \dim W$

Proof :

(i) Let $S = \{w_1, w_2, \ldots . w_m\}$ be a basis for W. Since W is a subspace of V, S is a part of a basis for V.

Hence $\dim W \leq \dim V$.

(ii) Let $\dim V = n$ and $\dim W = m$

Let $S = \{w_1, w_2, \ldots . w_m\}$ be a basis for W. Clearly S is a linearly Independent set of vectors in V.

Hence S is a part of a basis in V. Let $\{w_1, w_2, \ldots . w_m, v_1, v_2, \ldots . v_r\}$ be a basis for V. Then $m+r = n$.

Now, we claim $S' = \{W+v_1, W+v_2, +\ldots . . W+v_r\}$ is a basis for $\frac{V}{W}$.

$\alpha_1(W+v_1)+\alpha_2(W+v_2)+\ldots . . +\alpha_r(W+v_r) = W+0$

$\Rightarrow (W+\alpha_1 v_1)+(W+\alpha_2 v_2)+\ldots .+(W+\alpha_r v_r) = W$

$\Rightarrow W+\alpha_1 v_1+\alpha_2 v_2+\ldots . .+\alpha_r v_r = W$

$\Rightarrow \alpha_1 v_1+\alpha_2 v_2+\ldots . . +\alpha_r v_r \in W$

Now since $\{w, w_2, \ldots . . . w_n\}$ is a basis for W.

$\alpha_1 v_1+\alpha_2 v_2+\ldots . . \alpha_r v_r = \beta_1 w_1+\beta_2 w_2+\ldots . . +\beta_m w_m$

$\therefore \alpha_1 v_1+\alpha_2 v_2+\ldots . . \alpha_r v_r-\beta_1 w_1-\beta_2 w_2-\ldots . . -\beta_m w_m = 0$

$\therefore \alpha_1 = \alpha_2 = \ldots . . . = \alpha_r = \beta_1 = \beta_2 \ldots . . = \beta_m = 0$

$\therefore$ S′ is a linearly Independent set.

Now let $W+v \in \frac{V}{W}$

Let $v = \alpha_1 v_1+\ldots . . +\alpha_r v_r+\beta_1 w_1+\ldots . .+\beta_m w_m$

Then $W+v = W+(\alpha_1 v_1+\ldots . . +\alpha_r v_r+\beta_1 w_1+\ldots . .+\beta_m w_m)$

$= W+(\alpha_1 v_1+\ldots . . +\alpha_r v_r)$

(Since $\beta_1 w_1+\ldots . . +\beta_m w_m \in W$)

$$= (W+\alpha_1 v_1)+.....+(W+\alpha_r v_r)$$

$$= \alpha_1(W+v_1)+.....+\alpha_r(W+v_r)$$

Hence S′ spans $\frac{V}{W}$ so that S′ is a basis for $\frac{V}{W}$.

$\therefore \dim \frac{V}{W} = r = n-m = \dim V - \dim W.$

Theorem : 1.22 :

Let V be a finite dimensional vector space over a field F. Let A and B subspaces of V. Then $\dim(A+B) = \dim A + \dim B - \dim(A \cap B)$.

Proof :

A and B are subspaces of V. Hence $A \cap B$ is subspace of V.

Let $\dim(A \cap B) = r$

Let $S = \{v_1, v_2, ...v_r\}$ be a basis for $A \cap B$.

Since $A \cap B$ is a subspace of A and B, S is a part of a basis for A and B.

Let $\{v_1, v_2, ...v_r, u_1, u_2, ...u_s\}$ be a basis for A.

and $\{v_1, v_2 .....v_r, w_1, w_2 .....w_t\}$ be a basis for B.

We shall prove that $S' = \{v_1 v_2 .....v_r, u_1, u_2, ...u_s, w_1 w_2 .....w_t\}$ is a basis for A+B.

Let $\alpha_1 v_1+.....+\alpha_r v_r+\beta_1 u_1+.....+\beta_s u_s+\gamma_1 w_1+.....+\gamma_t w_t = 0$

Then $\beta_1 u_1+.....+\beta_s u_s = -(\gamma_1 w_1+.....+\gamma_t w_t)-(\alpha_1 v_1+.....+\alpha_r v_r) \in B$

Hence $\beta_1 u_1+.....+\beta_s u_s \in B$

Also $\beta_1 u_1+.....+\beta_s u_s \in A$

Hence $\beta_1 u_1+.....+\beta_s u_s \in A \cap B$

$\therefore \beta_1 u_1+.....+\beta_s u_s = \delta_1 v_1+.....\delta_r v_r$

$\therefore \beta_1 u_1+.....+\beta_s u_s-\delta_1 v_1-.....-\delta_r v_r = 0$

$\therefore \beta_1 = ... = \beta_s = \delta_1 = ... = \delta_r = 0$

(since $\{u_1.....u_s, v_1.....v_r\}$ is linearly Independent).

Similarly we can prove $\gamma_1 = \gamma_2 = \ldots\ldots = \gamma_t = 0$

$\therefore \alpha_i = \beta_j = \gamma_k = 0$ for $1 \le i \le r$

$1 \le j \le s$; $1 \le k \le t$.

Thus S′ is a linearly Independent set.

Clearly S′ spans A+B.

$\therefore$ S′ is a basis for A+B

Hence dim(A+B) = r+s+t

Also dim A = r+s, dim B = r+t, and dim (A$\cap$B) = r

$\therefore$ dim A + dim B –dim (A$\cap$B) = (r+s)+(r+t)–r

= r+s+t

= dim (A+B)

Aliter By the theorem, Let V be a vector space over a Field F. Let A and B be subspaces of V. Then $\frac{A+B}{A} \cong \frac{B}{A \cap B}$

Hence $\dim\left[\frac{A+B}{A}\right] = \dim\left[\frac{B}{A \cap B}\right]$

$\therefore$ dim(A+B)–dim A = dim B – dim (A$\cap$B)

$\therefore$ dim(A+B) = dim A + dim B – dim(A$\cap$B)

Corollary :

If V = A$\oplus$B, dim V = dim A + dim B.

Proof :

V = A$\oplus$B $\Rightarrow$ A+B = V

and A$\cap$B = {0}

$\therefore$ dim(A$\cap$B) = 0

Hence dim V = dim A + dim B.

Solved Problems :

1. Complete the set $\{(2,1,4,3), (2,1,2,0)\}$ to form a basis of $V_4(R)$.

Solution :

Since dim $V_4(R) = 4$, every basis of $V_4(R)$ contains four vectors,

Let $v_1 = (2,1,4,3)$,

$v_2 = (2,1,2,0)$

Then $\alpha_1 v_1+\alpha_2 v_2 = 0 \Rightarrow (2\alpha_1+2\alpha_2, \alpha_1+\alpha_2, 4\alpha_1+2\alpha_2, 3\alpha_1)$

$= (0,0,0,0)$

$\Rightarrow 2\alpha_1+2\alpha_2 = 0$

$\alpha_1+\alpha_2 = 0$

$4\alpha_1+2\alpha_2 = 0$

$3\alpha_1 = 0$

$\Rightarrow \alpha_1 = \alpha_2 = 0$

$A = \{v_1, v_2\}$ is linearly independent.

Now $L(A) = \{\alpha v_1+\beta v_2 / \alpha,\beta \text{ scalars}\}$

$= \{(2\alpha+2\beta, \alpha+\beta, 4\alpha+2\beta, 3\alpha) / \alpha,\beta \text{ scalars}\}$

We choose a vector outside this span L(A) and get an enlarged linearly independent set.

For each vector in L(A), the first co-ordinate $2\alpha+2\beta$ is twice the second co-ordinate $\alpha+\beta$. Hence $e_1 = (1,0,0,0)$ is not in this span. Thus, we get the enlarged linearly independent set $S_1 = \{v_1, v_2, R_1\}$.

Now $L(S_1) = \{\alpha v_1+\beta v_2+\gamma e_1 / \alpha,\beta,\gamma \text{ scalars}\}$

$= \{(2\alpha+2\beta+\gamma, \alpha+\beta, 4\alpha+2\beta, 3\alpha) / \alpha,\beta,\gamma \text{ scalars}\}$

As before we enlarge S_1 by choosing a vector outside $L(S_1)$

Clearly, $e_2 = (0,1,0,0)$ is not in $L(S_1)$

Let $S_2 = \{v_1, v_2, e_1, e_2\}$.

Then S_2 is linearly independent and hence is a basis for $V_4(R)$.

2. Let A = {(1,1,1,1), (1,2,1,2)} be a linearlyindependent subset of $V_4(R)$. Extend it to a basis of $V_4(R)$.

Solution :

$$\text{Let } v_1 = (1, 1, 1, 1), \quad v_2 = (1, 2, 1, 2)$$

$$\text{Then } L(A) = \{\alpha v_1+\beta v_2 / \alpha,\beta \text{ scalars}\}$$

$$= \{(\alpha+\beta, \alpha+2\beta, \alpha+\beta, \alpha+2\beta) / \alpha, \beta \text{ scalars}\}$$

For each vector in L(A) the first and third co-ordinates are equal to $\alpha+\beta$.

$\therefore$ $v_3 = (0,2,1,2)$ is not in L(A). Thus, we get th enlarged linearly independent set

$$S_1 = \{v_1,v_2,v_3\}$$

Now $$L(S_1) = \{\alpha v_1+\beta v_2+\gamma v_3 / \alpha,\beta,\gamma \text{ scalars}\}$$

$$= \{(\alpha+\beta, \alpha+2\beta+2\gamma, \alpha+\beta+\gamma, \alpha+2\beta+2\gamma) / \alpha,\beta,\gamma \text{ scalars}\}$$

Clearly $v_4 = (2,5,3,6)$ is not in $L(S_1)$

$\therefore$ $S_2 = \{v_1,v_2,v_3,v_4\}$ is a basis for $V_4(R)$

Exercise :

1. Find the dimension of the subspace spanned by the following vectors in $V_3(R)$.

   (a) (1,1,1), (–1,–1,–1)

   (b) (1,0,2), (2,0,1), (1,0,1)

   (c) (1,2,–3); (0,0,1), (–1,2,1)

   (d) (1,1,2) (–1,1,0)

2. Find the dimension of the subspace spanned by the following vectors in $V_4(R)$.

   (a) e_1, e_2, e_3, e_4

   (b) e_1, e_2

   (c) e_1, e_2, e_3

   (d) e_1

3. In $V_3(R)$, find dim(A+B) and dim(A$\cap$B) where

   (a) A is the subspace spanned by (1,1,1) and

   B is the subspace spanned by (–1,–1,–1)

   (b) A is the subspace spanned by (1, 1, 1) and

   B is the subspace spanned by (1, 2, 1)

   (c) A is the subspace spanned by (1, 1, 1) and (1,2,1) and

   B is the subspace spanned by (0, 0, 1)

   (d) A is the subspace spanned by (1, 1, 1) and (1, 2, 1) and

   B is the subspace spanned by (1, –1, 1) and (–1, 1, –1)

4. Let V_1 and V_2 be subspaces of V such that $V_1 \cap V_2$ is the zero space.

   Prove that dim V_1 + dim $V_2 \leq$ dim V.

5. Let V_1 and V_2 be subspaces of V such that every vector $v \in V$ can be represented as $v=v_1+v_2$ where $v_1 \in V_1$ and $v_2 \in V_2$ prove that dim V_1 + dim $V_2 \geq$ dim V.

6. If A and B are finite dimensional subspaces of V such that $A \subseteq B$ and dim A = dim B then show that A = B.

7. Let S be a subspace of a finite - dimensional vector space V. If dim V = dim S then prove that S = V.

8. Let W_1 and W_2 be two subspaces of a finite dimensional vector space V.

   If dim V = dim W_1 + dim W_2 and

   $W_1 \cap W_2 = \{0\}$ prove that, $V = W_1 \oplus W_2$.

Answers:

1.	(a)	1	(b)	2	(c)	3	(d)	2
2.	(a)	4	(b)	2	(c)	3	(d)	1
3.	(a)	1;1	(b)	2;0	(c)	3;0	(d)	3;0

2.1. SUBSPACE

Definition :

Let V be a vector space over a field F. A non-empty subset W of V is called a **subspace** of V if W itself is a vector space over F under the operations of V.

Theroem 2.1 :

Let V be a vector space over F. A non-empty subset W of V is a subspace of V iff W is closed with respect to vector addition and scalar multiplication in V.

Proof :

Let W be a subspace of V.

Then W itself is a vector space and hence W is closed with respect to vector addition and scalar multiplication. Conversely, let W be a non-empty subset of V such that

$$u, v \in W \Rightarrow u+v \in W$$

and $u \in W$ and $\alpha \in F \Rightarrow \alpha u \in W$

We prove that W is a subspace of V.

Since W is non-empty, these exists an element $u \in W$.

$\therefore 0u = 0 \in W$

Also $v \in W \Rightarrow (-1)v = -v \in W$

Thus W contains 0 and the additive inverse of each of its element.

Hence W is an additive subgroup of V.

Also $u \in W$ and $\alpha \in F \Rightarrow \alpha u \in W$.

Since the elements of W are the elements of V the other axioms of a vector space are true in W.

Hence W is a subspace of V.

Theorem 2.2 :

Let V be a vector space over a field F. A non-empty subset W of V is a subspaceof V iff $u,v \in W$ and $\alpha,\beta \in F \Rightarrow \alpha u+\beta v \in W$.

Proof :

Let W be a subspace of V.

Let $u,v \in W$ and $\alpha,\beta \in F$

Then by theorem 2.1, αu and $\beta v \in W$ and hence $\alpha u+\beta v \in W$.

Conversely,

Let $u,v \in W$ and $\alpha,\beta \in F \Rightarrow \alpha u+\beta v \in W$.

Taking $\alpha=\beta=1$, we get $u,v \in W \Rightarrow u+v \in W$.

Taking $\beta=0$, we get

$\alpha \in F$ and $u \in W \Rightarrow \alpha u \in W$

Hence by theorem 2.1.

W is a subspace of V.

Examples :

1) $\{0\}$ and V are subspaces of any vector space V. They are called the trivial subspaces of V.

2) $W = \{(a,0,0)/a \in R\}$ is a subspace of R^3 for, let $u = (a,0,0)$, $v=(b,0,0) \in W$ and $\alpha,\beta \in R$.

Then $$\alpha u+\beta v = \alpha(a,0,0)+\beta(b,0,0)$$
$$= (\alpha a+\beta b,0,0) \in W$$

Hence W is a subspace of R^3.

Note : Geometrically W consists of all points on the x-axis in the Euclidean 3 space.

3) In R^3, $W = \{(ka, kb, kc) / k \in R\}$ is a subspace of R^3.

For if $u = (k_1a, k_1b, k_1c)$

and $v = (k_2a, k_2b, k_2c) \in W$ and $\alpha,\beta \in R$

Then $\alpha u+\beta v = \alpha(k_1a, k_1b, k_1c)+\beta(k_2a, k_2b, k_2c)$

$= (\alpha k_1+\beta k_2)a, (\alpha k_1+\beta k_2)b, (\alpha k_1+\beta k_2)c \in W$

Hence W is a subspace of R^3.

Note :

Geometrically W consists of all points of the line $\frac{x}{a}=\frac{y}{b}=\frac{z}{c}$ provided a, b, c are not all zero.Thus the set of all points on a line through the origin is a subspace of R^3. However a line not passing through the origin is not a subspace of R^3, sinc the additive identity (0,0,0) does not lie on the line.

4. $W = \{(a,b,0)/a,b\in R\}$ is a subspace of R^3 W consists of all points of the xy-plane.

5. Let W be the set of all points in R^3 satisfying the equation $lx+my+nz=0$. W is a subspace of R^3. For, let $u=(a_1,b_1,c_1)$ and $v = (a_2,b_2,c_2)\in W$ and $\alpha,\beta\in R$

Then we have

$la_1+mb_1+nc_1 = 0 = la_2+mb_2+nc_2$

Hence $\alpha(la_1+mb_1+nc_1)+\beta(la_2+mb_2+nc_2) = 0$

(i.e.,) $l(\alpha a_1+\beta a_2)+m(\alpha b_1+\beta b_2)+n(\alpha c_1+\beta c_2) = 0$

(i.e.,) $\alpha u+\beta v\in W$ so that W is a subspace of R^3.

Note :

Geometrically W consists of all points on the plane $lx+my+nz = 0$, which passes through the origin.

6. Let $W = \{f/f\in F[x]$ and $f(a) = 0\}$

(i.e.,) W is the set of all polynomials in F[x] having a as a root where $a\in F$. Then W is a vector space over F.

W observe that $x-a\in W$ and hence W is non-empty.

Let $f,g\in F[x]$ and $\alpha,\beta\in F$

To prove that $\alpha f+\beta g\in W$. We have to show that a is a root of $\alpha f+\beta g$.

Now, $(\alpha f+\beta g)(a) = \alpha f(a)+\beta g(a)$

$= \alpha 0+\beta 0 = 0$

Hence a is a root of $\alpha f+\beta g$.

$\therefore \alpha f+\beta g \in W$ and hence W is a subspace of F[x].

7. $W = \left\{ \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} / a,b \in R \right\}$ is a subspace of $M_2(R)$.

Solved Problems :

Problem 1:

Prove that the intersection of two subspaces of a vector space is a subspace.

Solution :

Let A and B be two subspaces of a vector space V over a field F.

We claim that $A \cap B$ is a subspace of V.

Clearly $0 \in A \cap B$ and hence $A \cap B$ is non-empty.

Now, let $u,v \in A \cap B$ and $\alpha,\beta \in F$

Then $u,v \in A$ and $u,v \in B$

$\therefore \alpha u+\beta v \in A$ and $\alpha u+\beta v \in B$ (since A and B subspaces)

$\therefore \alpha u+\beta v \in A \cap B$

Hence $A \cap B$ is a subspace of V.

Problem 2 :

Prove that the union of two subspaces of a vector space need not be a subspace.

Solution :

Let $A = \{(a,0,0)/a \in R\}$

$B = \{(0,b,0)/b \in R\}$

Clearly A and B are subspaces of R^3 (by example 2).

However $A \cup B$ is not a subspace of R^3.

For $(1,0,0)$ and $(0,1,0) \in A \cup B$

But $(1,0,0)+(0,1,0) = (1,1,0) \notin A \cup B$

Problem 3 :

Prove that the union of two subspaces of a vector space is a subspace iff one is contained in the other.

Proof :

Let H and K be two subspaces of G such that one is contained in the other. Hence either $H \subseteq K$ or $K \subseteq H$.

$\therefore$ $H \cup K = K$ (or) $H \cup K = H$

Hence $H \cup K$ is a subspace of G.

Conversely, suppose $H \cup K$ is a subspace of G.

We claimthat $H \subseteq K$ (or) $K \subseteq H$.

Suppose that H is not contained in K and K is not contained in H. Then there exist elements a, b such that

$a \in H$ and $a \notin K$ ---------(1)

$b \in K$ and $b \notin H$ ---------(2)

Clearly $a,b \in H \cup K$. Since $H \cup K$ is a subspace of G, $a,b \in H \cup K$. Hence $ab \in H$ (or) $ab \in K$.

Case (i) :

Let $ab \in H$. Since $a \in H$, $a^{-1} \in H$.

Hence $a^{-1}(ab) = b \in H$ which is contradiction to (2).

Case (ii) :

Let $ab \in K$. Since $b \in K$, $b^{-1} \in K$.

Hence $(ab)b^{-1} = a \in K$ which is a contradiction to (1). Hence our assumption that H is not contained in K and K is not contained in H is false.

$\therefore$ $H \subseteq K$ or $K \subseteq H$.

Problem 4 :

If A and B are subspaces of V prove that A+B = {v∈V/v=a+b, a∈A, b∈B} is a subspace of V. Further show that A+B is the smallest subspace containing A and B. (i.e.,) If W is any subspace of V containing A and B then W contains A+B.

Solution :

Let $v_1, v_2 \in A+B$ and $\alpha \in F$

Then $v_1 = a_1+b_1$, $v_2 = a_2+b_2$ where $a_1, a_2 \in A$ and $b_1, b_2 \in B$

Now, $v_1+v_2 = (a_1+b_1)+(a_2+b_2)$

$= (a_1+a_2)+(b_1+b_2) \in A+B$

Also, $\alpha(a_1+b_1) = \alpha a_1+\alpha b_1 \in A+B$

Hence A+B is a subspace of V. Clearly $A \subseteq A+B$ and $B \subseteq A+B$

Now, let W be any subspace of V containing A and B.

We shall prove tht $A+B \subseteq W$.

Let $v \in A+B$. Then v=a+b where $a \in A$ and $b \in B$.

Since $A \subseteq W$, $a \in W$. Similarly $b \in W$

$a+b = v \in W$

$\therefore A+B \subseteq W$ so that A+B is the smallest subspace of V containing A and B.

Problem 5 :

Let A and B be subspace of a vector space V. Then $A \cap B = \{0\}$ iff every vector $v \in A+B$ can be uniquely expressed in the form v=a+b where $a \in A$ and $b \in B$.

Solution :

Let $A \cap B = \{0\}$

Let $v \in A+B$

Let $v = a_1+b_1 = a_2+b_2$ where $a_1, a_2 \in A$ and $b_1, b_2 \in B$

Then $a_1-a_2 = b_2-b_1$

But $a_1-a_2 \in A$ and $b_2-b_1 \in B$

Hence $a_1-a_2, b_2-b_1 \in A \cap B$

Since $A \cap B = \{0\}$, $a_1 - a_2 = 0$ and $b_2 - b_1 = 0$ so that $a_1 = a_2$ and $b_1 = b_2$. Hence the expression of v in the form a+b where $a \in A$ and $b \in B$ is unique.

Conversely suppose that any element in A+B can be uniquely expressed in the form a+b where $a \in A$ and $b \in B$.

We claim that $A \cap B = \{0\}$

If $A \cap B \neq \{0\}$, let $v \in A \cap B$ and $v \neq 0$

Then $0 = v - v = 0 + 0$

Thus 0 has been expressed in the form a+b in two different ways which is a contradiction.

Hence $A \cap B = \{0\}$

Definition :

Let A and B be subspaces of a vector space V. Then V is called the **direct sum** of A and B if (i) $A+B = V$ (ii) $A \cap B = \{0\}$. If V is the direct sum of A and B we write $V = A \oplus B$.

Note :

$V = A \oplus B$ iff every element of V can be uniquely expressed in the form a+b where $a \in A$ and $b \in B$.

Examples :

1. In $V_3(R)$. Let $A = \{(a,b,0)/a,b \in R\}$ and $B = \{(0,0,c)/c \in R\}$. Clearly A and B are subspaces of V and $A \cap B = \{0\}$. Also let $v = (a,b,c) \in V_3(R)$.

Then $v = (a,b,0)+(0,0,c)$ so that

$A+B = V_3(R)$

Hence $V_3(R) = A \oplus B$.

2. In $M_3(R)$, let A be the set of all matrices of the form $\begin{bmatrix} a & b \\ 0 & 0 \end{bmatrix}$ and B be the set of all matrices of the form $\begin{bmatrix} 0 & 0 \\ c & d \end{bmatrix}$. Clearly A and B are subspaces of $M_2(R)$ and $A \cap B = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ and $A+B = M_2(R)$

Hence $M_2(R) = A \oplus B$

3. If a_1, a_2, a_3 are fixed elements of a field F, then the set W of all ordered trials (x_1,x_2,x_3) of element F, such that $a_1x_1+a_2x_2+a_3x_3 = 0$ is a subspace of $V_3(F)$.

Solution :

Let $\alpha = (x_1,x_2,x_3)$ & $\beta = (y_1,y_2,y_3)$ be any two elements of w. Then x_1, x_2, x_3, y_1, y_2, y_3 are element of F and are such that

$$a_1x_1+a_2x_2+a_3x_3 = 0$$

$$a_1y_1+a_2y_2+a_3y_3 = 0$$

If a, b be any elements of F,

We have $a\alpha+b\beta = a(x_1,x_2,x_3) + b(y_1,y_2,y_3)$

$= (ax_1,ax_2,ax_3)+(by_1,by_2,by_3)$

$= (ax_1+by_1, ax_2+by_2, ax_3+by_3)$

Now, $a_1(ax_1+by_1)+a_2(ax_2+by_2)+a_3(ax_3+by_3)$

$= a(a_1x_1+a_2x_2+a_3x_3)+b(a_1y_1+a_2y_2+a_3y_3)$

$= a0+b0$

$= 0$

$\therefore\ a\alpha+b\beta = \{ax_1+by_1, ax_2+by_2, ax_3+by_3\} \in W$

Hence W is a subspace of $V_3(F)$.

Exercise :

1. Show that the following subsets of R^3 are subspaces Interpret them geometrically.

   a) $\{(a,0,c)/a,c \in R\}$

   b) $\{(a,b,c)/a=b=c\}$

   c) $\{(a,b,c)/a=b+c\}$

   d) $\{(a,b,a+b)/a,b \in R\}$

2. Show that the set of all polynomials in R[x] having atleast one rational root is not a subspace of R[x].

3. Let W be a subspace of V and U be a subspace of W. Then show that U is a subspace of V.

4. Show that each of the following subsets of $V_3(R)$ is not a subspace.

   i) $S = \{(x,y,z) / x^2+y^2+z^2 \leq 1\}$

   ii) $S = \{(x,y,z) / x+y+z = 1\}$

   iii) $S = \{(x,y,z) / x \geq y \geq z\}$

Theorem (U.Q) 2.3 :

Let V be a vector space over F and W a subspace of V. Let $\frac{V}{W} = \{W+v/v \in V\}$.

Then $\frac{V}{W}$ is a vector space over F under the following operations

i) $(W+v_1)+(W+v_2) = W+v_1+v_2$

ii) $\alpha(W+v_1) = W+\alpha v_1$

Proof :

Since W is a subspace of V it is a subgroup of (V, +). Since (V, +) is abelian, W is a normal subgroup of (V, +) so that (i) is a well defined operation.

Now, we shall prove that (ii) is a well defined operation.

$$W+v_1 = W+v_2 \Rightarrow v_1-v_2 \in W$$

$$\Rightarrow \alpha(v_1-v_2) \in W \text{ (since W is a subspace)}$$

$$\Rightarrow \alpha v_1-\alpha v_2 \in W$$

$$\Rightarrow \alpha v_1 \in W+\alpha v_2$$

$$\Rightarrow W+\alpha v_1 = W+\alpha v_2$$

Hence (ii) is a well defined operation.

Now, let $W+v_1, W+v_2, W+v_3 \in \frac{V}{W}$.

Then $(W+v_1)+[(W+v_2)+(W+v_3)] = (W+v_1)+(W+v_2+v_3)$

$= (W+v_1)+(W+v_2+v_3)$

$= W+v_1+v_2+v_3$

$= (W+v_1+v_2)+(W+v_3)$

$= [(W+v_1)+(W+v_2)]+(W+v_3)$

Hence '+' is associative.

$W+0 = W \in \frac{V}{W}$ is the additive identity element.

For $(W+v_1)+(W+0) = W+v_1$

$= (W+0)+(W+v_1) \ \forall \ v_1 \in V$

Also $W-v_1$ is the additive inverse of $W+v_1$.

Hence $\frac{V}{W}$ is a group under +.

Further $(W+v_1)+(W+v_2) = W+v_1+v_2$

$= W+v_2+v_1$

$= (W+v_2)+(W+v_1)$

Hence $\frac{V}{W}$ is an abelian group. Now, let $\alpha,\beta \in F$.

$\alpha[(W+v_1)+(W+v_2)] = \alpha(W+v_1+v_2)$

$= W+\alpha(v_1+v_2)$

$= W+\alpha v_1+\alpha v_2$

$= (W+\alpha v_1)+(W+\alpha v_2)$

$= \alpha(W+v_1)+\alpha(W+v_2)$

$(\alpha+\beta)(W+v_1) = W+(\alpha+\beta)v_1$

$= W+\alpha v_1+\beta v_1$

$= (W+\alpha v_1)+(W+\beta v_1)$

$(\alpha+\beta)(W+v_1) = \alpha(W+v_1)+\beta(W+v_1)$

$\alpha[\beta(W+v_1)] = \alpha(W+\beta v_1)$

$= W+\alpha\beta v_1$

$$\alpha[\beta(W+v_1)] = (\alpha\beta)(W+v_1)$$
$$1(W+v_1) = W+1.v_1$$
$$= W+v_1$$

Hence $\frac{V}{W}$ is a vector space.

The vector space $\frac{V}{W}$ is called the **quotient space** of V by W.

2.2. LINEAR TRANSFORMATION

Definition :

Let V and W be vector spaces over a field F A mapping $T:V\to W$ is called a homomorphism if

i) $T(u+v) = T(u)+T(v)$ and

ii) $T(\alpha u) = \alpha T(u)$ where $\alpha\in F$ and $u,v\in V$.

A homomorphism T of vector spaces is also called a **linear Transformation.**

i) If T is 1–1 then T is called **monomorphism.**

ii) If T isonto then T is called an **epimorphism.**

iii) If T is 1–1 and onto T is called an **isomorphism.**

iv) Two vector spaces V and W are said to be isomorphic if there exists an isomorphism T from V to W and we write $V\cong W$.

v) A linear transformation $T:V\to F$ is called a **linear functional.**

Examples :

1. $T:V\to W$ defined by $T(v) = 0\ \forall v\in V$ is a **trivial linear transformation.**

2. $T:V\to V$ defined by $T(v) = v\ \forall v\in V$ is the **identity lienar transformation.**

3. Let V be a vector space over a field F and W a subspaceof V. Then $T:V\to\frac{V}{W}$ defined by $T(v) = W+v$ is a linear transformation for

$$T(v_1+v_2) = W+(v_1+v_2)$$
$$= (W+v_1)+(W+v_2)$$
$$= T(v_1)+T(v_2)$$

Also $T(\alpha v_1) = W+\alpha v_1$

$= \alpha(W+v_1) = \alpha T(v_1)$

This is called the **natural homomorphism** from V to $\frac{V}{W}$.

Clearly T is onto and hence T is an epimorphism.

4. $T:V_3(R)\to V_3(R)$ defined by

$T(a,b,c) = (a,0,0)$ is a linear transformation.

5. Let V be the set of all polynomial of degree $\leq n$ in R[x] including the zero polynomial $T:V\to V$ defined by $T(f) = \frac{df}{dx}$ is a linear transformation.

For, $T(f+g) = \frac{d(f+g)}{dx} = \frac{df}{dx}+\frac{dg}{dx}$

$= T(f)+T(g)$

Also $T(\alpha f) = \frac{d(\alpha f)}{dx} = \alpha\frac{df}{dx} = \alpha T(f)$

6. Let V be as in Example 5. Then $T:V\to V_{n+1}(R)$ defined by $T(a_0+a_1x+....+a_0x^n) = (a_0,a_1,.....a_n)$ is a linear transformation.

For, Let $f = a_0+a_1x+.....+a_nx^n$ and

$g = b_0+b_1x+.....b_nx^n$

Then $f+g = (a_0+b_0)+(a_1+b_1)x+.....+....(a_n+b_n)x^n$

$T(f+g) = ((a_0+b_0),(a_1+b_1),.....(a_n+b_n))$

$= (a_0,a_1,.....,a_n)+(b_0,b_1,.....,b_n)$

$T(f+g) = T(f)+T(g)$

Also $T(\alpha f) = (\alpha a_0, \alpha a_1,.....\alpha a_n)$

$= \alpha(a_0,a_1,.....,a_n)$

$= \alpha T(f).$

Clearly T is 1–1 and onto and hence T is an isomorphism.

7. Let V denote the set of all sequences in R.

$T : V\to V$ defined by $T(a_1,a_2,.....,a_n,....) = (0,a_1,a_2,.....,a_n,....)$ is a linear transformation.

8. $T:R^2 \to R^2$ defined by $T(a,b) = (2a-3b, a+4b)$ is a linear transformation.

Let $u = (a, b)$ and $v = (c, d)$ and $\alpha \in R$.

$\therefore T(u+v) = T((a,b) + (c,d))$

$= T(a+c, b+d)$

$= (2(a+c)-3(b+d), (a+c)+4(b+d))$

$= (2a+2c-3b-3d, a+c+4b+4d)$

$= (2a-3b+2c-3d, a+4b+c+4d)$

$= (2a-3b, a+4b)+(2c-3d, c+4d)$

$= T(a,b)+T(c,d)$

$\therefore T(u+v) = T(u)+T(v)$

Also $T(\alpha u) = T(\alpha(a,b))$

$= T(\alpha a, \alpha b)$

$= (2\alpha a-3\alpha b, \alpha a+4\alpha b)$

$= \alpha(2a-3b, a+4b)$

$= \alpha T(a, b)$

$T(\alpha u) = \alpha T(u)$

$\therefore$ T is a linear transformation.

Theorem 2.4 :

Let $T:V \to W$ be a linear transformation.

Then $T(V) = \{T(v)/v \in V\}$ is a subspace of W.

Proof :

Let w_1 and $w_2 \in T(V)$ and $\alpha \in F$. Then there exists $v_1, v_2 \in V$ such that $T(v_1)=w_1$ and $T(v_2) = w_2$.

Hence $w_1+w_2 = T(v_1)+T(v_2)$

$= T(v_1+v_2) \in T(v)$

Similarly $\alpha w_1 = \alpha T(v_1)$

$= T(\alpha v_1) \in T(v)$

Hence T(v) is a subspace of W.

Definition :

Let V and W be vector spaces over a field F and T:V→W be a linear transformation. Then the **Kernal** of T is defined to be {v/v∈V and T(v)=0} and is denoted by Ker T.

Thus Ker T = {v/v∈V and T(v) = 0}

Example :

1. T:V→W defined by T(v) = 0 ∀ v∈V is trial linear transformation.

   ∴ Ker T = V.

2. T:V→V defined by T(v)=v ∀v∈V is the identity linear transformation. Ker T={0}

Note : Let T:V→W be a linear transformation. Then T is a monomorphism iff Ker T={0}.

3. Let V be the set of all polynomials of degree ≤ n in R[x] including the zero polynomial T:V→V defined by $T(f) = \frac{df}{dx}$ is a linear transformation.

For, $$T(f+g) = \frac{d(f+g)}{dx} = \frac{df}{dx} + \frac{dg}{dx} = T(f)+T(g)$$

Also $$T(\alpha f) = \frac{d(\alpha f)}{dx} = \alpha \frac{df}{dx} = \alpha T(f)$$

∴ T:V→W be a linear transformation Ker T is the set of all constant polynomials.

(U:Q) Theorem 2.5 : (Fundamental Theorem of homomorphism)

Let V and W be vector spaces over a field F and T:V→W be an epimorphism.

Then i) Ker T = V_1 is a subspace of V and

ii) $\frac{V}{V_1} \cong W$

Proof :

i) Given V_1 = Ker T

= {v/v∈V and T(v) = 0}

Clearly T(0) = 0

Hence $0 \in \text{Ker } T = V_1$

$\therefore$ V_1 is non-empty subset of V.

Let $u,v \in \text{Ker } T$ and $\alpha,\beta \in F$

$\therefore$ $T(u) = 0$ and $T(v) = 0$

Now $T(\alpha u+\beta v) = T(\alpha u)+T(\beta v)$

$= \alpha T(u)+\beta T(v)$

$= \alpha 0+\beta 0$

$T(\alpha u+\beta v) = 0$

$\alpha u+\beta v \in \text{Ker } T$.

By the theorem :

Let V be a vector space over a field F. A non-empty subset W of V is a subspace of V iff $u,v \in W$ and $\alpha,\beta \in F \Rightarrow \alpha u+\beta v \in W$.

Ker T is a subspace of V.

ii) We define a map $\phi: \frac{V}{V_1} \to W$ by $\phi(V_1+v) = T(v)$.

ϕ is well defined.

Let $V_1+v = V_1+w$

$\therefore$ $v \in V_1+w$

$\therefore$ $v = v_1+w$ where $v_1 \in V_1$

$\therefore$ $T(v) = T(v_1+w) = T(v_1)+T(w)$

$= 0+T(w) = T(w)$.

$\therefore$ $\phi(V_1+v) = \phi(V_1+w)$

ϕ is 1–1 :

$\phi(V_1+v) = \phi(V_1+w)$

$\Rightarrow T(v) = T(w)$

$\Rightarrow T(v)-T(w) = 0$

$\Rightarrow T(v)+T(-w) = 0$

$\Rightarrow T(v-w) = 0$

$\Rightarrow v-w \in \text{Ker } T = V_1$

$\Rightarrow v \in V_1+w$

$\Rightarrow V_1+v = V_1+w$

ϕ is onto :

Let $w \in W$

Since T is onto there exists $v \in V$. Such that $T(v) = w$.

$\therefore \quad \phi(V_1+v) = w$

$\phi(V_1+v) = w$

ϕ is a homomorphism.

$$\phi[(V_1+v)+(V_1+w)] = \phi[V_1+(v+w)]$$
$$= T(v+w)$$
$$= T(v)+T(w)$$
$$= \phi(V_1+v)+\phi(V_1+w)$$

Also $\phi[\alpha(V_1+v)] = \phi(V_1+\alpha v)$

$$= T(\alpha v)$$
$$= \alpha T(v)$$
$$= \alpha T(V_1+v)$$

$\therefore \phi$ is an isomorphism from $\frac{V}{V_1}$ onto W.

$$\therefore \quad \frac{V}{V_1} \cong W$$

Theorem 2.6 (U.Q) :

Let V be a vector space over a field F. Let A and B subspaces of V. Then

$$\frac{A+B}{A} \cong \frac{B}{A \cap B}$$

Proof :

We know that A+B is a subspace of V containing A.

Hence $\frac{A+B}{A}$ is also a vector space over F.

An element of $\frac{A+B}{A}$ is of the form A+(a+b) where $a \in A$ and $b \in B$.

But A+a = A

Hence an element of $\frac{A+B}{A}$ is of the form A+b.

Now, consider $f:B \to \frac{A+B}{A}$ defined by f(b) = A+b.

Clearly f is onto.

Also $f(b_1+b_2) = A+(b_1+b_2)$

$= (A+b_1)+(A+b_2)$

$f(b_1+b_2) = f(b_1)+f(b_2)$

and $f(\alpha b_1) = A+\alpha b_1$

$= \alpha(A+b_1)$

$= \alpha f(b_1)$

Hence f is a linear transformation.

Let K be the Kernel of f.

Then $K = \{b/b \in B, A+b=A\}$

Now, $A+b = A$ iff $b \in A$

Hence $K = A \cap B$

By the Fundamental theorem of homomorphism.

$$\frac{B}{A \cap B} \cong \frac{A+B}{A}$$

Theorem 2.7 :

Let V and W be vector spaces over a field F. Let L(V,W) represent the set of all linear transformations from V to W. Then L(V,W) itself is a vector space over F under addition and scalar multiplication defined by

$$(f+g)(v) = f(v)+g(v)$$

$$\text{and } (\alpha f)(v) = \alpha f(v)$$

Proof :

Let $f, g \in L(V,W)$ and $v_1, v_2 \in V$.

Then $(f+g)(v_1+v_2) = f(v_1+v_2)+g(v_1+v_2)$

$= f(v_1)+f(v_2)+g(v_1)+g(v_2)$

$= f(v_1)+g(v_1)+f(v_2)+g(v_2)$

$= (f+g)(v_1)+(f+g)(v_2)$

Also $(f+g)(\alpha v) = f(\alpha v)+g(\alpha v)$

$= \alpha f(v)+\alpha g(v)$

$= \alpha[f(v)+g(v)]$

$= \alpha(f+g)(v)$

Hence $(f+g) \in L(V,W)$

Now, $(\alpha f)(v_1+v_2) = (\alpha f)(v_1)+(\alpha f)(v_2)$

$= \alpha f(v_1)+\alpha f(v_2)$

$= \alpha[f(v_1)+f(v_2)]$

$= \alpha f(v_1+v_2)$

Also $(\alpha f)(\beta v) = \alpha[f(\beta v)] = \alpha(\beta f(v))$

$= \beta[\alpha f(v)] = \beta[(\alpha f)(v)]$

Hence $\alpha f \in L(V,W)$

Addition defined on L(V,W) is obviously commutative and associative.

The function $f:V \to W$ defined by $f(v)=0$ for all $v \in V$ is clearly a linear transformation and is the additive identity of L(V,W).

Further $(-f):V \to W$ defined by

$(-f)(v) = -f(v)$ is the additive inverse of f.

Thus L(V,W) is an abelian group under addition.

The remaining axioms for a vector space can be easily verified.

Hence L(V,W) is a vector space over F.

Exercises :

1. Let V and W be vector spaces over a field F. Show that a mapping $T:V \to W$ is a linear transformation iff

$T(\alpha v_1+\beta v_2) = \alpha T(v_1)+\beta T(v_2)$ for all $v_1, v_2 \in V$ and $\alpha, \beta \in F$.

2. **Find the Kernel of the following linear transformations.**

i) $T:V_4(R)\rightarrow V_4(R)$ defined by

$T(x_1,x_2,x_3,x_4) = (x_1,0,x_3,0)$

ii) $T:V_3(R)\rightarrow V_3(R)$ defined by

$T(a, b, c) = (a, b, 0)$

3. **Prove that** $T:V_3(R)\rightarrow R$ **defined by** $T(x,y,z) = x^2+y^2+z^2$ **is not a linear transformation.**

Answers :

2. **(i) Not 1–1, not onto; Kernel** $\{0, a, 0, b) / a,b\in R\}$

(ii) Not 1–1, not onto; Kernel $\{(0, 0, c) / c\in R\}$

2.3. MATRIX OF A LINEAR TRANSFORMATION

Definition :

Let V and W be finite dimensional vector spaces over a field F.

Let dim V = m and dim W = n.

Fix an ordered basis $\{v_1,v_2,....v_m\}$ for V and an ordered basis $\{w_1,w_2,....,w_n\}$ for W.

Let $T:V\rightarrow W$ be a linear transformation. We have seen that T is completely specified by the elements $T(v_1)$, $T(v_2)$.....$T(v_m)$.

Now, let

$$\left.\begin{aligned} T(v_1) &= a_{11}w_1+a_{12}w_2+.....+a_{1n}w_n \\ T(v_2) &= a_{21}w_1+a_{22}w_2+.....+a_{2n}w_n \\ &\cdots\cdots\quad\cdots\cdots\quad\cdots\cdots \\ &\cdots\cdots\quad\cdots\cdots\quad\cdots\cdots \\ T(v_m) &= a_{m1}w_1+a_{m2}w_2+.....+a_{mn}w_n \end{aligned}\right\} \quad \text{---------(1)}$$

Hence $T(v_1)$, $T(v_2)$.....$T(v_m)$ are completely specified by the mn elements a_{ij} of th field F. These a_{ij} can be conveniently arranged in the form of m rows and n columns as follows :

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

Such an array of mn elements of F arranged in m rows and n columns is known as m×n matrix over the field F and is denoted by (a_{ij}). Thus to every linear transformation T there is associated withit an m×n matrix over F.

Conversely, any m×n matrix over F defines a linear transformation $T:V\to W$ given by the formula (1).

Note :

The m×n matrix which we have associated with a linear transformation $T:V\to W$ depends on the choice of the basis for V and W.

For example, consider the linear transformation $T:V_2(R)\to V_2(R)$ given by $T(a,b) = (a,a+b)$.

Choose $\{e_1, e_2\}$ as a basis both for the domain and the range.

Then $T(e_1) = (1, 1) = e_1+e_2$

$T(e_2) = (0, 1) = e_2$

Hence the matrix representing T is $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$.

Now, we choose $\{e_1, e_2\}$ as a basis for the domain and $\{(1,1), (1, -1)\}$ as a basis for the range.

Let $w_1 = (1, 1)$ and $w_2 = (1, -1)$.

Then $T(e_1) = (1, 1) = w_1$, and $T(e_1) = (0, 1) = \left(\frac{1}{2}\right)w_1 - \left(\frac{1}{2}\right)w_2$

Hence the matrix representing T in $\begin{bmatrix} 1 & 0 \\ 1/2 & -1/2 \end{bmatrix}$

Solved Problems :

1. Obtain the matrix representing the linear transformation $T:V_3(R)\to V_3(R)$ given by $T(a,b,c) = (3a, a-b, 2a+b+c)$ w.r.t. the standard basis $\{e_1, e_2, e_3\}$.

Solution :

$$T(e_1) = T(1,0,0) = (3,1,2) = 3e_1+e_2+2e_3$$

$$T(e_2) = T(0,1,0) = (0,-1,1) = -e_2+e_3$$

$$T(e_3) = T(0,0,1) = (0,0,1) = e_3$$

Thus the matrix representing T is $\begin{bmatrix} 3 & 1 & 2 \\ 0 & -1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$

2. Find the linear transformation

$T : V_3(R) \to V_3(R)$ determined by the matrix $\begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ -1 & 3 & 4 \end{bmatrix}$ w.r.t. the standard basis $\{e_1, e_2, e_3\}$.

Solution :

$$T(e_1) = e_1+2e_2+e_3 = (1, 2, 1)$$

$$T(e_2) = 0e_1+e_2+e_3 = (0, 1, 1)$$

$$T(e_3) = -e_1+3e_2+4e_3 = (-1, 3, 4)$$

Now, $(a, b, c) = a(1, 0, 0)+b(0, 1, 0)+c(0, 0, 1)$

$$= ae_1+be_2+ce_3$$

$$\therefore T(a, b, c) = T(ae_1+be_2+ce_3)$$

$$= aT(e_1)+bT(e_2)+cT(e_3)$$

$$= a(1,2,1)+b(0,1,1)+c(-1,3,4)$$

$$T(a, b, c) = (a-c, 2a+b+3c, a+b+4c)$$

$\therefore$ This is the required linear transformation.

Exercises :

1. Obtain the matrices for the following linear transformations.

   a) $T:V_2(R) \to V_2(R)$ given by $T(a,b) = (-b, a)$ w.r.t.

      i) standard basis

      ii) the basis $\{(1, 2), (1, -1)\}$ for both domain and range.

b) $T:V_3(R)\to V_2(R)$ given by $T(a, b, c) = (a+b, 2c-a)$ w.r.t.

i) standard basis

ii) $\{(1,0,-1), (1, 1, 1), (1,0,0)\}$ as a basis for $V_3(R)$ and $\{(0,1) (1,0)\}$ for $V_2(R)$.

2. Obtain the linear transformation determined by the following matrices.

a) $T:V_2(R)\to V_2(R)$ given by $\begin{bmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{bmatrix}$ w.r.t. the standard basis.

b) $T:V_3(R)\to V_3(R)$ given by $\begin{bmatrix}a & b & c\\ b & c & a\\ c & a & b\end{bmatrix}$ w.r.t. the standard basis.

c) $T:V_2(R)\to V_3(R)$ given by $\begin{bmatrix}2 & 1 & -1\\ 1 & 1 & -1\end{bmatrix}$ w.r.t. the standard basis.

Answers :

1. (a) (i) $\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}$ (ii) $\begin{bmatrix}-\frac{1}{3} & -\frac{5}{3}\\ \frac{2}{3} & \frac{1}{3}\end{bmatrix}$

(b) (i) $\begin{bmatrix}1 & 1\\ 1 & 0\\ 0 & 2\end{bmatrix}$ (ii) $\begin{bmatrix}-3 & 1\\ 1 & 2\\ -1 & 1\end{bmatrix}$

2. (a) $T(a, b) = (a\cos\theta+b\sin\theta, -a\sin\theta+b\cos\theta)$

(b) $T(x, y, z) = (ax+by+cz, bx+cy+az, cx+ay+bz)$

(c) $T(a, b) = (2a+b, a+b, -a-b)$

MATRICES

In this chapter we shall develop the general theory of matrices. Throughout this chapter we deal with matrices whose entries are from the field F of real or complex numbers.

3.1. ALGEBRA OF MATRICES

Definition :

Let F be an arbitrary field. A rectangular array of the form

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

where the a_{ij} are scalars in F, is called a matrix over F or simply a matrix.

The above matrix is also denoted by

(a_{ij}), i = 1,2,.....,m, j = 1,2,.....,n or simply by (a_{ij})

The m horizontal n-tuples

$(a_{11}, a_{12},.....,a_{1n})$, $(a_{21},a_{22},.....a_{2n})$,.....$(a_{m1}, a_{m2},.....a_{mn})$ are the rows of the matrix, and the n vertical m-tuples.

$$\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}, \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix}, \cdots\cdots\cdots \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix}$$ are its column.

Note that the element a_{ij}, called the ij - entry or ij - component, appears in the i^{th} row and the j^{th} column.

A matrix with m rows and n column is called an m by n matrix, or m×n matrix; the pair of numbers (m, n) is called its size or shape.

If m = n, A is called a **square matrix** of order n.

Definition :

Two matrices A = (a_{ij}) and B= (b_{ij}) are said to be **equal** if A and B have the same number of rows and columns and the corresponding entires in the two matrices are same.

3.2. MATRIX ADDITION AND SCALAR MULTIPLICATION

Definition :

Let A and B be two matrices with the same size. (i.e.,) the same number of rows and columns, say m×n matrices.

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \quad \text{and } B = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{bmatrix}$$

The sum of A and B, written A+B, is the matrix obtained by adding corresponding entries.

$$A+B = \begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \cdots & a_{2n}+b_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \cdots & a_{mn}+b_{mn} \end{bmatrix}$$

Example :

$$\text{If } A = \begin{bmatrix} 1 & 3 \\ 2 & 5 \\ 8 & 7 \end{bmatrix} \text{ and } B = \begin{bmatrix} 2 & 0 \\ 1 & 3 \\ 0 & -1 \end{bmatrix}$$

$$\text{Then } A+B = \begin{bmatrix} 3 & 3 \\ 3 & 8 \\ 8 & 6 \end{bmatrix}$$

The product of a scalar k by the matrix A, written k.A or simply kA, is the matrix obtained by multiplying each entry of A by k :

$$kA = \begin{bmatrix} ka_{11} & ka_{12} & \dots & ka_{1n} \\ ka_{21} & ka_{22} & \dots & ka_{2n} \\ \dots & \dots & \dots & \dots \\ ka_{m1} & ka_{m2} & \dots & ka_{mn} \end{bmatrix}$$

We also define $-A = -1.\ A$ and $A-B = A+(-B)$.

Note 1 :

A+B and kA are also m×n matrices.

Note 2 :

The sum of matrices with different sizes is not defined.

3.3. MATRIX MULTIPLICATION

Definition

Let $A = (a_{ij})$ be an m×n matrix and $B=(b_{ij})$ be an n×p matrix. We define the product AB as the m×p matrix (C_{ij}) where the ij^{th} entry C_{ij} is given by

$$C_{ij} = a_{i1}b_{1j}+a_{i2}b_{2j}+.....+a_{in}b_{nj} = \sum_{k=1}^{n} a_{ik}b_{kj}$$

Note 1 :

The product AB of two matrices is defined only when the number of columns of A is equal to the number of rows of B.

Note 2 :

The entry C_{ij} of the product AB is found by multiplying i^{th} row of A and the j^{th} column of B. To multiply a row and a column. We multiply the corresponding entries and add.

Examples :

1. Let $A = \begin{bmatrix} 1 & -2 & 4 \\ -3 & 0 & 2 \\ 7 & 4 & 3 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 3 & -3 \\ 0 & 0 & 1 \end{bmatrix}$. Find AB.

$$AB = \begin{bmatrix} 1 & -2 & 4 \\ -3 & 0 & 2 \\ 7 & 4 & 3 \end{bmatrix}\begin{bmatrix} 0 & 1 & 0 \\ -1 & 3 & -3 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & -5 & 10 \\ 0 & -3 & 2 \\ -4 & 19 & -9 \end{bmatrix}$$

2. Let $A = \begin{bmatrix} 1 & 0 & 2 & 3 \\ 0 & 2 & 1 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 1 \\ 1 & 5 \\ 3 & 2 \\ 1 & 0 \end{bmatrix}$

A is a 3×4 matrix and B is a 4×2 matrix. Hence the product AB is a 3×2 matrix and

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

$$= \begin{bmatrix} 10 & 5 \\ 6 & 12 \\ 2 & 1 \end{bmatrix}$$

Note that in this example the product BA is not defined. Even if the product BA is defined, AB need not be equal to BA.

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

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

Then AI = IA = A

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

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

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

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

$\therefore$ **AI = IA = A**

4. Consider the square matrix of order n given by

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

Let A be any m×n matrix. Then $I_nA = A$.

Also if A is an m×n matrix, $AI_n = A$

If A is any n×n matrix, $AI_n = I_nA = A$

In is called the identity matrix of order n.

We shall denote the identity matrix of any order by the symbol I.

Solved Problems :

1. If $A = \begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 2 \end{bmatrix}$ show that $A^2 - 4A - 5I = 0$

Solution :

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

$$= \begin{bmatrix} 9 & 8 & 10 \\ 8 & 9 & 10 \\ 10 & 10 & 12 \end{bmatrix}$$

$$4A = 4\begin{bmatrix}1 & 2 & 2\\ 2 & 1 & 2\\ 2 & 2 & 2\end{bmatrix} = \begin{bmatrix}4 & 8 & 8\\ 8 & 4 & 8\\ 8 & 8 & 8\end{bmatrix}$$

$$5I = 5\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix} = \begin{bmatrix}5 & 0 & 0\\ 0 & 5 & 0\\ 0 & 0 & 5\end{bmatrix}$$

$$A^2-4A-5I = \begin{bmatrix}9 & 8 & 10\\ 8 & 9 & 10\\ 10 & 10 & 12\end{bmatrix} - \begin{bmatrix}4 & 8 & 8\\ 8 & 4 & 8\\ 8 & 8 & 8\end{bmatrix} - \begin{bmatrix}5 & 0 & 0\\ 0 & 5 & 0\\ 0 & 0 & 5\end{bmatrix}$$

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

$\therefore A^2-4A-5I = 0$

2. Show that the matrix $A = \begin{bmatrix}2 & -3 & 1\\ 3 & 1 & 3\\ -5 & 2 & -4\end{bmatrix}$ satisfies the equation $A(A-I)(A+2I)=0$.

Solution :

$$A-I = \begin{bmatrix}2 & -3 & 1\\ 3 & 1 & 3\\ -5 & 2 & -4\end{bmatrix} - \begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}$$

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

$$A+2I = \begin{bmatrix}2 & -3 & 1\\ 3 & 1 & 3\\ -5 & 2 & -4\end{bmatrix} + \begin{bmatrix}2 & 0 & 0\\ 0 & 2 & 0\\ 0 & 0 & 2\end{bmatrix} = \begin{bmatrix}4 & -3 & 1\\ 3 & 3 & 3\\ -5 & 2 & -2\end{bmatrix}$$

Now,

$$A(A-I)(A+2I) = \begin{bmatrix} 2 & -3 & 1 \\ 3 & 1 & 3 \\ -5 & 2 & -4 \end{bmatrix}\begin{bmatrix} 1 & -3 & 1 \\ 3 & 0 & 3 \\ -5 & 2 & -5 \end{bmatrix}\begin{bmatrix} 4 & -3 & 1 \\ 3 & 3 & 3 \\ -5 & 2 & -2 \end{bmatrix}$$

$$= \begin{bmatrix} -12 & -4 & -12 \\ -9 & -3 & -9 \\ 21 & 7 & 21 \end{bmatrix}\begin{bmatrix} 4 & -3 & 1 \\ 3 & 3 & 3 \\ -5 & 2 & -2 \end{bmatrix}$$

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

Hence $A(A-I)(A+2I) = 0$.

3. Prove that $\begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}^n = \begin{bmatrix} \lambda^n & n\lambda^{n-1} \\ 0 & \lambda^n \end{bmatrix}$

Solution :

We prove this result by induction on n. When n = 1 result is obviously true.

Let us assume that the result is true for n = k.

$$\therefore \quad \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}^k = \begin{bmatrix} \lambda^k & k\lambda^{k-1} \\ 0 & \lambda^k \end{bmatrix}$$

$$\therefore \quad \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}^k\begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix} = \begin{bmatrix} \lambda^k & k\lambda^{k-1} \\ 0 & \lambda^k \end{bmatrix}\begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}$$

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

$$= \begin{bmatrix} \lambda^{k+1} & (k+1)\lambda^k \\ 0 & \lambda^{k+1} \end{bmatrix}$$

$\therefore$ The result is true for n = k+1

Hence the result is true for all positive integers n.

Exercises :

1. Write down six pairs of matrices A and B such that the product AB is defined and in each case compute the product AB and also define & compute BA.

2. Show that if A is an m×n matrix, then AB and BA are both defined iff B is an n×m matrix.

3. If A and B are two matrices such that AB and A+B are both defined, show that A, B are square matrices of the same order.

4. Let $A = \begin{bmatrix} 1 & 5 & 3 \\ 0 & 2 & -1 \\ 1 & 0 & -2 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & 1 & 2 \\ 0 & 2 & 3 \\ 1 & 3 & 1 \end{bmatrix}$ compute A, B^2, AB and BA.

5. If $A = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 2 & 1 \\ 2 & 0 & 3 \end{bmatrix}$ prove that $A^3-6A^2+7A+2I = 0$

6. Prove that if $A = \begin{bmatrix} 3 & -4 \\ 1 & -1 \end{bmatrix}$, then $A^k = \begin{bmatrix} 1+2k & -4k \\ k & 1-2k \end{bmatrix}$ for any positive integer k.

Theorem 3.1 :

Let A be an m×n matrix, B an n×p matrix and C be an p×q matrix. Then A(BC) = (AB)C.

Proof :

Let $A = (a_{ij})$, $B = (b_{ij})$ and $C = (C_{ij})$. Let us find the rs^{th} entry in A(BC).

The r^{th} row in A is $a_{r1}, a_{r2}, \ldots\ldots, a_{rm}$.

The s^{th} column in BC consists of the elements $\Sigma b_{1j}c_{js}, \ldots\ldots, \Sigma b_{nj}c_{js}$. Hence the rs^{th} entry in A(BC) is $a_{r1}\Sigma b_{1j}c_{js} + \ldots\ldots + a_{rn}\Sigma b_{nj}c_{js}$

$$= \sum_{i=1}^{n} a_{ri} \sum_{j=1}^{p} b_{ij}c_{js}$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{p} a_{ri}b_{ij}c_{js}$$

Let us now find the rs^{th} entry in (AB)C.

The r^{th} row in AB is

$\Sigma a_{ri}b_{i1}$, $\Sigma a_{ri}b_{i2}$,.....$\Sigma a_{ri}b_{ip}$

The s^{th} column in C is c_{1s}, c_{2s},.....,c_{ps}

Hence the rs^{th} entry in (AB)C is

$$(\Sigma a_{ri}b_{i1})c_{1s}+(\Sigma a_{ri}b_{i2})c_{2s}+.....+(\Sigma a_{ri}b_{ip})c_{ps} = \sum_{i=1}^{n}\sum_{j=1}^{p} a_{ri}b_{ij}c_{js}$$

Thus A(BC) = (AB)C.

Theorem 3.2 :

Let U,V,W be vector spaces of dimensions m,n&p respectively over a field F with respective bases $\{u_1,u_2,....,u_m\}$, $\{v_1,v_2,....,v_n\}$ and $\{w_1,w_2,....,w_p\}$.

Let $T_1:U\to V$ and $T_2:V\to W$ linear transformations and $M(T_1)$ and $M(T_2)$ their corresponding matrices with respect to these bases.

Then $M(T_2 \bullet T_1) = M(T_1)M(T_2)$

Proof :

$M(T_1)$ and is an $m\times n$ matrix and $M(T_2)$ is an $n\times p$ matrix. Hence the product $M(T_1)M(T_2)$ is defined and is an $m\times p$ matrix.

Let $M(T_1) = (a_{ij})$

and $M(T_2) = (b_{ij})$

Then, $T_1(u_i) = \sum_{j=1}^{n} a_{ij}v_j$

and $T_2(v_j) = \sum_{k=1}^{p} b_{jk}w_k$

$\therefore$ $(T_2 \bullet T_1)(u_i) = T_2\left(\sum_{j=1}^{n} a_{ij}v_j\right)$

$= \sum_{j=1}^{n} a_{ij}T_2(v_j)$

$$= \sum_{j=1}^{n} a_{ij} \sum_{k=1}^{p} b_{jk} w_k$$

$$= \sum_{j=1}^{n} \sum_{k=1}^{p} (a_{ij} b_{jk})(w_k)$$

Thus $M(T_2 \cdot T_1) = M(T_1)M(T_2)$

Note 1 :

Thus multiplication of two matrices is equivalent to the composition of their corresponding linear transformations in the reverse order. Since composition of linear transformation is associative we get matrix multiplication is associative.

Note 2 :

Let $M_n(F)$ denote the set of all square matrices of order n over the field F. Then matrix multiplication is an associative binary operation on $M_n(F)$. If A, B, $C \in M_n(F)$ the two distributive laws.

$A(B+C) = AB+AC$ and $(A+B)C = AC+BC$ can be verified.

Since $M_n(F)$ is already an abelian group under matrix addition we see that $M_n(F)$ is a ring.

Solved Problems :

1. Find for what values of x will

$$(x \quad 4 \quad 1)\begin{bmatrix} 2 & 1 & 0 \\ 1 & 0 & 2 \\ 0 & 2 & 4 \end{bmatrix}\begin{bmatrix} x \\ 4 \\ 1 \end{bmatrix} = 0$$

Solution :

$$(x \quad 4 \quad 1)\begin{bmatrix} 2 & 1 & 0 \\ 1 & 0 & 2 \\ 0 & 2 & 4 \end{bmatrix}\begin{bmatrix} x \\ 4 \\ 1 \end{bmatrix} = 0$$

$$(x \quad 4 \quad 1)\begin{bmatrix} 2x+4 \\ x+2 \\ 12 \end{bmatrix} = 0$$

$$2x^2+4x+4x+8+12 = 0$$

$$2x^2+8x+20 = 0$$

$$x^2+4x+10 = 0$$

$$x = \frac{-4\pm\sqrt{16-40}}{2}$$

$$= \frac{-4\pm\sqrt{-24}}{2} = \frac{-4\pm i2\sqrt{6}}{2} = -2\pm i\sqrt{6}$$

$\therefore$ $x = -2\pm i\sqrt{6}$

2. If the determinant $\begin{bmatrix} m & 1 & 2 \\ -1 & 0 & 3 \\ 5 & -1 & 4 \end{bmatrix}$ of this matrix is 7. Then Find the value of m.

Solution :

$$\begin{vmatrix} m & 1 & 2 \\ -1 & 0 & 3 \\ 5 & -1 & 4 \end{vmatrix} = 7$$

$$m(3)-1(-4-15)+2(1) = 7$$

$$3m+19+2 = 7$$

$$3m = 7-21$$

$$3m = -14$$

$\therefore$ $m = \frac{-14}{3}$

Exercises :

1. Using $A = \begin{bmatrix} 1 & -1 & 1 \\ 5 & 0 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 1 & 1 & 1 \end{bmatrix}$, $C = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}$ test the associative law A(BC) = (AB)C for matrix multiplication.

2. Compute $\begin{pmatrix} 2 & 1 & 3 \end{pmatrix}\begin{bmatrix} 2 & 1 & 3 \\ 4 & -1 & 2 \\ 0 & 1 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}$

3. Find for what values of x will $\begin{pmatrix}3 & x & 2\end{pmatrix}\begin{bmatrix}1 & 2 & 5\\ 0 & 1 & -1\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}1\\ x\\ 2\end{bmatrix} = 0$

4. Given that $\begin{bmatrix}1 & 0 & 0\\ 0 & 2 & 0\\ 0 & 0 & -3\end{bmatrix} A \begin{bmatrix}1 & 0 & 0\\ 0 & 0 & 1\\ 0 & 1 & 0\end{bmatrix} = \begin{bmatrix}1 & 2 & 3\\ 4 & 5 & 4\\ 3 & 2 & 1\end{bmatrix}$ find the matrix A.

3.4. THE TRANSPOSE OF A MATRIX

Definition :

Let $A = (a_{ij})$ be an m×n matrix. Then the n×m matrix $B = (b_{ij})$ where $b_{ij} = a_{ji}$ is called the **transpose** of the matrix A and it is denoted by A^T. Thus A^T is obtained from the matrix A by interchanging its rows and columns and the $(i, j)^{th}$ entry of $A^T = (j,i)^{th}$ entry of A.

For example,

If $A = \begin{bmatrix}1 & 3 & 4 & 5\\ 2 & 3 & 1 & 2\\ 0 & 1 & 2 & 0\end{bmatrix}$ then $A^T = \begin{bmatrix}1 & 2 & 0\\ 3 & 3 & 1\\ 4 & 1 & 2\\ 5 & 2 & 0\end{bmatrix}$

Clearly if A is an m×n matrix, then A^T is an n×m matrix.

Theorem 3.3 :

Let A and B be two m×n matrices. Then (i) $(A^T)^T = A$ (ii) $(A+B)^T = A^T+B^T$.

Proof :

(i) The $(i, j)^{th}$ entry of $(A^T)^T$ = $(j, i)^{th}$ entry of A^T

= $(i, j)^{th}$ entry of A

$\therefore\ (A^T)^T = A$

(ii) The $(i,j)^{th}$ entry of $(A+B)^T$ = $(j, i)^{th}$ entry of A+B

= $(j, i)^{th}$ entry of A+$(j, i)^{th}$ entry of B

= $(i, j)^{th}$ entry of A^T +$(i, j)^{th}$ entry of B^T

= $(i, j)^{th}$ entry of (A^T+B^T)

$\therefore\ (A+B)^T = A^T+B^T$.

Theorem 3.4 :

Let A be an m×n matrix and B be an n×p matrix. Then $(AB)^T = B^TA^T$.

Proof :

By hypothesis AB is defined and it is an m×p matrix. Hence $(AB)^T$ is a p×m matrix.

Further B^T is a p×n matrix and A^T is an n×m matrix.

Hence, the product B^TA^T is defined and it is a p×m matrix.

Now, Let $A = (a_{ij})$, $B = (b_{ij})$ and $AB = (c_{ij})$.

The $(i, j)^{th}$ entry of $AB = c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}$

∴ The $(i, j)^{th}$ entry of $(AB)^T = c_{ji} = \sum_{k=1}^{n} a_{jk}b_{ki}$

Now the i^{th} row of B^T is the i^{th} column of B and it consists of the elements $b_{1i}, b_{2i}, \ldots\ldots, b_{ni}$. Also the j^{th} column of A^T is the j^{th} row of A and it consists of the elements $a_{j1}, a_{j2}, \ldots\ldots, a_{jn}$.

Hence the $(i, j)^{th}$ entry of

$$\begin{aligned} B^TA^T &= b_{1i}a_{j1} + b_{2i}a_{j2} + \ldots\ldots + b_{ni}a_{jn} \\ &= \sum_{k=1}^{n} b_{ki}a_{jk} \\ &= (i, j)^{th} \text{ entry of } (AB)^T. \end{aligned}$$

Hence $(AB)^T = B^TA^T$.

Definition :

Let $A = (a_{ij})$ be a matrix with entries from the field of complex numbers. The **conjugate** of A, denoted by $\overline{A}$, is defined by $\overline{A} = \left(\overline{a_{ij}}\right)$.

$\overline{A}^T$ is called the **conjugate transpose** of the matrix A.

For example,

$$\text{If } A = \begin{bmatrix} 3 & 2+i & 1+i \\ 4-i & -2 & -1+2i \end{bmatrix} \text{ then } \overline{A} = \begin{bmatrix} 3 & 2-i & 1-i \\ 4+i & -2 & -1-2i \end{bmatrix}$$

Solved Problem :

1. Let $A = \begin{bmatrix} 2i & 3+4i & 0 \\ 1+i & 1-i & i \\ 3 & 2i & 4 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & 2 & 6 \\ -1 & 4 & 6 \\ 2 & 0 & 2 \end{bmatrix}$

Find $A^T, B^T, \overline{A}^T, \overline{A^T}, \overline{B}^T, \overline{B^T}, (A+B)^T, A^T+B^T, (AB)^T, B^TA^T$.

Solution :

$$A^T = \begin{bmatrix} 2i & 1+i & 3 \\ 3+4i & 1-i & 2i \\ 0 & i & 4 \end{bmatrix} \qquad B^T = \begin{bmatrix} 0 & -1 & 2 \\ 2 & 4 & 0 \\ 6 & 6 & 2 \end{bmatrix}$$

$$\overline{A^T} = \begin{bmatrix} -2i & 1-i & 3 \\ 3-4i & 1+i & -2i \\ 0 & -i & 4 \end{bmatrix}, \qquad \overline{B^T} = \begin{bmatrix} 0 & -1 & 2 \\ 2 & 4 & 0 \\ 6 & 6 & 2 \end{bmatrix}$$

$$\overline{A}^T = \begin{bmatrix} -2i & 1-i & 3 \\ 3-4i & 1+i & -2i \\ 0 & -i & 4 \end{bmatrix}, \qquad \overline{B}^T = \begin{bmatrix} 0 & -1 & 2 \\ 2 & 4 & 0 \\ 6 & 6 & 2 \end{bmatrix}$$

$$(A+B)^T = \begin{bmatrix} 2i & 5+4i & 6 \\ i & 5-i & 6+i \\ 5 & 2i & 6 \end{bmatrix} = \begin{bmatrix} 2i & i & 5 \\ 5+4i & 5-i & 2i \\ 6 & 6+i & 6 \end{bmatrix}$$

$$A^T+B^T = \begin{bmatrix} 2i & i & 5 \\ 5+4i & 5-i & 2i \\ 6 & 6+i & 6 \end{bmatrix}$$

$$(AB)^T = \begin{bmatrix} -3-4i & 12+20i & 18+36i \\ -1+3i & 6-2i & 12+2i \\ 8-2i & 6+8i & 26+12i \end{bmatrix}^T = \begin{bmatrix} -3-4i & -1+3i & 8-2i \\ 12+20i & 6-2i & 6+8i \\ 18+36i & 12+2i & 26+12i \end{bmatrix}$$

$$B^TA^T = \begin{bmatrix} -3-4i & -1+3i & 8-2i \\ 12+20i & 6-2i & 6+8i \\ 18+36i & 12+2i & 26+12i \end{bmatrix}$$

Note : $(AB)^T = B^TA^T$, $(A+B)^T = A^T+B^T$.

Theorem 3.5 :

Let A and B be matrices with entries from C. Then

(i) $\overline{(\overline{A})} = A$

(ii) $\overline{A+B} = \overline{A}+\overline{B}$

(iii) $\overline{kA} = \overline{k}\,\overline{A}$, where $k \in C$

(iv) $A = \overline{A} \Leftrightarrow$ all entries of A are real

(v) $\overline{AB} = \overline{A}\,\overline{B}$ provided AB is defined

(vi) $(\overline{A})^T = \overline{A^T}$

The proof of the above results are immediate consequences of the corresponding properties of complex numbers.

Exercises :

1. Let $A = \begin{bmatrix} 2 & 3 & 5 \\ 1 & 2 & 0 \\ 1 & 3 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 2 & 3 \\ 4 & 2 & 3 \end{bmatrix}$

Find A^T, B^T, $(A+B)^T$ $(AB)^T$ and B^TA^T.

2. Let $A = \begin{bmatrix} 3i & 2+i & 1 \\ 1-i & 1+2i & 0 \\ 4 & 2i & 3 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 3 & 2 \\ 2 & 5 & 2 \end{bmatrix}$

Find $\overline{A}$, $\overline{A+B}$, $\overline{AB}$, $\overline{A}\,\overline{B}$, $\overline{A}^T$, $\overline{B}^T$, $\overline{A}^TB$ and $A\overline{B}^T$.

3.5. THE INVERSE OF A MATRIX

A 2×2 matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ has an inverse iff $|A| = ad-bc \neq 0$ and the inverse of A is given by $\frac{1}{|A|}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$. Such matrices are called non-singular. In this section we shall describe the method of finding the inverse of any non-singular matrix of order n.

Determinants :

The determinant of a squase matrix A, denoted by det A or |A|, is a scalar. If the matrix is written out as an array of elements, then its determinant is indicated by replacing the brackets with vertical lines.

For 1×1 matrices,

det A = $|a_{11}| = a_{11}$.

For 2×2 matrices,

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

Determinants for n×n matrices with n>2 are calculated through a process of reduction on and expansion utilizing minors and co-factors, as follows.

Definition :

Let A = (a_{ij}) be an n×n matrix. If we delete the row and the column containing the element a_{ij} we obtain a square matrix of order n–1 and the determinant of this square matrix is called the **minor** of the element a_{ij} and is denoted by M_{ij}.

The minor M_{ij} multiplied by $(-1)^{i+j}$ is called the **cofactor** of the element a_{ij} and is denoted by A_{ij}.

$\therefore A_{ij} = (-1)^{i+j} M_{ij}$

Example :

$$\text{Let } A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

Corresponding to the 9 elements a_{ij}, we get 9 minors of A.

For example, the minor of a_{11} is $M_{11} = \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}$ and the minor of a_{23} is $M_{23} = \begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix}$

The cofactor of a_{11} is $A_{11} = (-1)^2 M_{11} = M_{11}$

The cofactor of a_{23} is $A_{23} = (-1)^{2+3} M_{23} = -M_{23}$

If $A = \begin{bmatrix} 0 & 1 & 2 \\ 3 & 4 & 5 \\ 6 & 7 & 8 \end{bmatrix}$ then $|A| = ?$

$$|A| = 0\begin{vmatrix} 4 & 5 \\ 7 & 8 \end{vmatrix} - 1\begin{vmatrix} 3 & 5 \\ 6 & 8 \end{vmatrix} + 2\begin{vmatrix} 3 & 4 \\ 6 & 7 \end{vmatrix}$$

$$= 0-1(24-30)+2(21-24) = 6-6 = 0$$

Definition :

A square matrix A is said to be **singular** if $|A| = 0$.

A is called a **non-singular** matrix if $|A| \neq 0$.

Remark :

The rule for multiplying two matrices is same as the rule for multiplying two determinants.

Hence it A and B are two $n \times n$ matrices $|AB| = |A|\,|B|$

Theorem 3.6 :

The product of any two non-singular matrices is non-singular.

Proof :

Let A and B be two non-singular matrices of the same order. Then $|A| \neq 0$ and $|B| \neq 0$.

$\therefore |AB| = |A|\,|B| \neq 0$

Hence AB is non-singular.

Note : Sum of two non-singular matrices need not be non-singular. For, if A is any non-singular matrix then $-A$ is also a non-singular matrix and $A+(-A)$ is the zero matrix which is obviously a singular matrix.

Definition :

Let $A = (a_{ij})$ be a square matrix. Let A_{ij} denote the co-factor of a_{ij}. The transpose of the matrix (A_{ij}) is called the **adjoint** or **adjugate** of the matrix A and is denoted by adj A.

Thus the $(i, j)^{th}$ entry of adj A is A_{ji}.

Note : If A is a square matrix of order n then adj A is also a square matrix of order n.

Example :

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

Then $$A_{11} = \begin{vmatrix} 1 & -1 \\ 1 & 3 \end{vmatrix} = 4$$

$$A_{12} = -\begin{vmatrix} 3 & -1 \\ -2 & 3 \end{vmatrix} = -7$$

Similarly other co-factors can be calculated and we get

$$\text{adj } A = \begin{bmatrix} A_{11} & A_{21} & A_{31} \\ A_{12} & A_{22} & A_{32} \\ A_{13} & A_{23} & A_{33} \end{bmatrix} = \begin{bmatrix} 4 & 2 & -2 \\ -7 & 7 & 7 \\ 5 & -1 & 1 \end{bmatrix}$$

$$A(\text{adj } A) = \begin{bmatrix} 1 & 0 & 2 \\ 3 & 1 & -1 \\ -2 & 1 & 3 \end{bmatrix}\begin{bmatrix} 4 & 2 & -2 \\ -7 & 7 & 7 \\ 5 & -1 & 1 \end{bmatrix} = \begin{bmatrix} 14 & 0 & 0 \\ 0 & 14 & 0 \\ 0 & 0 & 14 \end{bmatrix}$$

$$(\text{adj } A)A = \begin{bmatrix} 4 & 2 & -2 \\ -7 & 7 & 7 \\ 5 & -1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 2 \\ 3 & 1 & -1 \\ -2 & 1 & 3 \end{bmatrix} = \begin{bmatrix} 14 & 0 & 0 \\ 0 & 14 & 0 \\ 0 & 0 & 14 \end{bmatrix}$$

$\therefore$ $A(\text{adj } A) = (\text{adj } A)A$

Theorem 3.7 :

Let A be any square matrix of order n.

Then (adj A)A = A(adj A) = |A|I where I is the identity matrix of order n.

Proof :

The $(i, j)^{th}$ element of (A (adj A))

$$= \sum_{k=1}^{n} a_{ik}A_{jk}$$

$$= \begin{cases} 0 \text{ if } i \neq j \\ |A| \text{ if } i = j \end{cases}$$

$$\therefore \quad A(\text{adj } A) = \begin{bmatrix} |A| & 0 & \dots & 0 \\ 0 & |A| & \dots & 0 \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & |A| \end{bmatrix}$$

$$= |A| I$$

Similarly, (adj A)A = |A| I

Hence (adj A)A = A(adj A) = |A| I.

Note :

Suppose $|A| \neq 0$. Now, consider the matrix $B = \frac{1}{|A|}$ adj A.

Then $$AB = A\left[\frac{1}{|A|}(\text{adj } A)\right]$$

$$= \frac{1}{|A|}(A \text{ adj } A) = \frac{1}{|A|}|A|I = I$$

Similarly $$BA = I$$

Thus AB = BA = I.

Definition :

Let A be a square matrix of order n. A is said to be **invertible** in there exists a square matrix B of order n such that AB = BA = I and B is called the **inverse** of A and is denoted by A^{-1}.

Note :

The invertible matrices are precisely the units of the ring $M_n(F)$.

Theorem 3.8 :

A square matrix A of order n is non-singular iff A is invertible.

Proof :

Suppose A is invertible.

Then there exists a matrix B such that AB = BA = I.

Hence $|AB| = |I| = 1$

$\therefore |A|\,|B| = 1$

Hence $|A| \neq 0$ so that A is non-singular.

Conversely, let A be non-singular. Hence $|A| \neq 0$.

Now, Consider the matrix $B = \frac{1}{|A|}\text{adj}\,A$.

Then $AB = BA = I$

$\therefore$ A is invertible and B is the inverse of A.

Solved Problems :

1. Compute the inverse of the matrix $A = \begin{bmatrix} 2 & -1 & 1 \\ -15 & 6 & -5 \\ 5 & -2 & 2 \end{bmatrix}$

Solution :

$$|A| = \begin{vmatrix} 2 & -1 & 1 \\ -15 & 6 & -5 \\ 5 & -2 & 2 \end{vmatrix} = -1$$

Since $|A| \neq 0$, A is non-singular.

Hence A^{-1} exists and is given by $A^{-1} = \frac{\text{adj}\,A}{|A|}$

Now, we find adj $A = \begin{bmatrix} A_{11} & A_{21} & A_{31} \\ A_{12} & A_{22} & A_{32} \\ A_{13} & A_{23} & A_{33} \end{bmatrix}$

Where A_{ij}, (i, j = 1,2,3) are cofactors of a_{ij}.

$$A_{11} = \begin{vmatrix} 6 & -5 \\ -2 & 2 \end{vmatrix} = 2; \qquad A_{12} = -\begin{vmatrix} -15 & -5 \\ 5 & 2 \end{vmatrix} = 5$$

$$A_{13} = \begin{vmatrix} -15 & 6 \\ 5 & -2 \end{vmatrix} = 0; \qquad A_{21} = -\begin{vmatrix} -1 & 1 \\ -2 & 2 \end{vmatrix} = 0$$

$$A_{22} = \begin{vmatrix} 2 & 1 \\ 5 & 2 \end{vmatrix} = -1; \qquad A_{23} = -\begin{vmatrix} 2 & -1 \\ 5 & -2 \end{vmatrix} = -1$$

$$A_{31} = \begin{vmatrix} -1 & 1 \\ 6 & -5 \end{vmatrix} = -1; \qquad A_{32} = -\begin{vmatrix} 2 & 1 \\ -15 & -5 \end{vmatrix} = -5$$

$$A_{33} = \begin{vmatrix} 2 & -1 \\ -15 & 6 \end{vmatrix} = -3$$

Hence $$\text{adj } A = \begin{bmatrix} 2 & 0 & -1 \\ 5 & -1 & -5 \\ 0 & -1 & -3 \end{bmatrix}$$

$\therefore$ $$A^{-1} = \frac{1}{-1}\begin{bmatrix} 2 & 0 & -1 \\ 5 & -1 & -5 \\ 0 & -1 & -3 \end{bmatrix} = \begin{bmatrix} -2 & 0 & 1 \\ -5 & 1 & 5 \\ 0 & 1 & 3 \end{bmatrix}$$

2) If $w = e^{2\pi i/3}$ find the inverse the matrix $A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & \omega & \omega^2 \\ 1 & \omega^2 & \omega \end{bmatrix}$

Solution :

We note that $\omega^3 = 1$

$$|A| = \begin{bmatrix} 1 & 1 & 1 \\ 1 & \omega & \omega^2 \\ 1 & \omega^2 & \omega \end{bmatrix} = 1(\omega^2-\omega^4)-1(\omega-\omega^2)+1(\omega^2-\omega)$$

$$= \omega^2-\omega-\omega+\omega^2+\omega^2-\omega$$

$$= 3\omega^2-3\omega = 3(\omega^2-\omega)$$

Since $|A| \neq 0$, A is non-singular. Hence A^{-1} exists and is given by $A^{-1} = \frac{\text{adj } A}{|A|}$

Now, $$\text{adj } A = \begin{bmatrix} \omega^2-\omega & \omega^2-\omega & \omega^2-\omega \\ \omega^2-\omega & \omega-1 & 1-\omega^2 \\ \omega^2-\omega & 1-\omega^2 & \omega-1 \end{bmatrix}$$

$$\therefore \quad A^{-1} = \frac{1}{3(\omega^2-\omega)}\begin{bmatrix} \omega^2-\omega & \omega^2-\omega & \omega^2-\omega \\ \omega^2-\omega & \omega-1 & 1-\omega^2 \\ \omega^2-\omega & 1-\omega^2 & \omega-1 \end{bmatrix}$$

$$= \frac{1}{3\omega}\begin{bmatrix} \omega & \omega & \omega \\ \omega & 1 & -1-\omega \\ \omega & -1-\omega & 1 \end{bmatrix}$$

3) Show that a square matrix A is **orthogonal** iff $A^{-1} = A^T$.

Solution :

Suppose A is orthogonal . Then $AA^T = I$.

$\therefore \quad |AA^T| = |I| = 1$

$\therefore \quad |A|\,|A^T| = 1$

$\therefore \quad |A|\,|A| = 1$

$\therefore \quad |A| \neq 0$ and hence A is non-singular.

$\therefore A^{-1}$ exists.

Now, $A^{-1}(AA^T) = A^{-1}I$

$\therefore \quad (A^{-1}A)A^T = A^{-1}$

$\therefore \quad IA^T = A^{-1}$

$\therefore \quad A^T = A^{-1}$

Conversely, let $A^T = A^{-1}$

Then $AA^T = AA^{-1} = I$

Similarly, $A^TA = I$

Hence A is orthogonal.

4) Show that a square matrix A is involutory iff $A = A^{-1}$

Solution :

Suppose A is involutory. Then $A^2 = I$

Hence $|A^2| = 1$

$\therefore \quad |A^2| = |A|\,|A| = 1$

$\therefore$ $|A| \neq 0$ and hence A is non-singular.

$\therefore$ A^{-1} exists.

Now, $A^{-1}(AA) = A^{-1}I$

$(A^{-1}A)A = A^{-1}$

$IA = A^{-1}$

$A = A^{-1}$

Conversely, let $A = A^{-1}$

Then $A^2 = AA - AA^{-1} = I$

$\therefore$ A is involutory.

Theorem 3.9 :

Let V and W be vector spaces of dimension n over a field F with bases $v_1, v_2, \ldots v_n$ and $w_1, w_2, \ldots w_n$ respectively. Then a linear transformation $T:V \to W$ is non-singular iff the associated matrix is non-singular.

Proof :

Let $T:V \to W$ be a non-singular linear transformation.

Then T is 1–1 and onto.

Hence $T^{-1}:W \to V$ is also a linear transformation.

Let A and B be the matrices representing the linear tranformations T and T^{-1} with respect to the chosen bases.

By theorem 3.2 :

Multiplication of the matrices A and B is equivalent to the composition of the corresponding linear transformation T and T^{-1}.

Also $T \bullet T^{-1}$ and $T^{-1} \bullet T$ are identity transformations.

Hence $AB = BA = I$. Thus A has an inverse B. Hence A is non-singular.

Conversely, let A be a non-singular matrix. Then A^{-1} exists.

Let $S:W \to V$ be the linear transformation determined by the matrix A^{-1}.

It is easily verified that $T \bullet S = S \bullet T = I$

Hence T has an inverse linear transformation S.

Hence T is a non-singular linear transformation.

Exercises :

1. Compute the inverse of each of the following matrices.

(a) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$ (b) $\begin{bmatrix} 8 & -1 & -3 \\ -5 & 1 & 2 \\ 10 & -1 & -4 \end{bmatrix}$

(c) $\begin{bmatrix} 2 & 1 & -1 \\ 0 & 2 & 1 \\ 5 & 2 & -3 \end{bmatrix}$ (d) $\begin{bmatrix} 1 & 2 & 3 \\ 0 & -1 & 4 \\ -2 & 2 & 1 \end{bmatrix}$

(e) $\begin{bmatrix} 2 & 2 & -3 \\ -3 & 2 & 2 \\ 2 & -3 & 2 \end{bmatrix}$ (f) $\begin{bmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix}$

2. Show that the set of all non-singular matrices of order n over a field F is a group under matrix multiplication.

3. If A and B are non-singular matrices of order n prove that $(AB)^{-1} = B^{-1}A^{-1}$.

4. If A is a non-singular symmetric matrix prove that A^{-1} is also a symmetric matrix.

5. If A is a non-singular matrix, prove that $(A^T)^{-1} = (A^{-1})^T$.

6. If A is orthogonal, prove that A^{-1} is orthogonal.

3.6. TYPES OF MATRICES

Definition :

An $1\times n$ matrix is called a **row matrix**. Thus a row matrix consists of 1 row and n columns. It is of the form $(a_{11}, a_{12}, a_{13},.....,a_{1n})$.

Definition :

An $m\times 1$ matrix is called a **column matrix**. Thus a column matrix consists of m rows and 1 column and it is of the form $\begin{bmatrix} a_{11} \\ a_{12} \\ \vdots \\ a_{m1} \end{bmatrix}$

Definition :

Let $A = (a_{ij})$ be a square matrix. Then the elements $a_{11}, a_{22},.....,a_{nn}$ are called the diagonal elements of A and the diagonal elements constitute what is known as the principal diagonal of the matrix A. A square matrix is called a **diagonal matrix** if all the entries which do not belong to the principal are zero. Hence in a diagonal matrix $a_{ij} = 0$ if $i \neq j$.

For example, $\begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 5 \end{bmatrix}$ is a diagonal matrix.

Definition :

A diagonal matrix in which all the entries of the principal diagonal are equal is called a **scalar matrix.**

For example, $\begin{bmatrix} 5 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 5 \end{bmatrix}$ is a scalar matrix.

Definition :

A square matrix (a_{ij}) is called an **upper triangular** matrix if all the entries above the principal diagonal are zero.

Hence $a_{ij} = 0$ whenever $i<j$ in an upper triangular matrix.

Definition :

A square matrix (a_{ij}) is called a **lower triangular** matrix if all the entries below the principal diagonal are zero.

Hence $a_{ij} = 0$ whenever $i>j$ in an lower triangular matrix.

For example, $\begin{bmatrix} 2 & 3 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 5 \end{bmatrix}$ is a lower triangular $\begin{bmatrix} 2 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 1 & 1 & 4 & 0 \\ 3 & 2 & 1 & 1 \end{bmatrix}$ is upper triangular.

Clearly a square matrix is a diagonal matrix iff it is both lower triangular and upper triangular.

Definition :

A square matrix $A = (a_{ij})$ is said to be **symmetric** if $a_{ij} = a_{ji}$ for all i, j.

For example,

$$\begin{bmatrix} a & b \\ b & a \end{bmatrix}, \begin{bmatrix} f & a & b \\ a & g & c \\ b & c & h \end{bmatrix}, \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 0 & 6 & 5 \\ 3 & 6 & 3 & 6 \\ 4 & 5 & 6 & 7 \end{bmatrix}$$ are symmetric matrices.

Theorem 3.10 :

A square matrix A is symmetric iff $A = A^T$.

Proof :

Let A be a symmetric matrix.

Then the $(i, j)^{th}$ entry of A.

$= (j, i)^{th}$ entry of A

$= (i, j)^{th}$ entry of A^T.

Hence $A = A^T$

Conversely, let $A = A^T$

Then $(i, j)^{th}$ entry of $A = (i, j)^{th}$ entry of A^T

$= (j, i)^{th}$ entryof A

Hence A is symmetric.

Theorem 3.11 :

Let A be any square matrix. Then $A+A^T$ is symmetric.

Proof :

$$(A+A^T)^T = A^T+(A^T)^T$$
$$= A^T+A$$
$$= A+A^T$$

Hence $A+A^T$ is symmetric.

Theorem 3.12 :

Let A and B be symmetric matrices of order n. Then

(i) A+B is symmetric.

(ii) AB is symmetric iff AB = BA

(iii) AB+BAis symmetric

(iv) If A is symmetric, then kA is symmetric where $k \in F$

Proof :

(i) $(A+B)^T = A^T+B^T = A+B$ (since A and B are symmetric)

∴ A+B is symmetric.

(ii) AB is symmetric.

$\Leftrightarrow \quad (AB)^T = AB$

$\Leftrightarrow \quad B^TA^T = AB$ (by theorem 3.4)

$\Leftrightarrow \quad BA = AB$

(iii) $(AB+BA)^T = (AB)^T+(BA)^T$

$= B^TA^T+A^TB^T$

$=$ BA+AB (since A and B are symmetric)

$=$ AB+BA

∴ AB+BA is symmetric.

(iv) $(kA)^T = kA^T = kA$ (since A is symmetric)

∴ kA is symmetric.

Definition :

A square matrix $A = (a_{ij})$ is said to be **skew symmetric** if $a_{ij} = -a_{ji}$, for all i, j.

Note : Let A be a skew symmetric matrix. Then $a_{ii} = -a_{ii}$.

Hence $2a_{ii} = 0$

(i.e.,) $a_{ii} = 0$ for all i.

Thus in a skew symmetric matrix all the diagonal entries are zero.

$\begin{bmatrix} 0 & -a \\ a & 0 \end{bmatrix}, \begin{bmatrix} 0 & -3 & -1 \\ 3 & 0 & 2 \\ 1 & -2 & 0 \end{bmatrix}$ are examples of skew symmetric matrices.

Theorem 3.13 :

A square matrix A is skew symmetric matrix iff $A = -A^T$.

Proof :

Let A be a skew symmetrix matrix.

Then the $(i, j)^{th}$ entry of A

$= -(j, i)^{th}$ entry of A

$= -(i, j)^{th}$ entry of A^T

Hence $A = -A^T$

Conversely, let $A = -A^T$

Then $(i, j)^{th}$ entry of A $= -(i, j)^{th}$ entry of A^T

$= -(j, i)^{th}$ entry of A

Hence A is skew symmetric.

Theorem 3.14 :

Let A be any square matrix. Then $A-A^T$ is skew symmetric..

Proof :

$$(A-A^T)^T = A^T-(A^T)^T$$

$$= A^T-A = -(A-A^T)$$

Hence $A-A^T$ is skew symmetric.

Theorem 3.15 :

Any square matrix A can be expressed uniquely as the sum of a symmetric matrix and a skew symmetrix matrix.

Proof :

Let A be any square matrix.

Then $A+A^T$ is a symmetric matrix (by theorem 3.11)

$\therefore \frac{1}{2}(A+A^T)$ is also a symmetric matrix.

Also $\frac{1}{2}(A-A^T)$ is a skew symmetric matrix. (by theorem 3.14)

Now, $A = \frac{1}{2}(A+A^T)+\frac{1}{2}(A-A^T)$

$\therefore$ A is the sum of a symmetric matrix and a skew symmetric matrix.

Now, to prove the uniqueness, let A = R+S where S is a symmetric matrix and R is a skew symmetric matrix.

We claim that $S = \frac{1}{2}(A+A^T)$

and $R = \frac{1}{2}(A-A^T)$

$A = S+R$ --------(1)

$A^T = (S+R)^T = S^T+R^T$

$= S-R$ (since S is symmetric and R is skew symmetric)

$\therefore A^T = S-R$ --------(2)

From (1) & (2) we get,

$S = \frac{1}{2}(A+A^T)$

and $R = \frac{1}{2}(A-A^T)$

Theorem 3.16 :

Let A and B be skew symmetric matrices of order n. Then

(i) A+B is skew symmetric.

(ii) kA is skew symmetric, where $k \in F$.

(iii) A^{2n} is a symmetric matrix and A^{2n+1} is a skew symmetric matrix where n is any positive integer.

Proof :

Let A, B be skew symmetric.

(i) $(A+B)^T = A^T+B^T$

$= -A-B$ (by theorem 3.13)

$= -(A+B)$

$\therefore$ A+B is skew symmetric.

(ii) $(kA)^T = kA^T = -kA$ (since A is skew symmetric)

$\therefore$ kA is skew symmetric.

(iii) Let m be any positive integer.

Then $(A^m)^T = (AA.....\text{m times})^T$

$= (-A)(-A).....(-A)$ (m times) (since $A^T = -A$)

$= (-1)^m A^m$

$\therefore$ $(A^m)^T = \begin{cases} A^m \text{ if m is even} \\ -A^m \text{ if m is odd} \end{cases}$

$\therefore$ A^m is symmetric when m is even and skew symmetric when m is odd.

Definition :

A square matrix $A = (a_{ij})$ is said to be a **Hermitian matrix** if $a_{ij} = \overline{a_{ji}}$ for all i, j. A is said to be a **skew Hermitian matrix** iff $a_{ij} = -\overline{a_{ji}}$ for all i, j.

Example :

$\begin{bmatrix} 2 & -2+2i & 3 \\ -2-2i & 1 & 1+i \\ 3 & 1-i & 0 \end{bmatrix}$ is a Hermitian matrix.

$\begin{bmatrix} 0 & -a+ib \\ a+ib & 0 \end{bmatrix}, \begin{bmatrix} ib & c+id \\ -c+id & ib \end{bmatrix}$ are skew Hermitian matrices.

Note :

1. Any Hermitian matrix over R is a symmetric matrix and any skew Hermitian matrix over R is a skew symmetric matrix.
2. Let $A = (a_{ij})$ be a Hermitian matrix. Then $a_{ii} = \overline{a_{ii}}$ and hence a_{ii} is real for all i.
3. Let $A = (a_{ij})$ be a skew Hermitian matrix. Then $a_{ii} = -\overline{a_{ii}}$ and hence $a_{ii} = 0$ or purely imaginary for all i.

Theorem 3.17 :

Let A be a square matrix.

(i) A is Hermitian iff $A = \overline{A}^T$

(ii) A is skew Hermitian iff $A = -\overline{A}^T$

Proof :

(i) Let A be a Hermitian matrix. Then the (i, j)th entry of

$$A = (j, i)^{th} \text{ entry of } \overline{A}$$

$$= (i, j)^{th} \text{ entry of } \overline{A}^T$$

$$\text{Hence } A = \overline{A}^T$$

Conversely, let $A = \overline{A}^T$

Then $(i, j)^{th}$ entry of $A = (i, j)^{th}$ entry of $\overline{A}^T$

$$= (j, i)^{th} \text{ entry of } \overline{A}$$

Hence A is Hermitian.

(ii) Let A be a skew Hermitian matrix.

Then the $(i,j)^{th}$ entry of $A = -(j, i)^{th}$ entry of $\overline{A}$

$$= -(i, j)^{th} \text{ entry of } \overline{A}^T$$

Hence $A = -\overline{A}^T$

Conversely, let $A = -\overline{A}^T$

Then $(i, j)^{th}$ entry of $A = -(i, j)^{th}$ entry of $\overline{A}^T$

$$= -(j, i)^{th} \text{ entry of } \overline{A}$$

Hence A is skew Hermitian.

Theorem 3.18 :

Let A and B be square matrices of the same order. Then

(i) A, B are Hermitian $\Rightarrow$ A+B is Hermitian.

(ii) A, B are skew Hermitian $\Rightarrow$ A+B is skew Hermitian.

(iii) A is Hermitian $\Rightarrow$ iA is skew Hermitian.

(iv) A is skew Hermitian $\Rightarrow$ iA is Hermitian

(v) A is Hermitian and k is real $\Rightarrow$ kA is Hermitian

(vi) A is skew Hermitian and k is real $\Rightarrow$ kA is skew Hermitian

(vii) A, B are Hermitian $\Rightarrow$ AB+BA is Hermitian

(viii) A, B are Hermitian $\Rightarrow$ AB–BA is skew Hermitian.

Proof :

(i) $\overline{(A+B)}^T = (\overline{A}+\overline{B})^T$

$= \overline{A}^T + \overline{B}^T$

$= A+B$ (since A and B are Hermitian)

$\therefore$ A+B is Hermitian.

(ii) $-\overline{(A+B)}^T = -(\overline{A}+\overline{B})^T$

$= -\overline{A}^T - \overline{B}^T$

$= A+B$ (since A and B are skew Hermitian)

$\therefore$ A+B is skew Hermitian.

(iii) $\overline{-(iA)}^T = (-i\overline{A})^T$

$= i\overline{A}^T = iA$ (since A is Hermitian)

$\therefore$ iA is skew Hermitian.

(iv) $(\overline{iA})^T = -i\overline{A}^T$

$= iA$ (since A is skew Hermitian)

$\therefore$ iA is Hermitian.

(v) $(\overline{kA})^T = k\overline{A}^T$

$= kA$ (since A is Hermitian) i.e., $A = \overline{A}^T$

$\therefore$ kA is Hermitian.

(vi) $(-\overline{kA})^T = -k\overline{A}^T$

$= kA$ (since A is skew Hermitian) i.e., $A = -\overline{A}^T$

$\therefore$ kA is skew Hermitian.

(vii) $(\overline{AB+BA})^T = (\overline{AB}+\overline{BA})^T$

$= (\overline{A}\,\overline{B}+\overline{B}\overline{A})^T$

$$= (\overline{A}\,\overline{B})^T + (\overline{B}\,\overline{A})^T$$
$$= \overline{B}^T\overline{A}^T + \overline{A}^T\overline{B}^T$$
$$= BA+AB$$
$$= AB+BA$$

$\therefore$ AB+BA is Hermitian.

(viii) $$-(\overline{AB-BA})^T = -(\overline{AB}-\overline{BA})^T$$
$$= -(\overline{A}\,\overline{B}-\overline{BA})^T$$
$$= -\left((\overline{AB})^T-(\overline{BA})^T\right)$$
$$= -\left(\left(\overline{B}^T\overline{A}^T\right)-\left(\overline{A}^T\overline{B}^T\right)\right)$$
$$= -(BA-AB) \quad (\because \text{ A, B are Hermitian})$$
$$= AB-BA$$

$\therefore$ AB–BA is skew Hermitian.

Theorem 3.19 :

Let A be any square matrix. Then

(i) $A+\overline{A}^T$ is Hermitian.

(ii) $A-\overline{A}^T$ is skew Hermitian.

Proof :

(i) Let $A+\overline{A}^T = B$

Then $\overline{B} = \overline{A}+A^T$

$\therefore$ $\overline{B}^T = \left(\overline{A}+A^T\right)^T = \overline{A}^T+A = B$

Hence $A+\overline{A}^T$ is Hermitian.

(ii) Let $A-\overline{A}^T = B$

Then $\overline{B} = \overline{A}-A^T$

$\therefore$ $\overline{B}^T = \left(\overline{A}-A^T\right)^T = \overline{A}^T-A$

$$-\overline{B}^T = -\left(\overline{A}^T - A\right)$$

$$= A - \overline{A}^T = B$$

Hence $A - \overline{A}^T$ is skew Hermitian.

Theorem 3.20 :

Any square matrix A can be uniquely expressed as the sum of a Hermitian matrix and a skew Hermitian matrix.

Proof :

Let A be any square matrix.

Then $A + \overline{A}^T$ is Hermitian. (by theorem 3.19(i)).

$\therefore \frac{1}{2}\left(A + \overline{A}^T\right)$ is also a Hermitian matrix.

$A - \overline{A}^T$ is skew Hermitian (by theorem 3.19(ii))

$\therefore \frac{1}{2}\left(A - \overline{A}^T\right)$ is skew Hermitian matrix.

Now, $A = \frac{1}{2}\left(A + \overline{A}^T\right) + \frac{1}{2}\left(A - \overline{A}^T\right)$

$\therefore$ A is the sum of a Hermitian matrix and a skew Hermitian matrix.

Now, to prove the uniquesness, let A = R+S where S is a Hermitian matrix and R is a skew Hermitian matrix.

We claim that $S = \frac{1}{2}\left(A + \overline{A}^T\right)$

and $R = \frac{1}{2}\left(A - \overline{A}^T\right)$

$$A = S+R \qquad \text{--------(1)}$$

$\therefore$ $\overline{A}^T = \left(\overline{S+R}\right)^T$

$= \overline{S}^T + \overline{R}^T$

$=$ S–R (since S is Hermitian and R is skew Hermitian)

$\therefore$ $\overline{A}^T = S-R$ --------(2)

From (1) & (2)

$A+\overline{A}^T = 2S; \quad A-\overline{A}^T = 2R$

$\therefore S = \frac{1}{2}\left(A+\overline{A}^T\right); \quad \therefore R = \frac{1}{2}\left(A-\overline{A}^T\right)$

Definition :

A real square matrix A is said to be orthogonal if $AA^T = A^TA = I$.

Example :

$$A = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$$ is an orthogonal matrix.

We have to verify this.

$$A^T = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

$$AA^T = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

$$= \begin{bmatrix} \cos^2\theta+\sin^2\theta & -\cos\theta\sin\theta+\sin\theta\cos\theta \\ -\sin\theta\cos\theta+\cos\theta\sin\theta & \sin^2\theta+\cos^2\theta \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad (\because \cos^2\theta + \sin^2\theta = 1)$$

$$= I$$

$$A^TA = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}\begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$$

$$= \begin{bmatrix} \cos^2\theta+\sin^2\theta & \cos\theta\sin\theta-\cos\theta\sin\theta \\ \sin\theta\cos\theta-\sin\theta\cos\theta & \sin^2\theta+\cos^2\theta \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad (\because \cos^2\theta + \sin^2\theta = 1)$$

$$= I$$

$\therefore \quad AA^T = A^TA = I$

$\therefore$ A is an orthogonal matrix.

Theorem 3.21 :

Let A and B be orthogonal matrices of the same order. Then

(i) A^T is orthogonal

(ii) AB is orthogonal

Proof :

(i) $A^T(A^T)^T = A^TA = I$ (since A is orthogonal)

Similarly we can prove $(A^T)^TA^T = I$

$\therefore$ A^T is orthogonal.

(ii)
$$\begin{aligned}(AB)(AB)^T &= (AB)(B^TA^T)\\ &= A(BB^T)A^T\\ &= AIA^T \quad \text{(Since B is orthogonal)}\\ &= AA^T = I\end{aligned}$$

Similarly $(AB)^T(AB) = I$

Hence AB is orthogonal.

Definition :

A square matrix A is said to be an **unitary matrix** if $A\overline{A}^T = \overline{A}^TA = I$.

For example $\begin{bmatrix}0 & -i\\ i & 0\end{bmatrix}$ is unitary.

Note : Any unitary matrix over R is an orthogonal matrix.

Theorem 3.22 :

If A and B are unitary matrices of the same order, then AB is also an unitary matrix.

Proof :

$$\begin{aligned}(AB)(\overline{AB})^T &= (AB)(\overline{AB})^T = (AB)\left(\overline{B}^T\overline{A}^T\right)\\ &= A\left(B\overline{B}^T\right)\overline{A}^T = AI\overline{A}^T \quad \text{(since B is unitary)}\\ &= A\overline{A}^T = I\end{aligned}$$

Similarly $(\overline{AB})^T(AB) = I$

$\therefore$ Hence AB is unitary.

Exercises :

1. Give examples of each of the following types of matrices; upper triangular matrix, lower triangular matrix, diagonal matrix, scalar matrix, symmetric matrix, Hermitian matrix, skew Hermitian metrix, orthogonal matrix and unitary matrix.

2. Give examples of matrices over the field of complex numbers which are

   (a) Symmetric but not Hermitian

   (b) Skew symmetric but not skew Hermitian

3. Show that the product of two upper (lower) triangular matrices of the same order is again an upper (lower) triangular matrix.

4. Show that the product of two diagonal matrices of the same order is again a diagonal matrix.

5. Show that any two diagonal matrices of the same order commute.

6. For any square matrix A show that AA^T and A^TA are symmetric.

7. Show that if A is symmetric then A^T is symmetric.

8. Show that if A is skew symmetric then A^2 is symmetric and A^3 is skew symmetric.

9. Show that if A and B are symmetric matrices of the same order then AB–BA is skew symmetric.

10. Show that if A and B are skew symmetric matrices then AB is symmetric iff AB=BA.

11. Show that any Hermitian matrix A can be written as A = B+iC where B is a real symmetric matrix and C is a real skew symmetric matrix. State and prove a similar result for a skew Hermitian matrix.

12. Show that every square matrix A can be uniquely expressed as A = B+iC where B and C are Hermitian.

13. A square matrix A is caled an **idempotent matrix** if $A^2=A$

Show that $\begin{bmatrix} 2 & -3 & -5 \\ -1 & 4 & 5 \\ 1 & -3 & -4 \end{bmatrix}$ and $\begin{bmatrix} -1 & 3 & 5 \\ 1 & -3 & -5 \\ -1 & 3 & 5 \end{bmatrix}$ are idempotent matrices.

14. Show that if AB = A and BA = B then A and B are idempotent matrices.

15. Show that if A is an idempotent matrix, then B=I–A is also an idempotent matrix and AB=BA=0.

16. A square matrix A is said to be **nilpotent** if $A^n=0$ for some positive integer n.

Show that $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 3 \\ 5 & 2 & 6 \\ -2 & -1 & -3 \end{bmatrix}$ are nilpotent.

17. A square matrix A is said to be **involutory** if $A^2=I$.

Show that $\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$ and $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$ are involutory.

18. Show that a square matrix A is involutory iff (I+A)(I–A) = 0.

19. Show that $\frac{1}{2}\begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & -2 \\ -2 & 2 & -1 \end{bmatrix}$ is an orthogonal matrix.

20. Show that $\frac{1}{2}\begin{bmatrix} 1+i & -1+i \\ 1+i & 1-i \end{bmatrix}$ is an unitary matrix.

INNER PRODUCT SPACES

Introduction :

In this chapter we place an additional structure on a vector space V to obtain an inner product space, and in this context these concepts are defined.

We know that in the usual three dimensional vector space $V_3(R)$ it is possible to talk about the length of a vector and angle between two vectors. These concepts of length and angle can be defined in terms of the usual "dot Product" or "Scalar product" of two vectors. The dot product of $u = (a_1,b_1,c_1)$ and $v = (a_2,b_2,c_2)$ is defined by

$$u.v = a_1a_2 + b_1b_2 + c_1c_2$$

We note that the length of u is given by $\sqrt{u.u}$ and the angle θ between u and v is determined by $\cos\theta \frac{u.v}{\sqrt{u.u}\sqrt{v.v}}$. Hence u and v are perpendicular or orthogonal iff $u.v = 0$.

An inner product on a vector space is a generalisation of the dot product and in terms of such an inner product we can define the length of a vector and angle between two vectors. Our study about angle will be restricted to the concept of perpendicularity of two vectors.

Throughout this section we shall deal only with vector spaces over the field F of real or complex numbers.

4.1. Definition and Examples :

Definition : Let V be a vector space over F. An **inner product** on V is a function which assings to each ordered pair of vectors u,v in V a scalar in F denoted by $\langle u,v\rangle$ satisfying the following conditions.

(i) $\langle u+v+w\rangle = \langle u,w\rangle + \langle v,w\rangle$

(ii) $\langle \alpha u,v\rangle = \alpha\langle u,v\rangle$

(iii) $\langle u,v\rangle = \overline{\langle v,u\rangle}$ where $\overline{\langle v,u\rangle}$ is the complex conjugate of $\langle u,v\rangle$.

(iv) $\langle u,u\rangle \geq 0$ and $\langle u,u\rangle = 0$ iff $u = 0$.

A vector space with an inner product defined on it is called an **inner product space**. An inner product space is called an **Euclidean space** or **unitary space** according as F is the field of real numbers or complex numbers.

Note 1 : It F is the field of real numbers then condition (iii) takes the form $\langle u,v\rangle = \langle v,u\rangle$. Further (iii) asserts that $\langle u,u\rangle$ is always real and hence (iv) is meaningful whether F is the field of real or complex numbers.

Note 2 :

$$\langle u,\alpha v\rangle = \bar{\alpha}\langle u,v\rangle$$

For

$$\langle u,\alpha v\rangle = \overline{\langle \alpha v,u\rangle}$$

$$= \overline{\alpha\langle v,u\rangle}$$

$$= \overline{\alpha}\,\overline{\langle v,u\rangle}$$

$$= \bar{\alpha}\langle u,v\rangle$$

Note 3 :

$$\langle u,v+w\rangle = \langle u,v\rangle + \langle u,w\rangle$$

For,

$$\langle u,v+w\rangle = \overline{\langle v+w,u\rangle}$$

$$= \overline{\langle v,u\rangle + \langle w,u\rangle}$$

$$= \overline{\langle v,u\rangle} + \overline{\langle w,u\rangle}$$

$$= \langle u,v\rangle + \langle u,w\rangle.$$

Note 4 :

$$\langle u,0\rangle = \langle 0,v\rangle = 0$$

For,

$$\langle u,0\rangle = \langle u,00\rangle$$

$$= 0\langle u,0\rangle = 0$$

Similarly,

$\langle 0,v\rangle = 0$

Examples :

1. $V_n(R)$ is a real inner product space with inner product defined by

$$\langle x,y\rangle = x_1y_1+x_2y_2+.....+x_ny_n \text{ where}$$

$$x = (x_1,x_2,.....x_n) \text{ and}$$

$$y = (y_1,y_2,....y_n)$$

This is called the standard inner product on $V_n(R)$.

Proof :

Let $x,y,z \in V_n(R)$ and $\alpha \in R$

(i) $$\begin{aligned} \langle x+y,z\rangle &= (x_1+y_1)z_1 + (x_2+y_2)z_2 + \ldots + (x_n+y_n)z_n \\ &= (x_1z_1 + x_2z_2 + \ldots + x_nz_n) + (y_1z_1 + y_2z_2 + \ldots + y_nz_n) \\ &= \langle x,z\rangle + \langle y,z\rangle \end{aligned}$$

(ii) $$\begin{aligned} \langle \alpha x,y\rangle &= \alpha x_1y_1 + \alpha x_2y_2 + \ldots + \alpha x_ny_n \\ &= \alpha\,(x_1y_1 + x_2y_2 + \ldots + x_ny_n) \\ &= \alpha\,\langle x,y\rangle. \end{aligned}$$

(iii) $$\begin{aligned} \langle x,y\rangle &= x_1y_1 + x_2y_2 + \ldots + x_ny_n \\ &= y_1x_1 + y_2x_2 + \ldots + y_nx_n \\ &= \langle y,x\rangle. \end{aligned}$$

(iv) $$\langle x,x\rangle = x_1^2 + x_2^2 + \ldots + x_n^2 \geq 0 \text{ and}$$

$$\langle x,x\rangle = 0 \text{ iff } x_1 = x_2 = \ldots = x_n = 0$$

$$\therefore \quad \langle x,x\rangle = 0 \text{ iff } x = 0$$

2. $V_n(C)$ is a complex inner product space with inner product defined by

$$\langle x, y\rangle = x_1\bar{y}_1 + x_2\bar{y}_2 + \ldots + x_n\bar{y}_n$$

where $x = (x_1, x_2, \ldots, x_n)$

and $y = (y_1, y_2, \ldots y_n)$

Proof :

Let $x,y,z \in V_n(C)$ and $\alpha \in c$.

(i) $$\begin{aligned} \langle x+y,z\rangle &= (x_1+y_1)\,\bar{z}_1 + (x_2+y_2)\,\bar{z}_2 + \ldots + (x_n+y_n)\,\bar{z}_n \\ &= (x_1\bar{z}_1 + x_2\bar{z}_2 + \ldots + x_n\bar{z}_n) + (y_1\bar{z}_1 + y_2\bar{z}_2 + \ldots + y_n\bar{z}_n \\ &= \langle x,z\rangle + \langle y,z\rangle. \end{aligned}$$

(ii) $$\begin{aligned} \langle \alpha x,y\rangle &= (\alpha x_1\bar{y}_1 + \alpha x_2\bar{y}_2 + \ldots + \alpha x_n\bar{y}_n) \\ &= \alpha(x_1\bar{y}_1 + x_2\bar{y}_2 + \ldots + x_n\bar{y}_n) \\ &= \alpha\langle x,y\rangle. \end{aligned}$$

(iii) $$\langle \overline{y,x} \rangle = \overline{y_1\bar{x}_1 + y_2\bar{x}_2 + \ldots + y_n\bar{x}_n}$$

$$= \bar{y}_1x_1 + \bar{y}_2x_2 + \ldots + \bar{y}_nx_n$$

$$= x_1\bar{y}_1 + x_2\bar{y}_2 + \ldots + x_n\bar{y}_n$$

$$= \langle x,y \rangle.$$

(iv) $$\langle x,x \rangle = x_1\bar{x}_1 + \ldots + x_n\bar{x}_n$$

$$= |x_1|^2 + |x_2|^2 + \ldots + |x_n|^2 \geq 0$$

and $\langle x,x \rangle = 0$ iff $x = 0$

3. Let V be the set of all continuous real valued functions defined on the closed interval [0,1]. V is a real inner product space with inner product defined by

$$\langle f,g \rangle = \int_0^1 f(t)g(t)dt$$

Proof :

Let $f,g,h \in V$ and $\alpha \in R$.

(i) $$\langle f+g,h \rangle = \int_0^1 [f(t) + g(t)]h(t)dt$$

$$= \int_0^1 f(t)h(t)dt + \int_0^1 g(t)h(t)dt$$

$$= \langle f,h \rangle + \langle g,h \rangle$$

(ii) $$\langle \alpha f,g \rangle = \int_0^1 \alpha f(t)g(t)dt$$

$$= \alpha \int_0^1 f(t)g(t)dt$$

$$= \alpha \langle f,g \rangle.$$

(iii) $$\langle f,g \rangle = \int_0^1 f(t)g(t)dt$$

$$= \int_0^1 g(t)f(t)dt$$

$$= \langle g,f \rangle$$

(iv) $$\langle f,f \rangle = \int_0^1 [f(t)]^2 dt \geq 0 \text{ and}$$

$$\langle f,f \rangle = 0 \text{ iff } f = 0.$$

Exercises :

1. Show that in an inner product space

(i) $<\alpha u+\beta v, w> = \alpha <u, w> + \beta <v, w>$

(ii) $<u, \alpha v+\beta w> = \bar{\alpha} <u, v> + \bar{\beta} <u, w>$

(iii) $<\alpha u+\beta v, \gamma w+\delta z> = \alpha\bar{\gamma} <u, w> + \alpha\bar{\delta} <u, z> + \beta\bar{\upsilon} <v, w> + \beta\bar{\delta} <v, z)$ where

$$\alpha, \beta, \gamma, \delta \in F \text{ and } u, v, w, z \in V$$

2. Show that $V_2(R)$ is an inner product space with inner product defined by

$<x,y> = x_1y_1 + x_2y_1 - x_1y_2 + 4x_2y_2$

Where $x = (x_1, x_2)$ and $y = (y_1, y_2)$.

3. Show that $V_2(C)$ is an inner product space with inner product defined by

$<x,y> = 2x_1\bar{y}_1 + x_1\bar{y}_2 + y_2\bar{y}_1 + x_2\bar{y}_2$

Where $x = (x_1, x_2)$ and $y = (y_1, y_2)$.

4. Let V be the set of all continuous complex valued functions defined on the closed interval [0,1]. Show that V is a complex inner product space with inner product defined by

$$<f,g> = \int_0^1 f(t)\overline{g(t)}dt.$$

Definition :

Let V be an inner product space and let $x \in V$. The **norm** or **length** of x, denoted by $||x||$, is defined by $||x|| = \sqrt{<x, x>}$.

x is called a **unit vector** if $||x|| = 1$.

Solved problems :

1. Let V be the vector space of a polynomials with inner product given by $<f,g> = \int_0^1 f(t)g(t)dt$. Let $f(t) = t+2$ and $g(t) = t^2-2t-3$.

Find (i) $<f,g>$ (ii) $||f||$.

Solution :

(i) $$\langle f,g\rangle = \int_0^1 f(t)g(t)dt$$

$$= \int_0^1 (t+2)(t^2-2t-3)dt$$

$$= \int_0^1 (t^3-7t-6)dt$$

$$= \left[\frac{t^4}{4}-\frac{7t^2}{2}-6t\right]_0^1$$

$$= \frac{1}{4}-\frac{7}{2}-6$$

$$= -\frac{37}{4}.$$

(ii) $$\|f\|^2 = \langle f,f\rangle$$

$$= \int_0^1 [f(t)]^2 dt$$

$$= \int_0^1 (t+2)^2 dt$$

$$= \int_0^1 (t^2+4t+4)dt$$

$$= \left[\frac{t^3}{3}+2t^2+4t\right]_0^1$$

$$= \frac{1}{3}+2+4 = \frac{19}{3}.$$

$\therefore$ $$\|f\| = \frac{\sqrt{19}}{\sqrt{3}}$$

Exercises :

1) Find the norm of the following vectors in $V_3(R)$ with standard inner product.

(a) (1,1,1) (b) (1,2,3) (c) (3,-4,0)

(d) (4x+5y) where x = (1,–1,0) and y = (1,2,3)

2) Find the set of all unit vectors in $V_3(R)$ with standard norm.

Ans: 1.(a)$\sqrt{3}$,(b)$\sqrt{14}$,(c)5,(d)$3\sqrt{38}$

2. All points on the unit sphere with centre (0,0,0)

Theorem : 4.1

The norm defined in an inner product space V has the following properties.

(i) $\|x\| \geq 0$ and $\|x\| = 0$ iff $x = 0$

(ii) $\|\alpha x\| = |\alpha| \, \|x\|$.

(iii) $|<x,y>| \leq \|x\| \, \|y\|$ (Schwartz's inequality)

(iv) $\|x+y\| \leq \|x\|+\|y\|$ (Triangle inequality).

Proof :

(i) $\|x\| = \sqrt{<x,x>} \geq 0$ and $\|x\| = 0$ iff $x = 0$.

(ii)
$$\begin{aligned} \|\alpha x\|^2 &= <\alpha x, \alpha x> \\ &= \alpha <x, \alpha x> \\ &= \alpha \bar{\alpha} <x,x> \\ &= |\alpha|^2 \|x\|^2 \end{aligned}$$

Hence $\|\alpha x\| = |\alpha| \, \|x\|$.

(iii) The inequality is trivially true when $x = 0$ or $y = 0$. Hence let $x \neq 0$ and $y \neq 0$.

$$\text{consider, } z = y - \frac{<y,x>}{\|x\|^2} x$$

$$\begin{aligned} \text{Then } 0 &\leq <z,z> \\ &= <y - \frac{<y,x>}{\|x\|^2} x, y - \frac{<y,x>}{\|x\|^2} x> \\ &= <y,y> - \frac{<\overline{y,x}>}{\|x\|^2} <y,x> - \frac{<y,x>}{\|x\|^2} <x,y> + \frac{<y,x><\overline{y,x}>}{\|x\|^2 \|x\|^2} <x,x> \\ &= \|y^2\| - \frac{<\overline{y,x}><y,x>}{\|x\|^2} - \frac{<y,x><\overline{x,y}>}{\|x\|^2} + \frac{<y,x><\overline{y,x}>}{\|x\|^2} \\ &= \|y\|^2 - \frac{<\overline{x,y}><x,y>}{\|x\|^2} \end{aligned}$$

$\therefore\ 0 \leq \|x\|^2 \|y\|^2 - |\langle x,y\rangle|^2$

$\therefore\ |\langle x,y\rangle|^2 \leq \|x\|^2 \|y\|^2$

$\therefore\ |\langle x,y\rangle| \leq \|x\| \|y\|.$

(iv) $\|x+y\|^2 = \langle x+y, x+y\rangle$

$= \langle x,x\rangle + \langle x,y\rangle + \langle y,x\rangle + \langle y,y\rangle$

$= \|x\|^2 + \langle x,y\rangle + \langle \overline{x,y} \rangle + \|y\|^2$

$= \|x\|^2 + 2\text{Re}\langle x,y\rangle + \|y\|^2 \leq \|x\|^2 + 2|\langle x,y\rangle| + \|y\|^2$

$\leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2$ (by (iii)

$\leq (\|x\|+\|y\|)^2$

$\therefore\ \|x+y\| \leq \|x\|+\|y\|.$

Solved Problems :

1. (a) Show that in a real inner product space, if $\langle x,y\rangle=0$ then $\|x+y\|^2 = \|x\|^2 + \|y\|^2$.

Solution :

$\|x+y\|^2 = \langle x+y, x+y\rangle$

$= \langle x,x\rangle + \langle x,y\rangle + \langle y,x\rangle + \langle y,y\rangle$

$= \|x\|^2 + \langle x,y\rangle + \langle \overline{x,y} \rangle + \|y\|^2$

$= \|x\|^2+2\text{Re}\langle x.y\rangle+\|y\|^2$

$= \|x\|^2 + \|y\|^2 (\because \langle x,y\rangle = 0$

$\therefore\ \|x+y\|^2 = \|x\|^2 + \|y\|^2$

(b) Show that in a real inner product space, if

$\|x\|^2 + \|y\|^2 = \|x+y\|^2$, then $\langle x,y\rangle = 0$

Given $\|x\|^2 + \|y\|^2 = \|x+y\|^2$,

$\|x\|^2 + \|y\|^2 - \|x+y\|^2 = 0$

$= \langle x,x\rangle + \langle y,y\rangle - [\langle x,x\rangle + \langle x,y\rangle + \langle y,x\rangle + \langle y,y\rangle] = 0$

$$-[\langle x,y\rangle + \langle y,x\rangle] = 0$$

$$-\left[\langle x,y\rangle + \overline{\langle x,y\rangle}\right] = 0$$

$$-2\,\mathrm{Re}\,\langle x,y\rangle = 0$$

$$2\mathrm{Re}\,\langle x,y\rangle = 0$$

$$\therefore \langle x,y\rangle = 0.$$

Exercises:

1. Applying Schwartz's inequality to Vn(C) with standard inner product, show that

$$\left|\sum_{i=1}^{n} x_i \bar{y}_i\right| \le \left|\sum_{i=1}^{n} |x_i|^2\right|^{1/2} \left|\sum_{i=1}^{n} |y_i|^2\right|^{1/2}$$

2. Show that in any inner product space V

$$\|x+y\|^2 + \|x-y\|^2 = 2\left(\|x\|^2 + \|y\|^2\right)$$

3. Show that in any inner product space

$$\|\alpha x + \beta y\|^2 = |\alpha|^2\|x\|^2 + \alpha\bar{\beta}\langle x,y\rangle + \bar{\alpha}\beta\langle y,x\rangle + |\beta|^2\|y\|^2$$

4. Show that if equality is valid in schwartz's inequality or triangle inequality then x and y are linearly dependent. Is the converse true?

5. In an inner product space we define the distance between any two vectors x and y by d(x,y) = ||x-y||. Show that

(a) d (x,y) ≥0 and d (x,y) = 0 iff x = y

(b) d (x,y) = d (y,x)

(c) d (x,y) ≤ d (x,z) + d (z,y).

4.2. ORTHONORMAL BASIS

Definition :

Let V be an inner product space and let x,y∈V. x is said to be **orthogonal** to y if $\langle x,y\rangle = 0$

Note 1 : x is orthogonal to y $\Rightarrow \langle x,y\rangle = 0$

$$\Rightarrow \qquad \overline{\langle x,y\rangle} = \bar{0}$$

$\Rightarrow$ $\langle y,x\rangle = 0$

$\Rightarrow$ y is orthogonal to x.

Thus x and y are orthogonal iff $\langle x,y\rangle = 0$.

Note 2 : x is orthogonal to y $\Rightarrow$ αx is orthogonal to y.

Note 3 : x_1 and x_2 are orthogonal to y $\Rightarrow$ x_1+x_2 is orthogonal to y.

Note 4 : O is orthogonal to every vector in V and is the only vector with this property.

Definition :

Let V be an inner product space. A set S of vectors in V is said to be an orthogonal set if any two distinct vectors in S are orthogonal.

Definition :

S is said to be an **orthonormal** set if S is orthogonal and $\|x\| = 1$ for all $x \in S$.

Example :

The standard basis $\{e_1,e_2,.....e_n\}$ in R^n or C^n is an orthogonal set with respect to the standard inner product.

Theorem 4.2 :

Let $S = \{v_1,v_2,...v_n\}$ be an orthogonal set of non-zero vectors in an inner product space V. Then S is linearly independent.

Proof :

Let $\alpha_1v_1+\alpha_2v_2+.....+\alpha_nv_n = 0$

Then $\langle \alpha_1v_1+\alpha_2v_2+.....+\alpha_nv_n, v_1\rangle = \langle 0,v_1\rangle = 0$

$\therefore$ $\alpha_1\langle v_1,v_1\rangle + \alpha_2\langle v_2,v_1\rangle +...+\alpha_n\langle v_n,v_1\rangle = 0$

$\therefore$ $\alpha_1\langle v_1,v_1\rangle = 0$ (Since S is orthogonal)

$\therefore$ $\alpha_1 = 0$ (Since $v_1 \neq 0$)

Smilarly $\alpha_2 = \alpha_3 = ... = \alpha_n = 0$

Hence S is linearly independent.

Theorem 4.3 :

Let $S = \{v_1, v_2, \ldots v_n\}$ be an orthogonal set of non-zero vectors in V. Let $v \in V$ and $v = \alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n$. Then $\alpha_k = \dfrac{\langle v, v_k \rangle}{\|v_k\|^2}$.

Proof :

$$\begin{aligned} \langle v, v_k \rangle &= \langle \alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n, v_k \rangle \\ &= \alpha_1 \langle v_1, v_k \rangle + \alpha_2 \langle v_2, v_k \rangle + \ldots + \alpha_k \langle v_k, v_k \rangle + \ldots + \alpha_n \langle v_n, v_k \rangle \\ &= \alpha_k \langle v_k, v_k \rangle \text{ (Since S is orthogonal)} \\ &= \alpha_k \|v_k\|^2 \end{aligned}$$

$$\therefore \quad \alpha_k = \frac{\langle v, v_k \rangle}{\|v_k\|^2}$$

4.3. GRAM - SCHMIDT ORTHOGONALIZATION PROCESS

Theorem 4.4 :

Every finite dimensional inner product space has an orthonormal basis.

Proof :

Let V be a finite dimensional inner product space. Let $\{v_1, v_2, \ldots v_n\}$ be a basis for V. From this basis we shall construct an orthonormal basis $\{w_1, w_2, \ldots w_n\}$ by means of a construction known as Gram-Schmidt orthogonalisation process.

First we take $w_1 = v_1$

Let $w_2 = v_2 - \dfrac{\langle v_2, w_1 \rangle}{\|w_1\|^2} w_1$

We claim that $w_2 \neq 0$. For if $w_2 = 0$ then v_2 is a scalar multiple of w_1 and hence of v_1 which is a contradiction, since v_1, v_2 are linearly independent.

$$\text{Also, } \langle w_2, w_1 \rangle = \left\langle v_2 - \frac{\langle v_2, w_1 \rangle}{\|w_1\|^2} w_1, w_1 \right\rangle$$

$$= \left\langle v_2 - \frac{\langle v_2, v_1 \rangle}{\|v_1\|^2} v_1, v_1 \right\rangle \ (\because w_1 = v_1)$$

$$= \langle v_2, v_1 \rangle - \frac{\langle v_2, v_1 \rangle}{\|v_1\|^2} \langle v_1, v_1 \rangle$$

$$= \langle v_2, v_1 \rangle - \langle v_2, v_1 \rangle$$

$$= 0$$

Now, suppose that we have constructed non-zero orthogonal vectors $w_1, w_2, \ldots w_k$. Then put

$$w_{k+1} = v_{K+1} - \sum_{j=1}^{k} \frac{\langle v_{k+1}, w_j \rangle}{\|w_j\|^2} w_j$$

We claim that $w_{K+1} \neq 0$. For, if $w_{k+1} = 0$, then v_{k+1} is a linear combination of $w_1, w_2, \ldots w_k$ and hence is a linear combination of $v_1, v_2, \ldots v_k$ which is a contradiction since $v_1, v_2 \ldots v_{k+1}$ are lineanly independent.

Also,

$$\langle w_{k+1}, w_i \rangle = \langle v_{k+1}, w_i \rangle - \sum_{j=1}^{k} \frac{\langle v_{k+1}, w_j \rangle}{\|w_j\|^2} \langle w_j, w_i \rangle$$

$$= \langle v_{k+1}, w_i \rangle - \frac{\langle v_{k+1}, w_i \rangle}{\|w_i\|^2} \langle w_i, w_i \rangle$$

$$= \langle v_{k+1}, w_i \rangle - \langle v_{k+1}, w_i \rangle$$

$$= 0$$

Thus, continuing in this way we ultimately obtain a non-zero orthogonal set $\{w_1, w_2, \ldots w_n\}$

By theorem 4.2, this set is linearly independent and hence a basis.

To obtain an orthonormal basis we replace each w_i by $\frac{w_i}{\|w_i\|}$.

Solved problems

Problem 1 :

Apply Gram-Schmidt process to construct an orthonormal basis for $V_3(R)$ with the standard inner product for the basis $\{v_1,v_2,v_3\}$ where $v_1 = (1,0,1)$; $v_2 = (1,3,1)$ and $v_3 = (3,2,1)$.

Solution :

Take $w_1 = v_1 = (1,0,1)$

Then $\|w_1\|^2 = \langle w_1,w_1\rangle = 1^2+0^2+1^2 = 2$

and $\langle w_1,v_2\rangle = 1+0+1 = 2$

Put $$w_2 = v_2 - \frac{\langle v_2,w_1\rangle}{\|w_1\|^2} w_1$$

$$= (1,3,1) - (1,0,1)$$

$$= (0,3,0)$$

$\therefore$ $\|w_2\|^2 = 9$

Also, $\langle w_2,v_3\rangle = 0+6+0 = 6$ and

$\langle w_1,v_3\rangle = 3+0+1 = 4$

Now, $$w_3 = v_3 - \frac{\langle v_3,w_1\rangle}{\|w_1\|^2} w_1 - \frac{\langle v_3,w_2\rangle}{\|w_2\|^2} w_2$$

$$= (3,2,1) - \frac{4}{2}(1,0,1) - \frac{6}{9}(0,3,0)$$

$$= (3,2,1) - 2(1,0,1) - \frac{2}{3}(0,3,0)$$

$$= (1,0,-1)$$

$\therefore$ $\|w_3\|^2 = 2$

The orthogonal basis is $\{(1,0,1), (0,3,0), (1,0,-1)\}$.

Hence the orthonormal basis is $\left\{\left[\frac{1}{\sqrt{2}},0,\frac{1}{\sqrt{2}}\right],(0,1,0),\left[\frac{1}{\sqrt{2}},0,\frac{-1}{\sqrt{2}}\right]\right\}$

Problem 2 :

Let V be the set of all polynomials of degree ≤ 2 together with the zero polynomial. V is a real inner product space with inner product defined by $\langle f,g\rangle = \int_{-1}^{1} f(x)g(x)dx$. Starting with the basis $\{1,x,x^2\}$, obtain an orthonormal basis for V.

Solution :

Let $v_1 = 1$, $V_2 = x$ and $v_3 = x^2$

Let $w_1 = v_1$

Then $$\|w_1\|^2 = \langle w_1,w_1\rangle = \int_{-1}^{1} 1dx = 2$$

Hence $$\|w_1\| = \sqrt{2}$$

$$w_2 = v_2 - \frac{\langle v_2,w_1\rangle}{\|w_1\|^2} w_1$$

$$= x - \frac{1}{2}\int_{-1}^{1} xdx$$

$$= x.$$

$$\therefore \|w_2\|^2 = \langle w_2,w_2\rangle = \int_{-1}^{1} x^2dx = \frac{2}{3}$$

Now, $$w_3 = v_3 - \frac{\langle v_3,w_1\rangle}{\|w_1\|^2} w_1 - \frac{\langle v_3,w_2\rangle}{\|w_2\|^2} w_2$$

$$= x^2 - \frac{1}{2}\int_{-1}^{1} x^2dx - \left[\frac{3x}{2}\right]\int_{-1}^{1} x^3dx$$

$$= x^2 - \frac{1}{3}.$$

$$\therefore \|w_3\|^2 = \langle w_3,w_3\rangle = \int_{-1}^{1}\left(x^2 - \frac{1}{3}\right)^2 dx = \frac{8}{45}$$

Hence the orthogonal basis is $\left\{1, x, x^2 - \frac{1}{3}\right\}$

$\therefore$ The required orthonormal basis is $\left\{\frac{1}{\sqrt{2}}, \frac{\sqrt{3}}{\sqrt{2}}x, \frac{\sqrt{10}}{4}\left(3x^2 - 1\right)\right\}$.

Problem 3 :

Find a vector of unit length which is orthogonal to (1,3,4) in $V_3(R)$ with standard inner product.

Solution :

Let $x = (x_1,x_2,x_3)$ be any vector orthogonal to (1,3,4). Then $x_1+3x_2+4x_3 = 0$.

Any solution of this equation gives a vector orthogonal to (1,3,4).

For example $x = (1,1,-1)$ is orthogonal to (1,3,4).

Also, $\|x\| = \sqrt{3}$

Hence a unit vector orthogonal to (1,2,3) is given by $\left(\frac{1}{\sqrt{3}},\frac{1}{\sqrt{3}},\frac{-1}{\sqrt{3}}\right)$

Note :

The set of all vectors orthogonal to (1,3,4) are the points lying on the plane $x+3y+4z = 0$, which is a two dimensional subspace of $V_3(R)$.

Problem 4 :

Find an orthogonal basis containing the vector (1,3,4) for $V_3(R)$ with the standard inner product.

Solution :

(1,1,–1) is a vector orthogonal to (1,3,4) (Refer the above problem).

Now, let $y = (y_1,y_2,y_3)$ be a vector orthogonal to both (1,3,4) and (1,1,–1).

Then $y_1+3y_2+4y_3 = 0$

$y_1+y_2-y_3 = 0$

Any solution of this system of equations gives a vector orthogonal to (1,3,4) and (1,1,–1).

For example, (7,–5,2) is one such vector. (by cross multiplication method)

Hence {(1,3,4), (1,1,–1), (7,–5,2)} is an orthogonal basis containing (1,3,4).

Problem 5 :

Applying Gram–Schmidt process find the orthonormal basis of $V_3(R)$ with the standard inner product starting with the following basis. $\{v_1,v_2,v_3\}$ where $v_1 = (1,-1,0)$, $v_2 = (2,-1,-2)$ $v_3 = (1,-1,-2)$

Solution :

Take $w_1 = v_1 = (1,-1,0)$

Then $\|w_1\|^2 = \langle w_1, w_1 \rangle = 1^2+(-1)^2+0 = 1+1 = 2$

and $\langle w_1, v_2 \rangle = 2+1+0 = 3$

Put $$w_2 = v_2 - \frac{\langle v_2, w_1 \rangle}{\|w_1\|^2} w_1$$

$$= (2,-1,-2) - \frac{3}{2}(1,-1,0)$$

$$= \left(\frac{1}{2}, 1/2, -2\right)$$

$\therefore$ $$\|w_2\|^2 = \left(\frac{1}{2}\right)^2 + (1/2)^2 + (-2)^2$$

$$= \frac{1}{4}+\frac{1}{4}+4 = \frac{18}{4} = \frac{9}{2}$$

Also, $$\langle w_2, v_3 \rangle = \frac{1}{2}-\frac{1}{2}+4 = 4$$

$$\langle w_1, v_3 \rangle = 1+1+0 = 2$$

Now,

$$w_3 = v_3 - \frac{\langle v_3, w_1 \rangle}{\|w_1\|^2} w_1 - \frac{\langle v_3, w_2 \rangle}{\|w_2\|^2} w_2$$

$$= (1,-1,-2) - \frac{2}{2}(1,-1,0) - \frac{4\times 2}{9}\left(\frac{1}{2}, \frac{1}{2}, -2\right)$$

$$= (1,-1,-2) - (1,-1,0) - \frac{8}{9}\left(\frac{1}{2}, \frac{1}{2}, -2\right)$$

$$= (1,-1,-2) - (1,-1,0) - \left(\frac{4}{9}, \frac{4}{9}, -\frac{16}{9}\right)$$

$$= (0,0,-2) - \left(\frac{4}{9}, \frac{4}{9}, -\frac{16}{9}\right)$$

$$= \left(-\frac{4}{9}, -\frac{4}{9}, -\frac{2}{9}\right)$$

$$\therefore \quad \|w_3\|^2 = \left(-\frac{4}{9}\right)^2+\left(-\frac{4}{9}\right)^2+\left(-\frac{2}{9}\right)^2$$

$$= \frac{16}{81}+\frac{16}{81}+\frac{4}{81}$$

$$= \frac{36}{81}=\frac{4}{9}$$

$\therefore$ The orthogonal basis is

$$\left\{(1,-1,0),\left(\frac{1}{2},\frac{1}{2},-2\right),\left(-\frac{4}{9},-\frac{4}{9},-\frac{2}{9}\right)\right\}$$

Hence the orthonormal basis is

$$\left\{\left(\frac{1}{\sqrt{2}},\frac{-1}{\sqrt{2}},0\right),\left(\frac{\sqrt{2}}{6},\frac{\sqrt{2}}{6},\frac{-2\sqrt{2}}{3}\right),\left(\frac{-2}{3},\frac{-2}{3},\frac{-1}{3}\right)\right\}$$

Problem 6 :

Let V be the set of all polynomials of degree ≤ 2 over R with inner product defined by

$\langle f,g\rangle = \int_0^1 f(x)g(x)dx$. Starting with the basis

$\{1,x,x^2\}$. Obtain an orthonormal basis for V.

Solution :

Let $v_1 = 1$, $v_2 = x$ and $v_3 = x^2$

Let $w_1 = v_1$

Then $$\|w_1\|^2 = \langle w_1, w_1\rangle = \int_0^1 f(x)g(x)dx$$

$$= \int_0^1 1dx = [x]_0^1 = 1-0 = 1.$$

Hence $$\|w_1\| = 1.$$

$$W_2 = v_2 - \frac{\langle v_2, w_1\rangle}{\|w_1\|^2}w_1$$

$$= x - \int_0^1 xdx$$

$$= x - \left[\frac{x^2}{2}\right]_0^1 = \left(x - \frac{1}{2}\right)$$

$\therefore$ $\|w_2\|^2 = \langle w_2, w_2\rangle = \int_0^1 (x - 1/2)^2 dx$

$$= \int_0^1 \left(x^2 + \frac{1}{4} - x\right) dx$$

$$= \left[\frac{x^3}{3} + \frac{x}{4} - \frac{x^2}{2}\right]_0^1$$

$$= \frac{1}{3} + \frac{1}{4} - \frac{1}{2} = \frac{4+3-6}{12} = \frac{1}{12}$$

Now, $W_3 = v_3 - \frac{\langle v_3, w_1\rangle}{\|w_1\|^2} w_1 - \frac{\langle v_3, w_2\rangle}{\|w_2\|^2} w_2$

$$= x^2 - \int_0^1 x^2 dx - \frac{\left(x - \frac{1}{2}\right)}{\frac{1}{12}} \int_0^1 x^2 \left(x - \frac{1}{2}\right) dx$$

$$= x^2 - \int_0^1 x^2 dx - 12 \frac{(2x-1)}{2} \int_0^1 \left(x^3 - \frac{x^2}{2}\right) dx$$

$$= x^2 - \left[\frac{x^3}{3}\right]_0^1 - 6(2x-1)\left[\frac{x^4}{4} - \frac{x^3}{6}\right]_0^1$$

$$= x^2 - \frac{1}{3} - 6(2x-1)\left(\frac{1}{4} - \frac{1}{6}\right)$$

$$= x^2 - \frac{1}{3} - 6(2x-1)\left(\frac{1}{12}\right)$$

$$= x^2 - \frac{1}{3} - \frac{(2x-1)}{2}$$

$$= x^2 - \frac{1}{3} - x + \frac{1}{2} = x^2 - x + \frac{1}{6}$$

$$\therefore \quad \|w_3\|^2 = \langle w_3, w_3 \rangle = \int_0^1 \left(x^2 - x + \frac{1}{6}\right)^2 dx$$

$$= \int_0^1 \left(x^4 - 2x^3 + \frac{2x^2}{6} + x^2 - \frac{2x}{6} + \frac{1}{36}\right) dx$$

$$= \left[\frac{x^5}{5} - \frac{2x^4}{4} + \frac{2}{6}\frac{x^3}{3} + \frac{x^3}{3} - \frac{2}{6}\frac{x^2}{2} + \frac{1}{36}x\right]_0^1$$

$$= \frac{1}{5} - \frac{2}{4} + \frac{2}{18} + \frac{1}{3} - \frac{1}{6} + \frac{1}{36}$$

$$= \frac{1}{180}$$

Hence the orthogonal basis is $\left\{1, (x - 1/2), \left(x^2 - x + 1/6\right)\right\}$

$\therefore$ The required orthogonal basis is $\left\{1, (2x-1)\sqrt{3}, \left(6x^2 - 6x + 1\right)\sqrt{5}\right\}$

Exercises :

1. Applying Gram-Schmidt process find the orthonormal basis of $v_3(R)$ with the standard inner product starting with the following bases.

   (a) (2,–1,0), (4,–1,0), (4,0,–1)

   (b) (1,0,1), (1,0,–1), (0,3,4)

2. Obtain an orthogonal basis for $V_3(R)$ with standard inner product containing the vectors..

   (a) (1,1,–1) and (1,0,1)

   (b) (7,–1,1)

Answers :

1. (a) $\left\{\left(2\sqrt{5}/5, -\sqrt{5}/5, 0\right), \left(\sqrt{5}/5, 2\sqrt{5}/5, 0\right), (0,0,-1)\right\}$

   (b) $\left\{\left(1/\sqrt{2}, 0, 1/\sqrt{2}\right), \left(1/\sqrt{2}, 0, -1/\sqrt{2}\right), (0,1,0)\right\}$

4.4. ORTHOGONAL COMPLEMENT

Definition :

Let V be an inner product space. Let S be a subset of V. The orthogonal complement of S, denoted by $S^{\perp}$, is the set of all vectors in V which are orthogonal to every vector of S.

(i.e) $S^{\perp} = \{x/x \in V \text{ and } \langle x,u\rangle = 0 \text{ for all } u \in S\}$

Examples :

1. $V^{\perp} = \{0\}$ and $\{0\}^{\perp} = V$. Since 0 is the only vector which is orthogonal to every vector.

2. Let $S = \{(x,0,0)/x \in R\} \subseteq V_3(R)$ with standard inner product. Then

$S^{\perp} = \{(0,y,z)/y,z \in R\}$

(i.e.,) The orthogonal complement of the x axis is the yz–plane.

Theorem 4.5

If S is any subset of V then $S^{\perp}$ is a subspace of V.

Proof :

Clearly $0 \in S^{\perp}$ and hence $S^{\perp} \neq \phi$.

Now, let $x,y \in S^{\perp}$ and $\alpha,\beta \in F$.

Then $\langle x,u\rangle = \langle y,u\rangle = 0$ for all $u \in S$.

$\langle \alpha x + \beta y, u\rangle = \alpha\langle x,u\rangle + \beta\langle y,u\rangle = 0$ for all $u \in S$.

$\therefore \alpha x+\beta y \in S^{\perp}$.

Hence $S^{\perp}$ is a subspace of V.

Theorem 4.6 :

Let V be a finite dimensional inner product space. Let W be a subspace of V Then V is the direct sum of W and $W^{\perp}$.

(ie) $V = W \oplus W^{\perp}$.

Proof :

We shall prove that

(i) $W \cap W^{\perp} = \{0\}$ and

(ii) $W + W^{\perp} = V$.

(i) Let $v \in W \cap W^{\perp}$

Then $v \in W$ and $v \in W^{\perp}$

Now, $v \in W^{\perp} \Rightarrow v$ is orthogonal to every vector in W. In particular, v is orthogonal to itself.

$\therefore \langle v,v \rangle = 0$ and hence $v = 0$.

Hence $W \cap W^{\perp} = \{0\}$.

(ii) Let $\{v_1, v_2, \ldots . v_r\}$ be an orthonormal basis for W. Let $v \in V$.

consider,

$$v_0 = v - \langle v,v_1 \rangle v_1 - \langle v,v_2 \rangle v_2 \ldots\ldots - \langle v,v_r \rangle v_r$$

$\therefore$ $\langle v_0, v_i \rangle = \langle v,v_i \rangle - \langle v,v_i \rangle \langle v_i,v_i \rangle$

(Since $\langle v_i,v_j \rangle = 0$ if $i \neq j$)

$= \langle v,v_i \rangle - \langle v,v_i \rangle$

(Since $\langle v_i,v_i \rangle = 1$)

$= 0$

$\therefore$ v_0 is orthogonal to each of $v_1, v_2, \ldots . v_r$ and hence is orthogonal to every vector in W.

Hence $v_0 \in W^{\perp}$ and $v = \left[\langle v,v_1 \rangle v_1 + \langle v,v_2 \rangle v_2 + .. + \langle v,v_r \rangle v_r\right] + v_0 \in W + W^{\perp}$

$\therefore$ $V = W \oplus W^{\perp}$

Hence the theorem.

Corollary:

$\dim V = \dim W + \dim W^{\perp}$

Proof :

$\dim V = \dim(W \oplus W^{\perp}) = \dim W + \dim W^{\perp}$

Theorem 4.7

Let V be a finite dimensional inner product space. Let W be a subspace of V. Then $\left(W^{\perp}\right)^{\perp} = W$.

Proof :

Let $w \in W$. Then for any $u \in W^{\perp}$, $\langle w,u \rangle = 0$.

Hence $w \in \left(W^{\perp}\right)^{\perp}$.

Thus $W \subseteq \left(W^{\perp}\right)^{\perp}$ -------(1)

Now by theorem 4.6, $V = W \oplus W^{\perp}$

Also, $V = W^{\perp} \oplus \left(W^{\perp}\right)^{\perp}$

Hence $\dim W = \dim\left(W^{\perp}\right)^{\perp}$ -------(2)

From (1) and (2),

We get $W = \left(W^{\perp}\right)^{\perp}$

Solved Problems :

Problem 1 :

Let V be an inner product space and let S_1 and S_2 be subsets of V. Then $S_1 \subseteq S_2 \Rightarrow S_2^{\perp} \subseteq S_1^{\perp}$.

Solution :

Let $u \in S_2^{\perp}$

Then $\langle u,v \rangle = 0$ for all $v \in S_2$.

But $S_1 \subseteq S_2$. Hence $\langle u,v \rangle = 0$ for all $v \in S_1$

Hence $u \in S_1^{\perp}$

Thus $S_2^{\perp} \subseteq S_1^{\perp}$

Problem 2 :

Let W_1 and W_2 be subspaces of a finite dimensional linner product space. Then

(i) $(W_1 + W_2)^{\perp} = W_1^{\perp} \cap W_2^{\perp}$.

(ii) $(W_1 \cap W_2)^{\perp} = W_1^{\perp} + W_2^{\perp}$.

Solution :

(i) We know that $W_1 \subseteq W_1 + W_2$

$\therefore (W_1 + W_2)^{\perp} \subseteq W_1^{\perp}$ (by the previous problem)

Similarly, $(W_1 + W_2)^{\perp} \subseteq W_2^{\perp}$.

Hence $(W_1 + W_2)^{\perp} \subseteq W_1^{\perp} \cap W_2^{\perp}$ --------(1)

Now, let $w \in W_1^{\perp} \cap W_2^{\perp}$.

Then $w \in W_1^{\perp}$

and $w \in W_2^{\perp}$.

$\therefore \langle w,u \rangle = 0$ for all $u \in W_1$ and W_2.

Now, let $v \in W_1 + W_2$

Then $v = v_1 + v_2$ Where $v_1 \in W_1$ and $v_2 \in W_2$.

$\therefore \langle w,v \rangle = \langle w, v_1 + v_2 \rangle$

$= \langle w,v_1 \rangle + \langle w,v_2 \rangle$

$= 0+0$ (Since $v_1 \in W_1$ and $v_2 \in W_2$)

$= 0$

Hence $w \in (W_1 + W_2)^{\perp}$

$\therefore W_1^{\perp} \cap W_2^{\perp} \subseteq (W_1 + W_2)^{\perp}$ --------(2)

From (1) and (2) we get

$$(W_1 + W_2)^{\perp} = W_1^{\perp} \cap W_2^{\perp}$$

(ii) We know that $W_1 \cap W_2 \subseteq W_1$

$\therefore\ W_1^{\perp} \subseteq (W_1 \cap W_2)^{\perp}$ (by the problem 1)

Similarly, $W_1 \cap W_2 \subseteq W_2$

$W_2^{\perp} \subseteq (W_1 \cap W_2)^{\perp}$

Hence $W_1^{\perp} + W_2^{\perp} \subseteq (W_1 \cap W_2)^{\perp}$ ---------(1)

Now, let, $w \in (W_1 \cap W_2)^{\perp}$

$\therefore\ \langle w, u \rangle = 0$ for all $u \in W_1 \cap W_2$

$\langle w, u \rangle = 0$ for all $u \in W_1$ and $u \in W_2$

Now, let $v \in W_1 + W_2$

Then $v = v_1 + v_2$ where $v_1 \in W_1$ and $v_2 \in W_2$.

$\therefore\ \langle w, v \rangle = \langle w, v_1 + v_2 \rangle$

$= \langle w, v_1 \rangle + \langle w, v_2 \rangle$

$= 0+0$ (Since $v_1 \in w_1$ and $v_2 \in w_2$)

Hence $w \in W_1^{\perp} + W_2^{\perp}$

$(W_1 \cap W_2)^{\perp} \subseteq W_1^{\perp} + W_2^{\perp}$ --------(2)

From (1) & (2) we get,

$(W_1 \cap W_2)^{\perp} = W_1^{\perp} + W_2^{\perp}$

Exercises :

1. Let V be an inner product space. Let S be any subspace of V. Then show that $S^{\perp} = [L(s)]^{\perp}$.

2. Find a basis for the orthogonal complement of the subspace spanned by (2,1,–2) in $V_3(R)$.

CHARACTERISTIC EQUATION AND CAYLEY HAMILTON THEOREM

Definition :

An expression of the form $A_0+A_1x+A_2x^2+...+A_nx^n$ where A_0, A_1,...A_n are square matrices of the same order and $A_n \neq 0$ is called a **matrix polynomial** of degree n.

For example,

$\begin{pmatrix}1 & 2\\ 0 & 3\end{pmatrix}+\begin{pmatrix}1 & 1\\ 2 & 1\end{pmatrix}x+\begin{pmatrix}2 & 0\\ 3 & 1\end{pmatrix}x^2$ is a matrix

Polynomial of degree 2 and it is simply the matrix $\begin{pmatrix}1+x+2x^2 & 2+x\\ 2x+3x^2 & 3+x+x^2\end{pmatrix}$

Definition :

Let A be any square matrix of order n and let I be the identity matrix of order n. Then the matrix polynomial given by A–xI is called the **characteristic matrix** of A.

The determinent |A–xI| which is an ordinary polynomial in x of degree n is called the **characteristic polynomial** of A.

The equation |A–xI| = 0 is called the **characteristic equation** of A.

Example 1 :

Let $A = \begin{pmatrix}1 & 2\\ 3 & 4\end{pmatrix}$

Then the characteristic matrix of A is A–xI given by

$$A-xI = \begin{pmatrix}1 & 2\\ 3 & 4\end{pmatrix}-x\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$$

$$= \begin{pmatrix}1-x & 2\\ 3 & 4-x\end{pmatrix}$$

∴ The characteristic polynomial of A is

$$|A-xI| = \begin{pmatrix} 1-x & 2 \\ 3 & 4-x \end{pmatrix}$$

$$= (1-x)(4-x)-6$$

$$= x^2-5x-2$$

∴ The characteristic equation of A is $|A-xI| = 0$.

∴ $x^2-5x-2 = 0$ is the characteristic equation of A.

Example 2 :

Let $$A = \begin{bmatrix} 1 & 3 & 7 \\ 4 & 2 & 3 \\ 1 & 2 & 1 \end{bmatrix}$$

Then the characteristic matrix of A is A–xI given by

$$A-xI = \begin{bmatrix} 1 & 3 & 7 \\ 4 & 2 & 3 \\ 1 & 2 & 1 \end{bmatrix} - x \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1-x & 3 & 7 \\ 4 & 2-x & 3 \\ 1 & 2 & 1-x \end{bmatrix}$$

The characteristic polynomial of A is

$$|A-xI| = \begin{bmatrix} 1-x & 3 & 7 \\ 4 & 2-x & 3 \\ 1 & 2 & 1-x \end{bmatrix}$$

$$= (1-x)[(2-x)(1-x)-6] - 3[4(1-x)-3]+7[8-(2-x)]$$

$$= (1-x)[x^2-3x-4]-3[1-4x]+7[6+x]$$

$$= -x^3+4x^2+x-4-3+12x+42+7x$$

$$= -x^3+4x^2+20x+35$$

∴ The characteristic equation of A is $|A-xI| = 0$.

∴ $x^3-4x^2-20x-35 = 0$ is the characteristic equation of A.

Theorem 5.1. Cayley Hamilton theorem

Any square matrix A satisfies its characteristic equation.

(ie) If $a_0+a_1x+a_2x^2+...+a_nx^n$ is the characteristic polynomial of degree n of A then $a_0I+a_1A+a_2A^2+....+a_nA^n = 0$.

Proof :

Let A be a square matrix of order n.

Let $|A-xI| = a_0+a_1x+a_2x^2+....a_nx^n$ -------(1)

be the characteristic polynomial of A.

Now, adj (A–xI) is a matrix polynomial of degree n–1 since each entry of the matrix adj (A–xI) is a cofactor of A–xI and hence is a polynomial of degree $\leq$ n–1.

$\therefore$ Let adj $(A-xI) = B_0+B_1x+B_2x^2+...+B_{n-1}x^{n-1}$ -------(2)

Now, $(A-xI)$ adj $(A-xI) = |A-xI|\, I$

$(\because (adjA)A = A(adjA) = |A|I)$

$\therefore (A-xI)(B_0+B_1x+...+B_{n-1}x^{n-1})$

$= (a_0+a_1x+...+a_nx^n)\, I$ using (1) and (2)

$\therefore$ Equating the coefficients of the corresponding powers of x we get

$$\begin{aligned} AB_0 &= a_0I \\ AB_1 - B_0 &= a_1I \\ AB_2 - B_1 &= a_2I \\ \text{- - - -} \quad &\text{- - - - -} \\ \text{- - - -} \quad &\text{- - - - -} \\ AB_{n-1} - B_{n-2} &= a_{n-1}I \\ -B_{n-1} &= a_nI \end{aligned}$$

Pre–multiplying the above equations by I, A, A^2,.....A^n respectively and adding we get, $a_0I+a_1A+a_2A^2+.....+a_nA^n = 0$.

Note: The inverse of a non-singular matrix can be calculated by using the cayley Hamilton theorem as follows.

Let $a_0+a_1x+a_2x^2+..a_nx^n$ be the characteristic polynomial of A.

Then by the definition, we have

$$a_0I + a_1A+a_2A^2+...+a_nA^n = 0 \qquad \text{---------(3)}$$

Since $|A-xI| = a_0+a_1x+a_2x^2+..+a_nx^n$

We get $a_0 = |A|$ (by putting $x = 0$)

$\therefore a_0 \neq 0$ (since A is a non singular matrix)

$$\therefore \quad I = -\frac{1}{a_0}[a_1A+a_2A^2+...+a_nA^n] \text{ by (3)}$$

$$\therefore \quad A^{-1} = -\frac{1}{a_0}[a_1I-a_2A+...+a_nA^{n-1}]$$

Solved problems :

Problem 1 :

Find the characteristic equation of the matrix

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

Solution :

The characteristic matrix of A is A–xI given by

$$A-xI = \begin{pmatrix} 1-x & 0 & 2 \\ 0 & 1-x & 2 \\ 1 & 2 & 0-x \end{pmatrix}$$

The characteristic polynomial of A is

$$|A-xI| = \begin{vmatrix} 1-x & 0 & 2 \\ 0 & 1-x & 2 \\ 1 & 2 & -x \end{vmatrix}$$

$$= (1-x)[(1-x)(-x)-4] + 2[0-(1-x)]$$

$$= (1-x)[-x+x^2-4]-2(1-x)$$

$$= -x^3+2x^2+3x-4-2+2x$$

$$= -x^3+2x^2+5x-6$$

The characteristic equation of A is

$$-x^3+2x^2+5x-6 = 0$$

$$\text{(ie) } x^3-2x^2-5x+6 = 0$$

Problem 2 :

Find the characteristic equation of the matrix

$$A = \begin{bmatrix} 8 & -6 & 2 \\ -6 & 7 & -4 \\ 2 & -4 & 3 \end{bmatrix}$$

Solution :

The characteristic equation of A is given by $|A-\lambda I| = 0$.

$$\text{(ie) } \begin{bmatrix} 8-\lambda & -6 & 2 \\ -6 & 7-\lambda & -4 \\ 2 & -4 & 3-\lambda \end{bmatrix} = 0$$

$(8-\lambda)[(7-\lambda)(3-\lambda)-16]+6[-6(3-\lambda)+8]+2[24-2(7-\lambda)]=0$

$(8-\lambda)(\lambda^2-10\lambda+5)-6(6\lambda-10)+2(2\lambda+10)=0$

$(8\lambda^2-80\lambda+40-\lambda^3+10\lambda^2-5\lambda)+(36\lambda-60)+(4\lambda+20)=0$

$\lambda^3-18\lambda^2+45\lambda = 0$, which represents the characteristic equation of A.

Problem 3 :

Show that the non-singular matrix $A=\begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix}$ satisfies the equation $A^2-2A-5I=0$. Hence evaluate A^{-1}.

Solution :

The characteristic polynomial of A is

$$|A-xI| = \begin{vmatrix} 1-x & 2 \\ 3 & 1-x \end{vmatrix}$$

$$= x^2-2x-5$$

∴ By cayley – Hamilton theorem $A^2-2A-5I = 0$.

$$\therefore \quad I = \frac{1}{5}(A^2-2A)$$

$$\therefore \quad A^{-1} = \frac{1}{5}(A-2I)$$

$$= \frac{1}{5}\left[\begin{pmatrix}1 & 2\\ 3 & 1\end{pmatrix}-2\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\right]$$

$$= \frac{1}{5}\begin{pmatrix}-1 & 2\\ 3 & -1\end{pmatrix}.$$

Problem 4 :

Show that the matrix

$$A = \begin{bmatrix}2 & -3 & 1\\ 3 & 1 & 3\\ -5 & 2 & -4\end{bmatrix} \text{ satisfies the equation}$$

$$A(A-I)(A+2I) = 0.$$

Solution :

The characteristic polynomial of A is

$$|A-\lambda I| = \begin{vmatrix}2-\lambda & -3 & 1\\ 3 & 1-\lambda & 3\\ -5 & 2 & -4-\lambda\end{vmatrix}$$

$$= (2-\lambda)[(1-\lambda)(-4-\lambda)-6]+3\{3(-4-\lambda)+15]+1[6+5(1-\lambda)]$$

$$= (2-\lambda)(\lambda^2+3\lambda-10)+3(-3\lambda+3)+11-5\lambda$$

$$= -\lambda^3-\lambda^2+16\lambda-20-9\lambda+9+11-5\lambda$$

$$= -\lambda^3-\lambda^2+2\lambda.$$

By cayley – Hamilton theorem – $A^3-A^2+2A = 0$.

(ie) $A^3+A^2-2A = 0$

Hence $A(A^2+A-2I) = 0$

$$A(A+2I)(A-I) = 0.$$

Problem 5 :

Find the characteristic equation of the matrix

$$A = \begin{bmatrix} 1 & 0 & 3 \\ 2 & +1 & -1 \\ 1 & -1 & 1 \end{bmatrix}$$. Verify that the matrix satisfies it own characteristic equations. Also calculate A^{-1}.

Solution :

The matrix $[A-xI] = \begin{bmatrix} 1-x & 0 & 3 \\ 2 & +1-x & -1 \\ 1 & -1 & 1-x \end{bmatrix}$

The characteristic polynomial of A is

$$|A-xI| = \begin{bmatrix} 1-x & 0 & 3 \\ 2 & 1-x & -1 \\ 1 & -1 & 1-x \end{bmatrix}$$

$$= (1-x)\,[(1-x)\,(1-x)-1]+3\,[-2-(1-x)]$$

$$= (1-x)\,[x^2-2x]+3\,[x-3]$$

$$= -x^3+3x^2-2x+3x-9$$

$$= -x^3+3x^2+x-9$$

The characteristic equation of A is $x^3-3x^2-x+9 = 0$

We have to verify that $A^3-3A^2-A+9I = 0$

First we have to find A^2&A^3.

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

$$= \begin{bmatrix} 1+3 & -3 & 3+3 \\ 2+2-1 & 1+1 & 6-1-1 \\ 1-2+1 & -1-1 & 3+1+1 \end{bmatrix}$$

$$= \begin{bmatrix} 4 & -3 & 6 \\ 3 & 2 & 4 \\ 0 & -2 & 5 \end{bmatrix}$$

$$A^3 = A.A^2 = \begin{bmatrix}1 & 0 & 3\\ 2 & 1 & -1\\ 1 & -1 & 1\end{bmatrix}\begin{bmatrix}4 & -3 & 6\\ 3 & 2 & 4\\ 0 & -2 & 5\end{bmatrix}$$

$$= \begin{bmatrix}4 & -3-6 & 6+15\\ 8+3 & -6+2+2 & 12+4-5\\ 4-3 & -3-2-2 & 6-4+5\end{bmatrix}$$

$$= \begin{bmatrix}4 & -9 & 21\\ 11 & -2 & 11\\ 1 & -7 & 7\end{bmatrix}$$

$$A^3-3A^2+A+9I = \begin{bmatrix}4 & -9 & 21\\ 11 & -2 & 11\\ 1 & -7 & 7\end{bmatrix}-\begin{bmatrix}12 & -9 & 18\\ 9 & 6 & 12\\ 0 & -6 & 15\end{bmatrix}-\begin{bmatrix}1 & 0 & 3\\ 2 & 1 & -1\\ 1 & -1 & 1\end{bmatrix}+\begin{bmatrix}9 & 0 & 0\\ 0 & 9 & 0\\ 0 & 0 & 9\end{bmatrix}$$

$$= \begin{bmatrix}4-12-1+9 & -9+9-0-0 & 21-18-3+0\\ 11-9-2+0 & -2-6-1+9 & 11-12+1+0\\ 1-0-1+0 & -7+6+1+0 & 7-15-1+9\end{bmatrix}$$

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

The equation $A^3-3A^2-A+9I_3 = 0$ is verified.

$$9I_3 = -A^3+3A^2+A$$

$$I_3 = \frac{1}{9}[-A^3+3A^2+A]$$

pre multiplying by A^{-1} on both sides, we get

$$A^{-1} = \frac{1}{9}[-A^2+3A+I]$$

$$= \frac{1}{9}\left[\begin{pmatrix}-4 & 3 & -6\\ -3 & -2 & -4\\ 0 & 2 & -5\end{pmatrix}+\begin{pmatrix}3 & 0 & 9\\ 6 & 3 & -3\\ 3 & -3 & 3\end{pmatrix}+\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}\right]$$

$$= \frac{1}{9}\begin{bmatrix}0 & 3 & 3\\ 3 & 2 & -7\\ 3 & -1 & -1\end{bmatrix}$$

Problem 6 :

Using cayley – Hamilton theorem find the inverse of the matrix.

$$\begin{bmatrix} 7 & 2 & -2 \\ -6 & -1 & 2 \\ 6 & 2 & -1 \end{bmatrix}$$

Solution :

Let $$A = \begin{bmatrix} 7 & 2 & -2 \\ -6 & -1 & 2 \\ 6 & 2 & -1 \end{bmatrix}$$

The characteristic polynomial of $A = |A-xI|$

$$= \begin{vmatrix} 7-x & 2 & -2 \\ -6 & -1-x & 2 \\ 6 & 2 & -1-x \end{vmatrix}$$

$$= (7-x)[(-(1-x))^2-4]-2[6(1+x)-12]-2\,[-12+6(1+x)]$$

$$= (7-x)\,(x^2+2x-3)\,-2\,(6x-6)\,-2\,(6x-6)$$

$$= 7x^2+14x-21-x^3-2x^2+3x-12x+12-12x+12$$

$$= -x^3+5x^2-7x+3$$

∴ By cayley – Hamilton theorem,

$$-A^3+5A^2-7A+3I_3 = 0$$

$$\therefore A^3-5A^2+7A-3I_3 = 0$$

$$\therefore 3I_3 = A^3-5A^2+7A$$

$$I_3 = \frac{1}{3}(A^3-5A^2+7A)$$

premultiplying by A^{-1} on both sides we get

$$A^{-1} = \frac{1}{3}[A^2-5A+7I_3] \qquad \text{-------(1)}$$

Now, $$A^2 = \begin{bmatrix} 7 & 2 & -2 \\ -6 & -1 & 2 \\ 6 & 2 & -1 \end{bmatrix}\begin{bmatrix} 7 & 2 & -2 \\ -6 & -1 & 2 \\ 6 & 2 & -1 \end{bmatrix} = \begin{bmatrix} 25 & 8 & -8 \\ -24 & -7 & 8 \\ 24 & 8 & -7 \end{bmatrix}$$

From (1).

$$A^{-1} = \frac{1}{3}\left[\begin{pmatrix} 25 & 8 & -8 \\ -24 & -7 & 8 \\ 24 & 8 & -7 \end{pmatrix} - \begin{pmatrix} 35 & 10 & -10 \\ -30 & -5 & 10 \\ 30 & 10 & -5 \end{pmatrix} + \begin{pmatrix} 7 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & 7 \end{pmatrix}\right]$$

$$= \frac{1}{3}\begin{bmatrix} -3 & -2 & 2 \\ 6 & 5 & -2 \\ -6 & -2 & 5 \end{bmatrix}$$

Problem 7 :

Find the inverse of the matrix $\begin{bmatrix} 3 & 3 & 4 \\ 2 & -3 & 4 \\ 0 & -1 & 1 \end{bmatrix}$ using cayley – Hamilton theorem.

Solution :

The characteristic polynomial of A

$$= |A-xI| = \begin{bmatrix} 3-x & 3 & 4 \\ 2 & -3-x & 4 \\ 0 & -1 & 1-x \end{bmatrix}$$

$$= (3-x)\,[-(3+x)(1-x)+4]-3[2(1-x)]+4(-2)$$

$$= (3-x)\,(x^2+2x+1)-6\,(1-x)-8$$

$$= -x^3+x^2+5x+3-6+6x-8$$

$$= -x^3+x^2+11x-11$$

∴ By cayley – Hamilton theorem

$$-A^3+A^2+11A-11I_3 = 0$$

Hence $$11I_3 = -(A^3-A^2-11A)$$

$$I_3 = -\frac{1}{11}(A^3-A^2-11A)$$

pre (post) multiplying by A^{-1} on both sides we get

$$A^{-1} = -\frac{1}{11}[A^2-A-11I_3]$$

$$= -\frac{1}{11}\left[\begin{pmatrix}15 & -4 & 28\\ 0 & 11 & 0\\ -2 & 2 & -3\end{pmatrix}-\begin{pmatrix}3 & 3 & 4\\ 2 & -3 & 4\\ 0 & -1 & 1\end{pmatrix}-\begin{pmatrix}11 & 0 & 0\\ 0 & 11 & 0\\ 0 & 0 & 11\end{pmatrix}\right]$$

$$= -\frac{1}{11}\begin{bmatrix}1 & -7 & 24\\ -2 & 3 & -4\\ -2 & 3 & -15\end{bmatrix}$$

$$= \begin{bmatrix}\frac{-1}{11} & \frac{7}{11} & \frac{-24}{11}\\ \frac{2}{11} & \frac{-3}{11} & \frac{4}{11}\\ 2/11 & \frac{-3}{11} & 15/11\end{bmatrix}$$

Problem 8 :

Verify cayley Hamilton's theorem for the matrix

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

Solution :

The characteristic equation of A is

$$|A-\lambda I| = 0$$

$$\therefore \quad \begin{vmatrix}1-\lambda & 2\\ 4 & 3-\lambda\end{vmatrix} = 0$$

$$\therefore \quad (1-\lambda)(3-\lambda)-8 = 0$$

$$\therefore \quad \lambda^2-4\lambda-5 = 0$$

By cayley Hamilton's theorem A satisfies its characteristic equation.

We have $A^2 - 4A - 5I = 0$.

Now, $$A^2 = \begin{pmatrix}1 & 2\\ 4 & 3\end{pmatrix}\begin{pmatrix}1 & 2\\ 4 & 3\end{pmatrix} = \begin{pmatrix}9 & 8\\ 16 & 17\end{pmatrix}$$

$$4A = \begin{pmatrix}4 & 8\\ 16 & 12\end{pmatrix} \text{ and } 5I = \begin{pmatrix}5 & 0\\ 0 & 5\end{pmatrix}$$

$$\therefore \quad A^2-4A-5I = \begin{pmatrix} 9 & 8 \\ 16 & 17 \end{pmatrix} - \begin{pmatrix} 4 & 8 \\ 16 & 12 \end{pmatrix} - \begin{pmatrix} 5 & 0 \\ 0 & 5 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = 0$$

Thus cayley Hamilton's theorem is verified.

Problem 9 :

Using Cayley – Hamilton's theorem for the matrix $A = \begin{pmatrix} 1 & 0 & -2 \\ 2 & 2 & 4 \\ 0 & 0 & 2 \end{pmatrix}$

find (i) A^{-1}, (ii) A^4.

Solution :

The characteristic equation of A is $|A-xI| = 0$

$$\therefore \quad \begin{pmatrix} 1-x & 0 & -2 \\ 2 & 2-x & 4 \\ 0 & 0 & 2-x \end{pmatrix} = 0$$

$$(1-x)\,[2-x]^2 - 2\,(0) = 0$$

$$(1-x)\,(x^2+4-4x) = 0$$

$$-x^3+5x^2-8x+4 = 0$$

$$x^3-5x^2+8x-4 = 0$$

$\therefore$ By cayley Hamilton's theorem

$$A^3-5A^2+8A-4I = 0 \qquad \text{--------(1)}$$

$$4I = A^3-5A^2+8A$$

(i) To find A^{-1} pre multiplying by A^{-1} we get

$$4A^{-1} = A^{-1}A^3-5A^{-1}A^2+8A^{-1}A$$

$$= A^2-5A+8I$$

$$\therefore \quad A^{-1} = \frac{1}{4}\,[A^2-5A+8I] \qquad \text{--------(2)}$$

Now, $$A^2 = \begin{pmatrix} 1 & 0 & -2 \\ 2 & 2 & 4 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & 0 & -2 \\ 2 & 2 & 4 \\ 0 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & -6 \\ 6 & 4 & 12 \\ 0 & 0 & 4 \end{pmatrix}$$

From 2 :

$$A^{-1} = \frac{1}{4}\left[\begin{pmatrix}1 & 0 & -6\\ 6 & 4 & 12\\ 0 & 0 & 4\end{pmatrix} - \begin{pmatrix}5 & 0 & -10\\ 10 & 10 & 20\\ 0 & 0 & 10\end{pmatrix} + \begin{pmatrix}8 & 0 & 0\\ 0 & 8 & 0\\ 0 & 0 & 8\end{pmatrix}\right]$$

$$= \frac{1}{4}\begin{pmatrix}4 & 0 & 4\\ 4 & 2 & -8\\ 0 & 0 & 2\end{pmatrix} = \begin{pmatrix}1 & 0 & 1\\ 1 & 1/2 & -2\\ 0 & 0 & -1/2\end{pmatrix}$$

$\therefore$ $$A^{-1} = \begin{pmatrix}1 & 0 & 1\\ 1 & 1/2 & -2\\ 0 & 0 & -1/2\end{pmatrix}$$

(ii) To find A^4.

From (1) $A^3 = 5A^2 - 8A + 4I$

$\therefore$ $A^4 = 5A^3 - 8A^2 + 4A$

$= 5[5A^2 - 8A + 4I] - 8A^2 + 4A$ (using (1))

$= 17A^2 - 36A + 20I$

$$= 17\begin{pmatrix}1 & 0 & -6\\ 6 & 4 & 12\\ 0 & 0 & 4\end{pmatrix} - 36\begin{pmatrix}1 & 0 & -2\\ 2 & 2 & 4\\ 0 & 0 & 2\end{pmatrix} + 20\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}$$

$$= \begin{pmatrix}17 & 0 & -102\\ 102 & 68 & 204\\ 0 & 0 & 68\end{pmatrix} - \begin{pmatrix}36 & 0 & -72\\ 72 & 72 & 144\\ 0 & 0 & 72\end{pmatrix} + \begin{pmatrix}20 & 0 & 0\\ 0 & 20 & 0\\ 0 & 0 & 20\end{pmatrix}$$

$\therefore$ $$A^4 = \begin{pmatrix}1 & 0 & -30\\ 30 & 16 & 60\\ 0 & 0 & 16\end{pmatrix}$$

Exercises :

1. Obtain the characteristic polynomial for the following matrices.

(i) $\begin{pmatrix}2 & 3\\ 0 & 5\end{pmatrix}$ (ii) $\begin{pmatrix}2 & 2\\ 0 & 3\end{pmatrix}$

2. Find the characteristic equation of the following matrices.

(i) $\begin{pmatrix} -b & -c \\ 1 & 0 \end{pmatrix}$ (ii) $\begin{pmatrix} 2 & -1 & 1 \\ -1 & 2 & -1 \\ 1 & -1 & 2 \end{pmatrix}$

(iii) $\begin{pmatrix} -b & -c & -d \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ (iv) $\begin{pmatrix} 1 & 0 & 3 \\ 2 & 1 & -1 \\ 1 & -1 & 1 \end{pmatrix}$

3. Verify cayley – Hamilton theorem for the matrix

$A = \begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix}$ and hence find A^{-1}.

4. If $A = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 2 \\ 1 & 2 & 0 \end{pmatrix}$ Prove that $A^3-2A^2-5A+6I = 0$.

5. Using cayley Hamilton theorem, find the inverses of the matrices

(i) $\begin{bmatrix} 7 & 2 & -2 \\ -6 & -1 & 2 \\ 6 & 2 & -1 \end{bmatrix}$ (ii) $\begin{bmatrix} 1 & 0 & -2 \\ 2 & 2 & 4 \\ 0 & 0 & 2 \end{bmatrix}$ (iii) $\begin{bmatrix} 1 & 2 & 3 \\ 2 & -1 & 4 \\ 3 & 1 & -1 \end{bmatrix}$ (iv) $\begin{pmatrix} 13 & -3 & 5 \\ 0 & 4 & 0 \\ -15 & 9 & -7 \end{pmatrix}$

6. If $A = \begin{pmatrix} 2 & 4 \\ 1 & 1 \end{pmatrix}$ and find A^3 and A^{-3}

7. Calculate A^4 for the matrix $A = \begin{pmatrix} 2 & 1 & 2 \\ 0 & 2 & 3 \\ 0 & 0 & 5 \end{pmatrix}$

8. Verify that the matrix $A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & -1 & 4 \\ 3 & 1 & -1 \end{pmatrix}$ satisfies its own characteristic equation and hence find A^{-1} and A^4.

9. Find the characteristic roots for the matrix $A = \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix}$ and hence evaluate A^8.

5.2. EIGEN VALUES AND EIGEN VECTORS

Definition :

Let A be an nxn matrix. A number λ is called an **eigen value** of A if there exists a non zero vector $x = \begin{pmatrix} x_1 \\ x_2 \\ . \\ . \\ x_n \end{pmatrix}$ such that $AX = \lambda X$ and X is called an **eigen vector** corresponding to the eigen value λ.

Remark 1 :

It X is an eigen vector corresponding to the eigen value λ of A, then αX where α is any non zero number, is also an eigen vector corresponding to λ.

Remark 2 :

Let x be an eigen vector corresponding to the eigen value λ of A. Then $AX = \lambda X$. So that $(A-\lambda I)\ X = 0$. Thus X is a non – trivial solution of the system of homogeneous linear equations $(A-\lambda I)\ X = 0$. Hence $|A-\lambda I| = 0$, which is the characteristic polynomial of A.

Let $|A-\lambda I| = a_0\lambda^n + a_1\lambda^{n-1}+...+a_n$

The roots of this polynomial give the eigen values of A. Hence eigen values are also called **characteristic roots.**

Properties of Eigen Values

Property 1 : Let X be an eigen vector corresponding to the eigen values λ_1 and λ_2. Then $\lambda_1 = \lambda_2$.

Proof :

By definition $X \neq 0$, $AX = \lambda_1 X$ and $AX = \lambda_2 X$

$\therefore \quad \lambda_1 X = \lambda_2 X.$

$\therefore \quad (\lambda_1-\lambda_2)\ X = 0$

Since $X \neq 0$, $\lambda_1 = \lambda_2$.

Property 2 : Let A be a square matrix.

Then (i) the sum of the eigen values of A is equal to the sum of the diagonal elements (trace) of A.

(ii) product of eigen values of A is |A|.

Proof :

(i) Let $A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ . & . & . & . \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$

The eigen values of A are the roots of the characteristic equation

$$|A-\lambda I| = \begin{pmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-\lambda & \cdots & a_{2n} \\ . & . & . & . \\ a_{n1} & a_{n2} & \cdots & a_{nn}-\lambda \end{pmatrix} = 0 \qquad \text{-----(1)}$$

Let $|A-\lambda I| = a_0\lambda^n + a_1\lambda^{n-1} + \ldots + a_n$ -----(2)

From (1) and (2) we get

$a_0 = (-1)^n$; $a_1 = (-1)^{n-1}(a_{11}+a_{22}+\ldots+a_{nn})$; ..... ------(3)

Also by putting $\lambda = 0$ in (2) we get $a_n = |A|$.

Now, let $\lambda_1, \lambda_2, \ldots \lambda_n$ be the eigen values of A.

$\therefore$ $\lambda_1, \lambda_2, \ldots \lambda_n$ are the roots of (2)

$\therefore$ $\lambda_1+\lambda_2+\ldots+\lambda_n = -\frac{a_1}{a_0}$

$= a_{11}+a_{22}+\ldots+a_{nn}$ (using 3)

$\therefore$ Sum of the eigen values = trace of A.

(ii) Product of the eigen values = Product of the roots

$= \lambda_1 \lambda_2 \ldots \lambda_n$

$= (-1)^n \frac{a_n}{a_0} = \frac{(-1)^n a_n}{(-1)^n}$

$= a_n = |A|$.

Property 3 :

The eigen values of A and its transpose A^T are the same.

Proof :

It is enough if we prove that A and A^T have the same characteristic polynomial. Since for any square matrix M,

$$|M| = |M^T| \text{ we have,}$$

$$|A-\lambda I| = |(A-\lambda I)^T| = |A^T-(\lambda I)^T|$$

$$= |A^T-\lambda I|.$$

Hence the result.

Property 4 :

If λ is an eigen value of a non singular matrix A then $1/\lambda$ is an eigen value of A^{-1}.

Proof :

Let X be an eigen vector corresponding to λ. Then $AX = \lambda X$. Since A is non singular A^{-1} exists.

$\therefore$ $$A^{-1}(AX) = A^{-1}(\lambda X)$$

$$IX = \lambda A^{-1}X$$

$\therefore$ $$A^{-1}X = \left(\frac{1}{\lambda}\right)X$$

$\therefore \left(\frac{1}{\lambda}\right)$ is an eigen value of A^{-1}.

Corollary :

If $\lambda_1, \lambda_2, \ldots \lambda_n$ are the eigen values of a non singular matrix A then $\frac{1}{\lambda_1}, \frac{1}{\lambda_2}, \ldots \frac{1}{\lambda_n}$ are the eigen values of A^{-1}.

Property 5 :

It λ is an eigen value of A then $k\lambda$ is an eigen value of kA where k is a scalar.

Proof :

Let X be an eigen vector corresponding to λ.

Then $AX = \lambda X$ ------(1)

Now, $(kA)X = k(Ax)$

$= k(\lambda x)$ (by 1)

$= (k\lambda)x.$

$k\lambda$ is an eigen value of kA.

Property 6 :

It λ is an eigen value of A then λ^k is an eigen value of A^k where k is any positive integer.

Proof :

Let X be an eigen vector corresponding to λ.

Then $AX = \lambda X$ -------(1)

Now, $A^2X = (AA)X = A(AX)$

$= A(\lambda X)$ (by 1)

$= \lambda(AX) = \lambda(\lambda X)$ (by 1)

$= \lambda^2 X.$

λ^2 is an eigen value of A^2.

Proceeding like this we can prove that λ^k is an eigen value of A^k for any positive integer.

Corollary : If $\lambda_1, \lambda_2, ...\lambda_n$ are eigen values of A then $\lambda_1^K, \lambda_2^K, ...\lambda_n^K$ are eigen values of A^k for any positive integer k.

Property 7:

Eigen vectors corresponding to distinct eigen values of a matrix are linearly independent.

Proof :

Let $\lambda_1\ \lambda_2 ...\lambda_k$ be distinct eigen values of a matrix and Let X_i be the eigen vector corresponding to λ_i.

Hence $AX_i = \lambda_i X_i$ (i = 1,2,..k) -------(1)

Now, Suppose $X_1\ X_2 \ldots\ X_k$ are linearly dependent. Then there exist real numbers $\alpha_1\ \alpha_2 .. \alpha_k$ not all zero, such that $\alpha_1 X_1 + \alpha_2 X_2 + \ldots + \alpha_k X_k = 0$. Among all such realations, we choose one of shortest length say j.

By rearranging the vector $X_1\ X_2 \ldots\ X_k$ we may assume that

$$\alpha_1 X_1 + \alpha_2 X_2 + \ldots \alpha_j X_j = 0 \qquad \text{-------(2)}$$

$\therefore A(\alpha_1 X_1) + A(\alpha_2 X_2) + \ldots + A(\alpha_j X_j) = 0$

$\therefore \alpha_1 (AX_1) + \alpha_2 (AX_2) + \ldots + \alpha j\ (AXj) = 0$

$$\therefore \alpha_1 \lambda_1 X_1 + \alpha_2 \lambda_2 X_2 + \ldots + \alpha_j \lambda_j X_j = 0 \qquad \text{-------(3)}$$

Multiplying (2) by λ_1 and subtracting from (3) we get

$$\alpha_2 (\lambda_1 - \lambda_2) X_2 + \alpha_3 (\lambda_1 - \lambda_3) X_3 + \ldots + \alpha_j (\lambda_1 - \lambda_j) X_j = 0 \qquad \text{-------(4)}$$

and since $\lambda_1\ \lambda_2 \ldots \lambda_j$ are distinct and $\alpha_2 \ldots \alpha_j$ are non zero,

we have $\alpha_i (\lambda_1 - \lambda_i) \neq 0$; $i = 2,3 \ldots j$.

Thus (4) gives a relation whose length is j–1, giving a contradiction.

Hence $X_1\ X_2 \ldots\ldots X_k$ all linearly Independent.

Property : 8

The characteristic roots of a Hermitian matrix are all real.

Proof :

Let A be a Hermitian matrix

Hence $A = \overline{A}^T$ (by thrm 3.17) -------(1)

Let λ be a characteristic root of A and Let X be a characteristic vector corresponding to λ:

$$AX = \lambda X \qquad \text{------(2)}$$

Now,

$$AX = \lambda X \Rightarrow \overline{X}^T AX = \lambda \overline{X}^T X$$

$$\Rightarrow (\overline{X}^T AX)^T = \overline{X}^T X \text{ (since } X^T AX \text{ in a } 1\times 1 \text{ matrix)}$$

$$\Rightarrow X^T A^T (\overline{X}^T)^T = \lambda \overline{X}^T X$$

$$\Rightarrow \quad X^T A^T \overline{X} = \lambda \overline{X}^T X$$

$$\Rightarrow \quad \overline{X^T A^T \overline{X}} = \overline{\lambda X^T X}$$

$$\Rightarrow \quad \overline{X}^T \overline{A}^T X = \overline{\lambda} X^T \overline{X}$$

$$\Rightarrow \quad \overline{X}^T A X = \overline{\lambda} X^T \overline{X} \text{ (using 1)}$$

$$\Rightarrow \quad \overline{X}^T \lambda X = \overline{\lambda} X^T \overline{X} \text{ (using 2)}$$

$$\Rightarrow \quad \lambda(\overline{X}^T X) = \overline{\lambda}(X^T \overline{X}) \qquad \text{-------(3)}$$

Now $\overline{X}^T X = X^T \overline{X} = \overline{x}_1 x_1 + \overline{x}_2 x_2 + \ldots + \overline{x}_n x_n$

$$= |x_1|^2 + |x_2|^2 + \ldots + |x_n|^2$$

$$\neq 0$$

From (3) we get $\lambda = \overline{\lambda}$

Hence λ is real.

Corollary :

The characteristic roots of a real symmetric matrix are real.

Proof :

We know that any real symmetric matrix in Hermitian. Hence the result follows from the above property

Property 9 :

The characteristic roots of a skew Hermitian matrix are either purely imaginary or zero.

Proof :

Let A be a skew Hermitian matrix and λ be a characteristic root of A.

$\therefore |A - \lambda I| = 0$

$\therefore |iA - i\lambda I| = 0$

$\therefore i\lambda$ is a characteristic root of i A.

Since A is skew Hermitian iA is Hermitian. (Refer result (iv) theorem 3.18)

$\therefore$ By theorem $i\lambda$ is real. Hence λ is purely imaginary or zero.

Corollary :

The characteristic root of a real skew symmetric matrix are either purely imaginary or zero.

Proof :

We know that any real skew symmetric matrix is skew Hermitian.

Hence the result follows from the above property.

Property 10 :

Let λ be a characteristic root of an unitary matrix A. Then $1\lambda 1 = 1$ (ie) the characteristic roots of a unitary matrix are all the unit modulus.

Proof :

Let λ be a characteristic root of an unitary matrix A and X be a characteristic vector corresponding to λ.

$$\therefore \quad AX = \lambda X \qquad \text{-------(1)}$$

Taking conjugate and transpose in (1) we get

$$(\overline{AX})^T = (\overline{\lambda X})^T$$

$$\therefore \quad \overline{X}^T \overline{A}^T = \overline{\lambda} \overline{X}^T \qquad \text{-------(2)}$$

Multiplying (1) and (2) we get

$$\left(\overline{X^T A^T}\right)(AX) = \left(\overline{\lambda} \overline{X}^T\right)(\lambda X)$$

$$\therefore \quad \overline{X}^T \left(\overline{A}^T A\right) X = \overline{\lambda}\lambda \left(\overline{X}^T X\right)$$

Now, since A in an unitary matrix $\overline{A}^T A = I$

Hence $\overline{X}^T X = (\overline{\lambda}\lambda)\overline{X}^T X$

Since X is non–zero vector $\overline{X}^T$ in also non–zero vector and $\overline{X^T} X = |x_1|^2 + |x_2|^2 + \ldots + |x_n|^2 \neq 0$.

we get $\lambda\overline{\lambda} = 1$

Hence $|\lambda|^2 = 1 \Rightarrow$ Hence $|\lambda| = 1$.

Corollary :

Let λ be a characteristic root of an orthogonal matrix A. Then $1\lambda1 = 1$.

Since any orthogonal matrix is unitary the result follows from property 10.

Property 11 :

Zero in an eigen value of A if and only if A in a singular matrix.

Proof :

The eigen values of A are the roots of the characteristic equation $|A-\lambda I| = 0$
Now, 0 is an eigen value of A $\Leftrightarrow |A- 0I| = 0$

$\Leftrightarrow |A| = 0$

$\Leftrightarrow$ A is a singular matrix.

Property 12 :

If A and B are two square matrics of the same order then AB and BA have the same eigen values.

Proof :

Let λ be an eigen value of AB and X be an eigen vector corresponding to λ.

$\therefore$ $(AB)\,X = \lambda X$

$\therefore$ $B\,(AB)\,X = B(\lambda X) = \lambda(BX)$

$\therefore$ $(BA)\,(BX) = \lambda(BX)$

$\therefore$ $(BA)\,Y = \lambda Y$ where $Y = BX$.

Hence λ is an eigen value of BA.

Also BX is the corresponding eigen vector.

Property 13 :

If P and A all n×n matrices and P is a non singular matrix then A and $P^{-1}AP$ have the same eigen values

Proof :

Let $B = P^{-1}AP$.

To prove A and B have same eigen values, it is enough to prove that the characteristic polynomials of A and B are the same.

Now $|B-\lambda I| = |P^{-1}AP-\lambda I|$

$= |P^{-1}AP-P^{-1}(\lambda I)\, P|$

$= |P^{-1}(A-\lambda I)P|$

$= |P^{-1}||A-\lambda I||P|$

$= |P^{-1}||P||A-\lambda I|$

$= |P^{-1}P||A-\lambda I|$

$= |I||A-\lambda I|$

$= |A-I\lambda|$

∴ The characteristic equation of A and $P^{-1}AP$ are the same.

Property 14 :

If λ is a characteristic root of A then $f(\lambda)$ is a characteristic root of the matrix f(A) where f(x) is any polynomial.

Proof :

Let $f(x) = a_0x^n+a_1x^{n-1}+...+a_{n-1}x+a_n$

Where $a_0 \neq 0$and $a_1a_2....a_n$ are all real numbers.

∴ $f(A) = a_0A^n + a_1A^{n-1}+...+a_{n-1}\, A+a_nI$

Since λ in a charactertic root of A, λ^n in a characteristic root of A^n for any positive Integer n (refer property 6)

∴ $A^nX = \lambda^nX$

$A^{n-1}X = \lambda^{n-1}X$

--- ---

--- ---

$AX = \lambda X$

$a_0\, A^nX = a_0\lambda^nX$

∴ $a_1A^{n-1}X = a_1\lambda^{n-1}X$

--- ---

--- ---

--- ---

$a_{n-1}\, AX = a_{n-1}\, \lambda X.$

Adding the above equations we have

$a_0A^nX + a_1A^{n-1}X + \ldots\ldots + a_{n-1}AX$

$= a_0\lambda^nX + a_1\lambda^{n-1}X + \ldots + a_{n-1}\lambda X.$

$\therefore (a_0A^n + a_1A^{n-1} + \ldots + a_{n-1}A)X = (a_0\lambda^n + a_1\lambda^{n-1} + \ldots + a_{n-1}\lambda)X$

$\therefore (a_0A^n + a_1A^{n-1} + \ldots + a_{n-1}A + a_nI)X = (a_0\lambda^n + a_1\lambda^{n-1} + \ldots + a_{n-1}\lambda + a_n)X$

$\therefore f(A)X = f(\lambda)X$

Hence $f(\lambda)$ in a characteristic root of $f(A)$.

Problem :1

If X_1 X_2 are eigen vectors corresponding to an eigen value λ then aX_1+bX_2 (a,b non–zero scalars) in also an eigen vector corresponding to λ.

Solution :

Since $X_1$and X_2 are given vectors corresponding to λ, we have

$AX_1 = \lambda X_1$ and $AX_2 = \lambda X_2$

Hence $A(aX_1) = \lambda(aX_1)$ and $A(bX_2) = \lambda(bX_2)$

$\therefore A(aX_1+bX_2) = \lambda(aX_1+bX_2)$

$\therefore aX_1+bX_2$ in an eigen vector corresponding to λ.

Problem 2 :

If the eigen values of $A = \begin{bmatrix} 3 & 10 & 5 \\ -2 & -3 & -4 \\ 3 & 5 & 7 \end{bmatrix}$ are 2,2,3 find the eigen values of A^{-1} and A^2

Solution :

Since 0 is not an eigen value of A, A is a non singular matrix and hence A^{-1} exists.

Eigen values of A^{-1} are $\frac{1}{2},\frac{1}{2},\frac{1}{3}$ and eigen values of A^2 are 2^2, 2^2, 3^2

Problem 3 :

Find the eigen values of A^5 when

$$A = \begin{bmatrix} 3 & 0 & 0 \\ 5 & 4 & 0 \\ 3 & 6 & 1 \end{bmatrix}$$

Solution :

The characteristic equation of A is obviously $(3-\lambda)(4-\lambda)(1-\lambda) = 0$

Hence the eigen values of A are 3,4,1.

∴ The eigen values of A^5 are $3^5, 4^5, 1^5$.

Problem 4 :

Find the sum and product of the eigen values of the matrix

$\begin{bmatrix} 3 & -4 & 4 \\ 1 & -2 & 4 \\ 1 & -1 & 3 \end{bmatrix}$ without actually finding the eigen values.

Solution :

Let $A = \begin{bmatrix} 3 & -4 & 4 \\ 1 & -2 & 4 \\ 1 & -1 & 3 \end{bmatrix}$

Sum of the eigen values = trace of A = 3+ (–2) + 3 = 4 product of the eigen values = |A|

Now, $$|A| = \begin{bmatrix} 3 & -4 & 4 \\ 1 & -2 & 4 \\ 1 & -1 & 3 \end{bmatrix}$$

$$= 3(-6+4) + 4(3-4) - 4(-1+2)$$

$$= -6-4-4 = -14$$

∴ Product of the eigen values = –14

Problem 5 :

Find the characteristic roots of the matrix $\begin{pmatrix} \cos\theta & -\sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$

Solution :

Let $A = \begin{pmatrix} \cos\theta & -\sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$

The characteristic equation of A in given by $|A-\lambda I| = 0$.

$$\begin{vmatrix} \cos\theta-\lambda & -\sin\theta \\ -\sin\theta & \cos\theta-\lambda \end{vmatrix} = 0$$

$\therefore$ $(\cos\theta-\lambda)^2 - \sin^2\theta = 0$

$\therefore$ $(\cos\theta-\lambda-\sin\theta)\,(\cos\theta-\lambda+\sin\theta) = 0$

$\therefore$ $[\lambda-(\cos\theta-\sin\theta)]\,[\lambda-(\cos\theta+\sin\theta)] = 0$

$\therefore$ The two characteristic roots of the matrix are $(\cos\theta-\sin\theta)$ and $(\cos\theta+\sin\theta)$

Problem 6 :

Find the characteristic roots of the matrix

$$A = \begin{pmatrix} \cos\theta & -\sin\theta \\ -\sin\theta & -\cos\theta \end{pmatrix}$$

Solution :

The characteristic equation of A in given by $|A-\lambda I| = 0$.

(ie) $$\begin{vmatrix} \cos\theta-\lambda & -\sin\theta \\ -\sin\theta & -\cos\theta-\lambda \end{vmatrix} = 0$$

$\therefore$ $-(\cos^2\theta-\lambda^2) - \sin^2\theta = 0$

$\therefore$ $\lambda^2-(\cos^2\theta + \sin^2\theta) = 0$

$\therefore$ $\lambda^2 - 1 = 0$

$\therefore$ The characteristic roots are 1 and –1

Problem 7 :

Find the sum and product of the eigen values of the matrix.

$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ without finding the roots of the characteristic equation.

Solution :

Sum of the eigen values of

$$A = \text{trace of } A = a_{11}+a_{22}$$

product of the eigen values of

$$A = |A| = a_{11}\, a_{22} - a_{12}\, a_{21}$$

Problem 8 :

Verify the statement that the sum of the elements in the diagonal of a matrix is the sum of the eigen values of the matrix.

$$A = \begin{bmatrix} -2 & 2 & -3 \\ 2 & 1 & -6 \\ -1 & -2 & 0 \end{bmatrix}$$

Solution :

The characteristic equation of A is $|A-\lambda I| = 0$

$$\text{(ie)} \quad \begin{bmatrix} -2-\lambda & 2 & -3 \\ 2 & 1-\lambda & -6 \\ -1 & -2 & -\lambda \end{bmatrix} = 0$$

(ie) $(-2-\lambda)[(1-\lambda)(-\lambda)-12]-2[-2\lambda-6]-3[-4+(1-\lambda)] = 0$

(ie) $(-2-\lambda)\,(\lambda^2-\lambda-12) + 4(\lambda+3) + 3(\lambda+3) = 0$

(ie) $-2\lambda^2+2\lambda+24-\lambda^3+\lambda^2+12\lambda+4\lambda+12+3\lambda+9 = 0$

(ie) $-\lambda^3-\lambda^2+21\lambda+45 = 0$

(ie) $\lambda^3+\lambda^2-21\lambda-45 = 0.$

This is a cubic equation in λ and hence it has 3 roots and the three roots are the three eigen values of the matrix.

$$\text{The sum of the eigen values} = -\left(\frac{\text{coefficient of } \lambda^2}{\text{coefficient of } \lambda^3}\right) = -1$$

The sum of the elements on the diagonal of the matrix

$$A = -2+1+0 = -1$$

Hence the result.

Problem 9 :

The product of two eigen values of the matrix $A = \begin{pmatrix} 6 & -2 & 2 \\ -2 & 3 & -1 \\ 2 & -1 & 3 \end{pmatrix}$ is 16. Find the third eigen value. What is the sum of the eigen values of A ?

Solution :

Let $\lambda_1\ \lambda_2\ \lambda_3$ be the eigen values of A. Given, product of 2 eigen values (say) $\lambda_1\ \lambda_2$ is 16.

$\therefore\ \lambda_1\ \lambda_2 = 16.$

We know that the product of the eigen value is |A|

$$\text{(ie)} \qquad \lambda_1\ \lambda_2\ \lambda_3 = \begin{pmatrix} 6 & -2 & 2 \\ -2 & 3 & -1 \\ 2 & -1 & 3 \end{pmatrix}$$

$$\text{(ie)} \qquad 16\lambda_3 = 6(9-1) + 2(-6+2) + 2(2-6)$$

$$= 48-8-8$$

$$= 32$$

$$\therefore \qquad \lambda_3 = 2$$

$\therefore$ The third eigen value is 2.

Also we know that the sum of the eigen values of

$$A = \text{trace of } A = 6+3+3 = 12$$

Problem 10 :

The product of two eigen values of the matrix

$A = \begin{pmatrix} 2 & 2 & -7 \\ 2 & 1 & 2 \\ 0 & 1 & -3 \end{pmatrix}$ is – 12. Find the eigen values of A.

Solution :

Let $\lambda_1, \lambda_2, \lambda_3$ be the eigen values of A. Given product of 2 eigen values, say λ_1 and λ_2 is –12.

$\therefore \quad \lambda_1 \lambda_2 = -12$ -------(1)

We know that the product of the eigen values in |A|.

$$\therefore \quad \lambda_1 \lambda_2 \lambda_3 = \begin{vmatrix} 2 & 2 & -7 \\ 2 & 1 & 2 \\ 0 & 1 & -3 \end{vmatrix}$$

(ie) $12 \lambda_3 = -12$

$\therefore \quad \lambda_3 = 1$ -------(2)

Also we know sum of the eigen values = Trace of A.

$\therefore \quad \lambda_1+\lambda_2+\lambda_3 = 2+1-3 = 0$

$\therefore \quad \lambda_1+\lambda_2 = -1$ (using 2) -------(3)

using (3) in (1) we get

$$\lambda_1(-1-\lambda_1) = -12$$

$$\lambda_1^2+\lambda_1-12 = 0$$

$$(\lambda_1+4)(\lambda_1-3) = 0$$

$\therefore \quad \lambda_1 = 3$ or -4

Putting $\lambda_1 = 3$ in (1) we get $\lambda_2 = -4$ or putting $\lambda_1 = -4$ in (1) we get $\lambda_2 = 3$

Thus the three eigen values are 3, –4, 1.

Problem 11 :

Find the sum of the squares of the eigen values of

$$A = \begin{pmatrix} 3 & 1 & 4 \\ 0 & 2 & 6 \\ 0 & 0 & 5 \end{pmatrix}$$

Solution :

Let $\lambda_1, \lambda_2, \lambda_3$ be the eigen values of A. We know that $\lambda_1^2, \lambda_2^2, \lambda_3^2$ are the eigen values of A^2.

$$\therefore \quad A^2 = \begin{pmatrix} 3 & 1 & 4 \\ 0 & 2 & 6 \\ 0 & 0 & 5 \end{pmatrix} \begin{pmatrix} 3 & 1 & 4 \\ 0 & 2 & 6 \\ 0 & 0 & 5 \end{pmatrix}$$

$$= \begin{pmatrix} 9 & 5 & 38 \\ 0 & 4 & 42 \\ 0 & 0 & 25 \end{pmatrix}$$

$\therefore$ Sum of the eigen values of A^2 = Trace of A^2 = 9+4+25

(ie) $\lambda_1^2+\lambda_2^2+\lambda_3^2 = 38$.

$\therefore$ Sum of the squares of the eigen values of A = 38.

Problem 12 :

Find the eigen values and eigen vectors of the matrix.

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

Solution :

The characteristic equation of A in $|A-\lambda I| = 0$.

$$\therefore \begin{vmatrix} 1-\lambda & 1 & 3 \\ 1 & 5-\lambda & 1 \\ 3 & 1 & 1-\lambda \end{vmatrix} = 0$$

$\therefore (1-\lambda)[(5-\lambda)(1-\lambda)-1]-[(1-\lambda)-3]+3[1-3(5-\lambda)] = 0.$

$(1-\lambda)(\lambda^2-6\lambda+4)+(\lambda+2)+3(3\lambda-14) = 0$

$\lambda^2-6\lambda+4-\lambda^3+6\lambda^2-4\lambda+\lambda+2+9\lambda-42 = 0$

$\therefore -\lambda^3+7\lambda^2-36 = 0$ Hence $\lambda^3-7\lambda^2+36 = 0$

$\therefore (\lambda+2)(\lambda^2-9\lambda+18) = 0$

Hence $(\lambda+2)(\lambda-6)(\lambda-3) = 0$

$\therefore \lambda = -2,3,6$ are the three eigen values.

Case (i) :

Eigen vector corresponding to $\lambda = -2$.

Let $X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ be an eigen vector corresponding to $\lambda = -2$.

Hence $AX = -2X.$

(ie) $$\begin{bmatrix} 1 & 1 & 3 \\ 1 & 5 & 1 \\ 3 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -2x_1 \\ -2x_2 \\ -2x_3 \end{bmatrix}$$

$\therefore$ $$x_1+x_2+3x_3 = -2x_1$$
$$x_1+5x_2+x_3 = -2x_2$$
$$3x_1+x_2+x_3 = -2x_3$$

$\therefore$ $$3x_1+x_2+3x_3 = 0 \quad \text{-------(1)}$$
$$x_1+7x_2+x_3 = 0 \quad \text{-------(2)}$$
$$3x_1+x_2+3x_2 = 0 \quad \text{-------(3)}$$

Clearly this system of three equation reduces to two equations only. From (1) & (2)

We get $x_1 = -2k$; $x_2 = 0$ $x_3 = 2k$.

$\therefore$ It has only one independent solution and can be obtained by giving any value to k say k = 1.

$\therefore$ (–2,0,2) in an eigen vector corresponding to $\lambda = -2$.

Case (ii)

Eigen vector corresponding to $\lambda = 3$

Then $AX = 3X$ gives

$$-2x_1 + x_2 + 3x_3 = 0$$
$$x_1+2x_2+x_3 = 0$$
$$3x_1 + x_2 - 2x_3 = 0$$

Taking the first 2 equations we get

$$\frac{x_1}{-5} = \frac{x_2}{5} = \frac{x_3}{-5} = k \text{ (say)}$$

$\therefore$ $x_1 = -k$; $x_2 = k$; $x_3 = -k$.

Taking k = 1 (say) (–1,1,–1) is an eigen vector corresponding to $\lambda = 3$

Case (iii)

Eigen vector corresponding to $\lambda = 6$

we have $AX = 6X.$

Hence $-5x_1+x_2+3x_3 = 0$

$x_1-x_2+x_3 = 0$

$3x_1+x_2-5x_3 = 0$

Taking the first two equation we get

$$\frac{x_1}{4} = \frac{x_2}{8} = \frac{x_3}{4} = k.$$

$x_1 = k$; $x_2 = 2k$ $x_3 = k$. It satisfies the third equation also.

Taking $k = 1$ (say) (1,2,1) is an eigen vector corresponding to $\lambda = 6$.

Problem 13 :

Find the eigen values and eigen vectors of the matrix.

$$A = \begin{bmatrix} 6 & -2 & 2 \\ -2 & 3 & -1 \\ 2 & -1 & 3 \end{bmatrix}$$

Solution :

The characterstic equation of A is $|A-\lambda I| = 0$

$$\begin{vmatrix} 6-\lambda & -2 & 2 \\ -2 & 3-\lambda & -1 \\ 2 & -1 & 3-\lambda \end{vmatrix} = 0$$

$(6-\lambda)\,[(3-\lambda)^2-1] +2\,[(2\lambda-6)+2] +2\,(2-6+2\lambda) = 0$

$(6-\lambda)\,(8+\lambda^2-6\lambda) +4\lambda -8+4\lambda-8 = 0$

$48+6\lambda^2-36\lambda-8\lambda-\lambda^3+6\lambda^2+8\lambda-16 = 0$

$-\lambda^3+12\lambda^2-36\lambda+32 = 0$

Hence $\lambda^3-12\lambda^2+36\lambda-32 = 0$

$\therefore (\lambda-2)\,(\lambda-2)\,(\lambda-8) = 0.$

$\therefore$ The eigen values are 2,2,8. We now find the eigen vectors.

Case (i) :

$\lambda = 2$.

The eigen vector $X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ is got from $AX = 2X$

$$\therefore \quad 6x_1 - 2x_2 + 2x_3 = 2x_1$$

$$-2x_1 + 3x_2 - x_3 = 2x_2$$

$$2x_1 - x_2 + 3x_3 = 2x_3$$

$$\therefore \quad 4x_1 - 2x_2 + 2x_3 = 0$$

$$-2x_1 + x_2 - x_3 = 0$$

$$2x_1 - x_2 + x_3 = 0$$

The above three equations are equivalent to the single equation $2x_1 - x_2 + x_3 = 0$

The independent eigen vectors can be obtained by giving arbitrary values to any two of the unknowns $x_1\ x_2\ x_3$.

Giving $x_1 = 1$; $x_2 = 2$ we get $x_3 = 0$

Giving $x_1 = 3$; $x_2 = 4$ we get $x_3 = -2$

$\therefore$ The two Independent vectors corresponding to $\lambda = 2$ are (1,2,0) and (3,4,–2).

Case (ii)

$\lambda = 8$.

The eigen vector $X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ is got from

$$AX = 8X.$$

$$\therefore \quad -2x_1 - 2x_2 + 2x_3 = 0 \qquad \text{-------(1)}$$

$$-2x_1 - 5x_2 - x_3 = 0 \qquad \text{-------(2)}$$

$$2x_1 - x_2 - 5x_3 = 0 \qquad \text{-------(3)}$$

From (1) and (2) we get

$$\frac{x_1}{12} = \frac{x_2}{-6} = \frac{x_3}{6} = k \text{ (say)}$$

$\therefore x_1 = 2k\ ;\ x_2 = -k\ ;\ x_3 = k.$

Giving k = 1 we get an eigen vector corresponding to 8 as (2,–1,1).

Problem 14 :

Find the eigen values and eigen vectors of the matrix

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

Solution :

The characteristic equation of A in $|A-\lambda I| = 0$

$$\text{(ie)} \begin{bmatrix} 2-\lambda & -2 & 2 \\ 1 & 1-\lambda & 1 \\ 1 & 3 & -1-\lambda \end{bmatrix} = 0$$

$\therefore (2-\lambda)\,[-(1-\lambda)\,(1+\lambda)-3]+2[-(1+\lambda)-1] +2\,[3-(1-\lambda) = 0$

$\therefore (2-\lambda)\,(\lambda^2-4)-2\,(2+\lambda) +2\,(2+\lambda) = 0$

$\therefore 2\lambda^2-8-\lambda^3+4\lambda-4-2\lambda+4+2\lambda = 0$

$\therefore -\lambda^3+2\lambda^2+4\lambda-8 = 0$

Hence $\lambda^3-2\lambda^2-4\lambda+8 = 0$

$\therefore (\lambda-2)(\lambda^2-4) = 0$

Hence $(\lambda-2)\,(\lambda-2)\,(\lambda+2) = 0$

$\lambda = 2,2,-2$ are three eigen values.

Case (i)

$\lambda = 2.$

Let $X = (x_1\ x_2\ x_3)$ be an eigen vector corresponding to $\lambda = 2$, x is got from AX=2X.

(ie) $\begin{bmatrix} 2 & -2 & 2 \\ 1 & 1 & 1 \\ 1 & 3 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2x_1 \\ 2x_2 \\ 2x_3 \end{bmatrix}$

∴ The eigen vector corresponding to $\lambda = 2$ is given by the equations.

$$2x_1-2x_2+2x_3 = 2x_1$$
$$x_1+x_2+x_3 = 2x_2$$
$$x_1+3x_2-x_3 = 2x_3$$

(ie)
$$-x_2+x_3 = 0 \quad \text{-------(1)}$$
$$x_1-x_2+x_3 = 0 \quad \text{-------(2)}$$
$$x_1+3x_2-3x_3 = 0 \quad \text{-------(3)}$$

Taking (1) and (2) we get

$$\frac{x_1}{0} = \frac{x_2}{1} = \frac{x_3}{1} = k \text{ (say)}$$

∴ $x_1 = 0\ x_2 = k\ x_3 = k$.

Taking $k = 1$ we get (0,1,1) as an eigen vector corresponding to $\lambda = 2$.

Case (ii)

$\lambda = -2$.

Corresponding to $\lambda = -2$ we have $AX = -2X$.

∴
$$2x_1-2x_2+2x_3 = -2x_1$$
$$x_1+x_2+x_3 = -2x_2$$
$$x_1+3x_2-x_3 = -2x_3$$

∴
$$2x_1-x_2+x_3 = 0$$
$$x_1+3x_2+x_3 = 0$$
$$x_1+3x_2+x_3 = 0$$

∴ Taking the first two equation we get

$$\frac{x_1}{-4} = \frac{x_2}{-1} = \frac{x_3}{7} = k \text{ (say)}$$

∴ $x_1 = -4k$; $x_2 = -k$; $x_3 = 7k$.

Taking k = 1 we get (–4,–1,7) as an eigen vector corresponding to the eigen value $\lambda = -2$.

Hence the problem.

15. Find the eigen values and the eigen vectors of the matrix $\begin{bmatrix} 8 & -6 & 2 \\ -6 & 7 & -4 \\ 2 & -4 & 3 \end{bmatrix}$

Solution :

Let $A = \begin{pmatrix} 8 & -6 & 2 \\ -6 & 7 & -4 \\ 2 & -4 & 3 \end{pmatrix}$

The characteristic equation of the matrix A is

$$\begin{vmatrix} 8-\lambda & -6 & 2 \\ -6 & 7-\lambda & -4 \\ 2 & -4 & 3-\lambda \end{vmatrix} = 0$$

$(8-\lambda)[(7-\lambda)(3-\lambda)-16] +6[-6(3-\lambda)+8] +2[24-2(7-\lambda)] = 0$

$(8-\lambda)[\lambda^2-10\lambda+21-16] +6[-18+6\lambda+8] +2[24-14+2\lambda] = 0$

$(8-\lambda)[\lambda^2-10x+5]+6[6\lambda-10]+2[2\lambda+10] = 0$

$-\lambda^3+18\lambda^2-85\lambda+40+36\lambda-60+4\lambda+20 = 0$

$-\lambda^3+18\lambda^2-45\lambda = 0$

$\lambda^3-18\lambda^2+45\lambda = 0$

$\lambda(\lambda^2-18\lambda+45) = 0$

$\lambda(\lambda-3)(\lambda-15) = 0$

$\lambda = 0, 3, 15.$

Hence the eigen values of the given matrix are (0,3,15).

For the given matrix, the equation $(A-\lambda I)X = 0$

is $\begin{vmatrix} 8-\lambda & -6 & 2 \\ -6 & 7-\lambda & -4 \\ 2 & -4 & 3-\lambda \end{vmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0$

Hence the eigen vectors x is given by the equations

$$\left.\begin{aligned}(8-\lambda)x_1-6x_2+2x_3 &= 0\\ -6x_1+(7-\lambda)x_2-4x_3 &= 0\\ 2x_1-4x_2+(3-\lambda)x_3 &= 0\end{aligned}\right\} \qquad \text{-------(I)}$$

Case (i) When $\lambda = 0$, the equations (I) are becomes,

$$8x_1-6x_2+2x_3 = 0 \qquad \text{-------(1)}$$

$$-6x_1+7x_2-4x_3 = 0 \qquad \text{-------(2)}$$

$$2x_1-4x_2+3x_3 = 0 \qquad \text{-------(3)}$$

Since the equations are linearly dependent, we can omit one of them.

From (2) & (3), we have

$$\frac{x_1}{21-16} = \frac{x_2}{-8+18} = \frac{x_3}{24-14}$$

$$\Rightarrow \frac{x_1}{5} = \frac{x_2}{10} = \frac{x_3}{10}$$

(or) $\frac{x_1}{1} = \frac{x_2}{2} = \frac{x_3}{2}$

Hence the corresponding eigen vector is

$$X_1 = \begin{bmatrix}1\\2\\2\end{bmatrix}.$$

Also every non–zero multiple of this column vector is an eigen vector corresponding to $\lambda = 0$.

Case (ii)

When $\lambda = 3$, the eigen vector is given by the equations

$$5x_1-6x_2+2x_3 = 0 \qquad \text{-------(4)}$$

$$-6x_1+4x_2-4x_3 = 0 \qquad \text{-------(5)}$$

$$2x_1-4x_2+0.x_3 = 0 \qquad \text{-------(6)}$$

From equations (5) and (6) we have

$$\frac{x_1}{0-16} = \frac{x_2}{-8+0} = \frac{x_3}{24-8}$$

$$\frac{x_1}{-16} = \frac{x_2}{-8} = \frac{x_3}{16}$$

(or) $$\frac{x_1}{2} = \frac{x_2}{1} = \frac{x_3}{-2}$$

Hence the corresponding eigen vector is $x_2 = \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix}$

Case (iii)

When $\lambda = 15$, the eigen vector is given by the equations,

$$-7x_1 - 6x_2 + 2x_3 = 0 \qquad \text{-------(7)}$$

$$-6x_1 - 8x_2 - 4x_3 = 0 \qquad \text{-------(8)}$$

$$2x_1 - 4x_2 - 12x_3 = 0 \qquad \text{-------(9)}$$

From the equations (8) and (9), we have

$$\frac{x_1}{96-16} = \frac{x_2}{-8-72} = \frac{x_3}{24+16}$$

$$\frac{x_1}{80} = \frac{x_2}{-80} = \frac{x_3}{40}$$

(ie) $$\frac{x_1}{2} = \frac{x_2}{-2} = \frac{x_3}{1}.$$

Hence the corresponding eigen vector is $x_3 = \begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix}$

The three vectors [1, 2, 2], [2,1,–2] and [2,–2,1] are linearly independent.

16. Find the eigen values and the eigen vetors of

$$\begin{bmatrix} 3 & 10 & 5 \\ -2 & -3 & -4 \\ 3 & 5 & 7 \end{bmatrix}$$

Solution :

Let $$A = \begin{bmatrix} 3 & 10 & 5 \\ -2 & -3 & -4 \\ 3 & 5 & 7 \end{bmatrix}$$

The characteristic equation of the given matrix is

$$\begin{vmatrix} 3-\lambda & 10 & 5 \\ -2 & -3-\lambda & -4 \\ 3 & 5 & 7-\lambda \end{vmatrix} = 0$$

$(3-\lambda)\,[-(3+\lambda)\,(7-\lambda)+20] - 10\,[-2(7-\lambda)+12] + 5\,[-10+3(3+\lambda)] = 0$

$(3-\lambda)\,(\lambda^2-4\lambda-1) - 10\,(2\lambda-2) + 5\,(3\lambda-1) = 0$

$-\lambda^3+7\lambda^2-16\lambda+12 = 0$

$\lambda^3-7\lambda^2+16\lambda-12 = 0$

$\lambda^2(\lambda-2)-5\lambda(\lambda-2)+6\,(\lambda-2) = 0$

$(\lambda-2)\,[\lambda^2-5\lambda+6] = 0$

$(\lambda-2)\,(\lambda-2)\,(\lambda-3) = 0$

$\lambda = 2,2,3.$

(ie) the eigen values of the given matrix are (2,2,3)

Case (i)

Taking $\lambda = 3$, the corresponding eigen vector is given by the equations.

$$0x_1+10x_2+5x_3 = 0 \qquad \text{-------(1)}$$

$$-2x_1-6x_2-4x_3 = 0 \qquad \text{-------(2)}$$

$$3x_1+5x_2+4x_3 = 0 \qquad \text{-------(3)}$$

From equations (2) & (3), we have

$$\frac{x_1}{-24+20} = \frac{x_2}{-12+8} = \frac{x_3}{-10+18}$$

(ie) $$\frac{x_1}{-4} = \frac{x_2}{-4} = \frac{x_3}{8}$$

(ie) $$\frac{x_1}{1} = \frac{x_2}{1} = \frac{x_3}{-2}$$

Hence the corresponding eigen vector is

$$x_1 = \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}$$

Case (ii)

Let $\lambda = 2$ (the equal root), the components of the eigen vector are given by the equations

$$x_1+10x_2+5x_3 = 0 \qquad \text{-------(4)}$$

$$-2x_1-5x_2-4x_3 = 0 \qquad \text{-------(5)}$$

$$3x_1+5x_2+5x_3 = 0 \qquad \text{-------(6)}$$

From equations (5) and (6), we have

$$\frac{x_1}{-25+20} = \frac{x_2}{-12+10} = \frac{x_3}{-10+15}$$

$$\frac{x_1}{-5} = \frac{x_2}{-2} = \frac{x_3}{5}$$

(ie)
$$\frac{x_1}{5} = \frac{x_2}{2} = \frac{x_3}{-5}$$

Hence
$$X_2 = \begin{bmatrix} 5 \\ 2 \\ -5 \end{bmatrix}$$

The eigen vectors X_3 has also to be the same form as X_2 and hence it is linearly dependent on X_2.

Exercises :

1. For each of the following matrices find the characteristic vectors corresponding to each characteristic root.

(a) $\begin{bmatrix} 8 & -6 & 2 \\ -6 & 7 & -4 \\ 2 & -4 & 3 \end{bmatrix}$ (b) $\begin{bmatrix} 2 & 2 & 1 \\ 1 & 3 & 1 \\ 1 & 2 & 2 \end{bmatrix}$

2. For what value to k is 3 a characteristic root of

$$\begin{bmatrix} 3 & 1 & -1 \\ 3 & 5 & -k \\ 3 & k & -1 \end{bmatrix}$$

3. Find the characteristic equation of the matrix

$A = \begin{bmatrix} 2 & 3 & 1 \\ 3 & 1 & 2 \\ 1 & 2 & 3 \end{bmatrix}$ and prove that the matrix

$B = \begin{bmatrix} 3 & 1 & 2 \\ 1 & 2 & 3 \\ 2 & 3 & 1 \end{bmatrix}$ has the same characteristic equation.

4. Find the characteristic roots and the corresponding characteristic vectors of

A^3+A^2+A+I if $A = \begin{bmatrix} 1 & -1 & -1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix}$

5. Show that the matrix $\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$ has the two eigen vectors $\begin{pmatrix} 1 \\ i \end{pmatrix}$ and $\begin{pmatrix} 1 \\ -i \end{pmatrix}$

6. Find the eigen values and eigen vectors of the following matrices.

(i) $\begin{bmatrix} 3 & -4 & 4 \\ 1 & -2 & 4 \\ 4 & -1 & 3 \end{bmatrix}$ (ii) $\begin{pmatrix} 3 & 1 & 4 \\ 0 & 2 & 6 \\ 0 & 0 & 5 \end{pmatrix}$ (iii) $\begin{bmatrix} 1 & 1 & 3 \\ 1 & 5 & 1 \\ 3 & 1 & 1 \end{bmatrix}$ (iv) $\begin{pmatrix} 6 & -2 & 2 \\ -2 & 3 & -1 \\ 2 & -1 & 3 \end{pmatrix}$ (v) $\begin{pmatrix} 3 & -1 & 3 \\ 9 & -1 & 9 \\ 7 & -1 & 7 \end{pmatrix}$

5.3. RANK OF A MATRIX

We now proceed to introduce the concept of the rank of a matrix.

Definition :

Let $A = (a_{ij})$ be an mxn matrix. The rows $Ri = (a_{i1}, a_{i2}, \ldots a_{in})$ of A can be thought of as elements of F^n. The subspace of F^n generated by the m rows of A is called the **row space** of A.

Similarly, the subspace of F^m generated by the n columns of A is called the **column space** of A.

The dimension of the row space (column space) of A is called the **row rank (column rank)** of A.

Theorem 5.2 :

Any two row equivalent matrices have the same row space and have the same row rank.

Proof :

Let A be an mxn matrix.

It is enough if we prove that the row space of A is not altered by any elementary row operation. Obviously the row space of A is not altered by an elementary row operation of the type $R_i \leftrightarrow R_j$ Now, consider the elementary row operation.

$R_i \rightarrow cR_i$ where $c \in F-\{0\}$.

Since $L(\{R_1,R_2,..R_i,...,R_n\}) = L(\{R_1,R_2,....cR_i,...R_n\})$ the row space of A is not altered by this type of elementary row operation.

Similarly we can earily prove that the row space of A is not altered by an elementary row operation of the type $R_i \rightarrow R_i+cR_i$.

Hence row equivalent matrices have the same row space and hence the same row rank.

Theorem 5.3 :

Any two column equivalent matrices have the same column rank.

Proof :

Let A be an mxn matrix.

It is enough if we prove that the column space of A is not altered by any elementary column operation.

Obviously the column space of A is not altered by an elementary column operation of the type $C_i \leftrightarrow C_j$

Now, consider the elementary column operation. $C_i \rightarrow rC_i$ where $r \in F - \{0\}$.

Since $L(\{C_1,C_2,....C_i,...C_m\}) = L(\{C_1,C_2,...rC_i,...C_m\})$

The column space of A is not altered by this type of elementary row operation.

Similarly we can easily prove that the column space of A is not altered by an elementary column operation of the type $C_i \rightarrow C_i+rC_i$

Hence column equivalent matrices have the same column space and hence the same column rank.

Theorem 5.4 :

The row rank and the column rank of any matrix are equal.

Proof :

Let $A = (a_{ij})$ be an mxn matrix.

Let $R_1, R_2,\ldots,R_m$ denote the rows of A.

Hence $R_i = (a_{i1},a_{i2},\ldots.a_{in})$

Suppose the row rank of A is r.

Then the dimension of the row space is r.

Let $v_1 = (b_{11},\ldots..,b_{1n})$, $v_2 = (b_{21},b_{22},\ldots..b_{2n}),\ldots..$

$v_r = (b_{r1},\ldots.b_{rn})$ be a basis for the rowspace of A.

Then each row is a linear combination of the vectors $v_1,v_2,\ldots.v_r$.

Let

$$R_1 = k_{11}v_1 + k_{12}v_2+\ldots\ldots+k_{ir}v_r$$

$$R_2 = k_{21}v_1+k_{22}v_2+\ldots.+k_{2r}v_r$$

$$R_m = k_{m1}v_1+k_{m2}v_2+\ldots k_{mr}v_r$$

Where $k_{ij}\in F$

Equating the ith component of each of the above equations, we get

$$a_{1i} = k_{11}b_{1i} + k_{12}b_{2i} +\ldots..+k_{1r}b_{ri}$$

$$a_{2i} = k_{21}b_{1i} + k_{22}b_{2i} +\ldots..+k_{2r}b_{ri}$$

$$a_{mi} = k_{m1}b_{1i}+k_{m2}b_{2i}+\ldots\ldots\ldots+k_{mr}b_{ri}$$

$$\begin{bmatrix} a_{1i} \\ . \\ . \\ a_{mi} \end{bmatrix} = b_{1i}\begin{bmatrix} k_{11} \\ . \\ . \\ k_{m1} \end{bmatrix} + b_{2i}\begin{bmatrix} k_{12} \\ . \\ . \\ k_{m2} \end{bmatrix} +\ldots\ldots+b_{ri}\begin{bmatrix} k_{1r} \\ . \\ . \\ k_{mr} \end{bmatrix}$$

Thus each column of A is a linear combination of r vectors.

Hence the dimension of the column space $\leq$ r.

$\therefore$ Column rank of $A \leq r$ = row rank of A.

Similarly, row rank of $A \leq$ column rank of A.

Hence the row rank and the column rank of A are equal.

Definition :

The **rank** of a matrix A is the common value of its row and column rank.

In other words, the rank of a matrix is the largest order of any non–vanishing minor of the matrix.

Note 1 : Since the row rank and column rank of a matrix are unaltered by elementary row and column operations, equivalent matrices have the same rank.

In particular if a matrix A is reduced to its canonical form, $\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$ then rank of $A = r$.

Thus to find the rank of a matrix A, we reduce A to the canonical form and find the number of non–zero entries in the diagonal.

Note that in the canonical form of the matrix A, there exists an r×r sub–matrix, namely Ir, whose determinant is not zero.

Further every (r+1) × (r+1) sub–matrix contains a row of zeros and hence its determinant is zero.

Also under any elementary row or column operation the value of a determinant is either unaltered or multiplied by a non–zero constant.

Hence the matrix A is also such that

(i) there exists an r×r sub–matrix whose determinant is non zero.

(ii) The determinant of every (r+1)×(r+1) sub–matrix is zero.

Hence one can also define the rank of a matrix A to be r if A satisfies (i) and (ii)

Note 2 : Any non–singular matrix of order n is equivalent to the identity matrix and hence its rank is n.

Note 3 : The rank of a matrix is not altered on multiplication by non–singular matrices, since premultiplication by a non–singular matrix is equivalent to applying elementary row operations and post–multiplication by a non–singular matrix is equivalent to applying elementary column operations.

Solved problems :

Problem 1 :

1. Find the rank of the matrix $A = \begin{bmatrix} 4 & 2 & 1 & 3 \\ 6 & 3 & 4 & 7 \\ 2 & 1 & 0 & 7 \end{bmatrix}$

Solution :

$$A = \begin{bmatrix} 4 & 2 & 1 & 3 \\ 6 & 3 & 4 & 7 \\ 2 & 1 & 0 & 7 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 2 & 4 & 3 \\ 4 & 3 & 6 & 7 \\ 0 & 1 & 2 & 7 \end{bmatrix} C_1 \leftrightarrow C_3$$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 4 & -5 & -10 & -5 \\ 0 & 1 & 2 & 7 \end{bmatrix} \begin{matrix} C_1 \to C_2 - 2C_1 \\ C_3 \to C_3 - 4C_1 \\ C_4 \to C_4 - 3C_1 \end{matrix}$$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -5 & -10 & -5 \\ 0 & 1 & 2 & 7 \end{bmatrix} R_2 \to R_2 - 4R_1$$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -5 & 0 & 0 \\ 0 & -1 & 0 & 6 \end{bmatrix} \begin{matrix} C_3 \to C_3 - 2C_2 \\ C_4 \to C_4 - C_2 \end{matrix}$$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -5 & 0 & 0 \\ 0 & 0 & 0 & 6 \end{bmatrix} R_3 \to R_3 + \frac{1}{5} R_2$$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -5 & 0 & 0 \\ 0 & 0 & 6 & 0 \end{bmatrix} C_2 \leftrightarrow C_3$$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{matrix} R_2 \to \frac{-1}{5} R_2 \\ R_3 \to \frac{1}{6} R_3 \end{matrix}$$

∴ Rank of A = 3.

2. Find the rank of the matrix $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ -2 & -3 & -1 \end{bmatrix}$

Solution :

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

$$\sim \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 0 & 0 & 0 \end{bmatrix} R_3 \rightarrow R_2 + R_3$$

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

$\therefore$ Rank of $A \neq 3$.

But there is atleast one non-zero minor of order 2, namely $\begin{vmatrix} 1 & 2 \\ 2 & 3 \end{vmatrix}$ which is $= -1$.

Hence Rank of $A = 2$.

3. Find the rank of the matrix $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{bmatrix}$

Solution :

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

$$|A| = \begin{vmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{vmatrix}$$

$$= 1(6-1)-2(4-3)+3(2-9)$$

$$= 5-2-21 = -18 \neq 0$$

$\therefore$ Rank of $A = 3$

4. Find the rank of matrix $\begin{bmatrix} 1 & 2 & 1 & 2 \\ 1 & 3 & 2 & 2 \\ 2 & 4 & 3 & 4 \\ 3 & 7 & 4 & 6 \end{bmatrix}$

Solution :

Let A be the given matrix.

$$A \sim \begin{bmatrix} 1 & 2 & 1 & 2 \\ 1 & 3 & 2 & 2 \\ 2 & 4 & 3 & 4 \\ 0 & 1 & 1 & 0 \end{bmatrix} R_4 \to R_4 - 3R_1$$

$$\sim \begin{bmatrix} 1 & 2 & 1 & 2 \\ 1 & 3 & 2 & 2 \\ 0 & -1 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix} R_3 \to R_3 - (R_1 + R_2)$$

$$\sim \begin{bmatrix} 1 & 2 & 1 & 2 \\ 0 & 1 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix} R_2 \to R_2 - R_1$$

In this final form of A, the fourth order determinant

$$= \begin{vmatrix} 1 & 1 & 0 \\ -1 & 0 & 0 \\ 1 & 1 & 0 \end{vmatrix} = 0$$

The minor of order 3 namely

$$\begin{vmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & -1 & 0 \end{vmatrix} = 1(1) = 1 \neq 0$$

$\therefore$ Rank of A = 3.

5. Find the rank of the matrix.

$$A = \begin{bmatrix} 1 & -7 & 3 & -3 \\ 7 & 20 & -2 & 25 \\ 5 & -2 & 4 & 7 \end{bmatrix}$$

Solution :

$$A \sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 7 & 69 & -23 & 46 \\ 5 & 33 & -11 & 22 \end{bmatrix} \begin{matrix} C_2 \to C_2 + 7C_1 \\ C_3 \to C_3 - 3C_1 \\ C_4 \to C_4 + 3C_2 \end{matrix}$$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 7 & 23 & -23 & 23 \\ 5 & 11 & -11 & 11 \end{bmatrix} \begin{matrix} C_2 \to C_2 \div 3 \\ C_4 \to C_4 \div 2 \end{matrix}$$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 7 & 23 & 0 & 0 \\ 5 & 11 & 0 & 0 \end{bmatrix} \begin{matrix} C_3 \to C_3 + C_2 \\ C_4 \to C_4 - C_2 \end{matrix}$$

In this final form of A, the fourth order and third order determinants are 0.

The leading minor of order 2 is

$$\begin{vmatrix} 1 & 0 \\ 7 & 23 \end{vmatrix} = 23 \neq 0$$

∴ The rank of A = 2.

6. Find the rank of the matrix $A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 4 & 1 & 0 & 2 \\ 0 & 3 & 4 & 2 \end{bmatrix}$ by examining the determinant minors.

Solution :

$$\begin{vmatrix} 1 & 1 & 1 \\ 4 & 1 & 0 \\ 0 & 3 & 4 \end{vmatrix} = 0 = \begin{vmatrix} 1 & 1 & 1 \\ 1 & 0 & 2 \\ 3 & 4 & 2 \end{vmatrix}$$

$$\begin{vmatrix} 1 & 1 & 1 \\ 4 & 1 & 2 \\ 0 & 3 & 2 \end{vmatrix} = 0 = \begin{vmatrix} 1 & 1 & 1 \\ 4 & 0 & 2 \\ 0 & 4 & 2 \end{vmatrix}$$

∴ Every 3×3 submatrix of A has determinant zero.

Also, $\begin{vmatrix} 1 & 1 \\ 4 & 1 \end{vmatrix} = -3 \neq 0$

∴ Rank of A = 2.

Exercises :

1. Find the rank of the following matrices.

(a) $\begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 0 & 2 & 2 \end{bmatrix}$ (b) $\begin{bmatrix} 2 & 3 & 4 \\ 3 & 1 & 2 \\ 2 & 2 & 2 \end{bmatrix}$ (c) $\begin{bmatrix} 2 & 3 & 1 \\ 4 & 6 & 2 \\ -6 & -9 & -3 \end{bmatrix}$

(d) $\begin{bmatrix} 0 & 1 & 2 & 1 \\ 2 & -3 & 0 & -1 \\ 1 & 1 & -1 & 0 \end{bmatrix}$ (e) $\begin{bmatrix} 1 & -1 & 0 & 2 & 1 \\ 3 & 1 & 1 & -1 & 2 \\ 4 & 0 & 1 & 0 & 3 \\ 9 & -1 & 2 & 3 & 7 \end{bmatrix}$

2. Find the column rank of the matrices.

(a) $\begin{bmatrix} 1 & 2 & -1 & 3 \\ 2 & 4 & 1 & -2 \\ 3 & 6 & 3 & -7 \end{bmatrix}$ (b) $\begin{bmatrix} 3 & 1 & -5 & -1 \\ 1 & -2 & 1 & -5 \\ 1 & 5 & -7 & 2 \end{bmatrix}$

(Hint : Row rank = rank of the matrix = column rank).

3. Find the row rank of the matrix

$$\begin{bmatrix} 1 & 3 & 1 & -2 \\ 1 & 4 & 3 & -1 \\ 2 & 3 & -4 & -7 \\ 3 & 8 & 1 & -7 \end{bmatrix}$$

Answers :

1. (a) 3 (b) 2 (c) 1 (d) 3 (e) 3
2. (a) 2 (b) 3
3. 2

6.1. REDUCTION TO NORMAL FORMS

Definition :

By means of elementary operations any non-zero matrix can be reduced to a simple form called the **normal form** of the matrix.

Theorem 6.1 :

Every non-zero m×n matrix A can be reduced to a matrix of the form $\begin{pmatrix} I_r & O_{r,n-r} \\ O_{m-r,r} & O_{m-r,n-r} \end{pmatrix}$ by successive applications of a finite sequence of elementary row and column operations where Op,q is the p×q zero matrix.

Proof :

Proof is by induction on the number of rows of A. Let m = 1, (i.e.,) A has only one row, say $A = (a_{11}, a_{12},.....,a_{1n})$.

Since $A \neq 0$, by interchanging coloums, if necessary we can bring a non-zero entry α in the first place.Multiplying A by α^{-1} we get 1 in the first place Make the other entries of A zero by adding suitable multiples of 1 of them. Thus, the theorem is true when m=1.

Assume that the theroem is true for any non-zero matrix with m–1 rows. Let A be a non-zero m×n matrix.

Let a_{ij} be a non-zero entry of A. Interchange the 1st and i^{th} rows; then interchange the 1st and j^{th} columns. We then have a_{ij} in the (1, 1) position. Multiplying the first row by a_{ij}^{-1} we get 1 in the (1, 1) position.

All other entries in the first column can be made zero by adding suitable multiples of first row to each row. Similarly, all other entries in the first row can be made zero. We thus arrive at a matrix of the form

$$\begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & & & & \\ \vdots & & B & & \\ 0 & & & & \end{bmatrix}$$

, where B is an (m–1)×(n–1) matrix. By induction hypothesis, B can be reduced to the desired form by elementary operations.

However, any elementary row or column operation on B can be considered as an elementary operation on the corresponding rows or columns of A and does not alter the first row or the first column of A. Hence the theorem.

Corollary 1 :

Let A be a non-zero m×n matrix. Then these exist non-singular square matrices P and Q of orders m and n respectively such that PAQ is a matrix of the form $\begin{pmatrix} I_r & O_{r,n-r} \\ O_{m-r,r} & O_{m-r,n-r} \end{pmatrix}$.

Proof :

We know that every elementary operation on A is equivalent to multiplying A by an elementary matrix.

∴ The previous theorem can be stated as

$$P_1.....P_iAQ_1.....Q_j = \begin{pmatrix} I_r & O_{r,n-r} \\ O_{m-r,r} & O_{m-r,n-r} \end{pmatrix} \qquad \text{---------(1)}$$

where P's and Q's are elementary matrices.

But elementary matrices are non-singular and any product of non-singular matrices is also non-singular.

∴ if $P_1,.....P_i = P$ and $Q_1,.....Q_j = Q$

where P,Q are non-singular matrices, then from (1) $PAQ = \begin{pmatrix} I_r & O_{r,n-r} \\ O_{m-r,r} & O_{m-r,n-r} \end{pmatrix}$

Corollary 2 :

Any non-singular square matrix A of order n is equivalent to the unit matrix I_n.

Proof :

By corollary 1, there exist non-singular square matrices P and Q such that

$$PAQ = \begin{pmatrix} I_r & O_{r,n-r} \\ O_{n-r,r} & O_{n-r,n-r} \end{pmatrix}$$

Since P, A and Q are non-singular the matrix PAQ is also non-singular.

∴ $\begin{pmatrix} I_r & O_{r,n-r} \\ O_{n-r,r} & O_{n-r,n-r} \end{pmatrix}$ is non-singular. This is possible iff $\begin{pmatrix} I_r & O_{r,n-r} \\ O_{n-r,r} & O_{n-r,n-r} \end{pmatrix} = I_n$.

Examples :

1. Find whether the matrix $A = \begin{bmatrix} 1 & 2 & 1 \\ -1 & 0 & 2 \\ 2 & 1 & -3 \end{bmatrix}$ is equivalent to I_3

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

$$\sim \begin{bmatrix} 1 & 2 & 1 \\ 0 & 2 & 3 \\ 0 & -3 & -5 \end{bmatrix} \begin{matrix} R_2 \to R_2 + R_1 \\ R_3 \to R_3 - 2R_1 \end{matrix}$$

$$\sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 3 \\ 0 & -3 & -5 \end{bmatrix} \begin{matrix} C_2 \to C_2 - 2C_1 \\ C_3 \to C_3 - C_1 \end{matrix}$$

$$\sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 3/2 \\ 0 & 3 & 5 \end{bmatrix} \begin{matrix} R_2 \to \frac{1}{2}R_2 \\ R_3 \to (-1)R_3 \end{matrix}$$

$$\sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 3/2 \\ 0 & 0 & 1/2 \end{bmatrix} R_3 \to R_3 - 3R_2$$

$$\sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1/2 \end{bmatrix} R_2 \to R_2 \to 3R_3$$

$$\sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} R_3 \to 2R_3$$

$$= I_3$$

∴ A is equivalent to I_3.

2. Reduce the matrix $\begin{bmatrix} 2 & -2 & 0 & 6 \\ 4 & 2 & 0 & 2 \\ 1 & -1 & 0 & 3 \\ 1 & -2 & 1 & 2 \end{bmatrix}$ to its normal form

Solution :

$$\begin{bmatrix} 2 & -2 & 0 & 6 \\ 4 & 2 & 0 & 2 \\ 1 & -1 & 0 & 3 \\ 1 & -2 & 1 & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & 0 & 3 \\ 4 & 2 & 0 & 2 \\ 1 & -1 & 0 & 3 \\ 0 & -1 & 1 & -1 \end{bmatrix} \begin{matrix} R_1 \to \frac{1}{2}R_1 \\ \\ \\ R_4 \to R_4 - R_3 \end{matrix}$$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 4 & 6 & 0 & -10 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 1 & -1 \end{bmatrix} \begin{matrix} \\ C_2 \to C_2 + C_1 \\ \\ C_4 \to C_4 - 3C_1 \end{matrix}$$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -5/3 \\ 0 & -1 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{matrix} \\ R_2 \to \frac{1}{6}R_2 \\ R_3 \Leftrightarrow R_4 \\ \end{matrix}$$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -1 & 1 & -8/3 \\ 0 & 0 & 0 & 0 \end{bmatrix} C_4 \to C_4 + \frac{5}{3}C_2$$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -8/3 \\ 0 & 0 & 0 & 0 \end{bmatrix} C_2 \to C_2 + C_3$$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} C_4 \to C_4 + \frac{8}{3}C_3$$

$$\sim \begin{pmatrix} I_3 & O_{3,1} \\ O_{1,3} & O_{1,1} \end{pmatrix}$$

Exercises :

1. Find the matrix obtained from $A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ -1 & 2 & -3 \end{bmatrix}$ by applying in this order

   (i) the elementary row operations $R_1 \Leftrightarrow R_3$, $R_2 \to 2R_2$, $R_1 \to R_1 - R_3$ and

   (ii) the elementary column operations $C_1 \to -2C_1$, $C_2 \Leftrightarrow C_3$, $C_2 \to C_2 - C_1$.

2. Reduce to normal form the matrices,

(i) $\begin{bmatrix} 1 & 2 & 0 & -1 \\ 3 & 4 & 1 & 2 \\ -2 & 3 & 2 & 5 \end{bmatrix}$ (ii) $\begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{bmatrix}$ (iii) $\begin{pmatrix} 1 & 0 & -7 \\ 0 & 1 & 2 \end{pmatrix}$

6.2. SIMILAR AND CONGRUENT MATRICES

Definition :

Let V be an n dimensional vector space over a field F. Let $T \in A(V)$ have the matrix $M_1(T)$ in the basis $\{v_1, v_2, \ldots\ldots, v_n\}$ and the matrix $M_2(T)$ in the basis $\{w_1, w_2, \ldots . w_n\}$. We are interested in knowing whether there exists any relationship between the matrices $M_1(T)$ and $M_2(T)$. The following theorem gives an answer to this.

Theorem 6.2 :

Let V be an n dimensional vector space over a field F. Let $T \in A(V)$ have the matrix $M_1(T)$ in the basis $\{v_1, v_2, \ldots\ldots, v_n\}$ and the matrix $M_2(T)$ in the basis $\{w_1, \ldots . w_n\}$. Then there exists a non-singular matrix P of order n such that $M_2(T) = P^{-1}M_1(T)P$.

Proof :

Let $M_1(T) = (a_{ij})$ and $M_2(T) = (b_{ij})$

By definition,

$$T(v_j) = a_{1j}v_1 + a_{2j}v_2 + \ldots\ldots + a_{nj}v_n$$

$$T(w_j) = b_{1j}w_1 + b_{2j}w_2 + \ldots + b_{nj}w_j, \qquad j = 1, 2, \ldots, n.$$

Define a map $S: V \to V$ by $S(v_j) = w_j$

Clearly S is a linear map.

Also $v \in V \Rightarrow v = \alpha_1 w_1 + \alpha_2 w_2 + \ldots\ldots + \alpha_n w_n$

$\Rightarrow v = \alpha_1 S(v_1) + \alpha_2 S(v_2) + \ldots\ldots + \alpha_n S(v_n)$

$\Rightarrow v = S(\alpha_1 v_1 + \ldots\ldots + \alpha_n v_n)$

$\Rightarrow v = S(v^1)$ where $v^1 = \alpha_1 v_1 + \ldots\ldots + \alpha_n v_n \in V$

$\therefore$ S is onto.

By using the theorem,

Let V be finite dimensional over F and let $T \in A(V)$. Then the following are equivalent.

(a) T is regular

(b) T is non-singular

(c) T is onto.

S is regular

$\therefore$ S^{-1} exists in A(V).

Now $T(wj) = b_{1j} w_1 + \ldots\ldots + b_{nj} w_n$

$= b_{1j} S(v_1) + \ldots\ldots + b_{nj} S(v_n)$

$= S(b_{1j} v_1 + \ldots\ldots + b_{nj}(v_n))$

$\therefore$ $(S^{-1}TS)(v_j) = (S^{-1}T)(S(v_j))$

$= (S^{-1}T)(w_j)$

$= S^{-1}(T(w_j))$

$= (S^{-1}S)(b_{1j} v_1 + \ldots\ldots + b_{nj} v_n)$

$= b_{1j} v_1 + \ldots\ldots + b_{nj} v_n$

Hence the matrix of $S^{-1}TS$ in the basis $\{v_1, v_2, \ldots\ldots, v_n\}$ is (b_{ij}).

(i.e.,) $M_2(T) = (b_{ij}) = M_1(S^{-1}TS)$

$= M_1(S^{-1}) M_1(T) M_1(S)$

$= [M_1(S)]^{-1} M_1(T) M_1(S)$

$= P^{-1} M_1(T) P$, where $P = M_1(S)$.

Example :

Let $T:R^2 \to R^2$ be the linear transformation defined by $T(x,y) = (2y, 3x-y)$. Find (i) the matrix $M_1(T)$ in the basis $\{(1,0), (0,1)\}$; (ii) the matrix $M_2(T)$ in the basis $\{(1,3),(2,5)\}$; and (iii) a non-singular matrix P such that $M_2(T) = P^{-1}M_1(T)P$.

(i) Let $e_1 = (1, 0)$, $e_2 = (0, 1)$

Then $T(e_1) = (0, 3) = 0e_1+3e_2$

and $T(e_2) = (2, -1) = 2e_1-1e_2$

$\therefore$ $M_1(T) = \begin{pmatrix} 0 & 2 \\ 3 & -1 \end{pmatrix}$

(ii) Let $v_1 = (1, 3)$, $v_2 = (2, 5)$

Then $T(v_1) = (6, 0) = -30v_1+18v_2$

and $T(v_2) = (10, 1) = -48v_1+29v_2$

$\therefore$ $M_2(T) = \begin{pmatrix} -30 & -48 \\ 18 & 29 \end{pmatrix}$

(iii) Define $S:R^2 \to R^2$ by $S(e_j) = v_j$. Then S is linear and

$S(e_1) = (1, 3) = 1e_1+3e_2$

$S(e_2) = (2, 5) = 2e_1+5e_2$

$\therefore$ Matrix of S in the basis $\{e_1, e_2\}$ is

$$M_1(S) = \begin{pmatrix} 1 & 2 \\ 3 & 5 \end{pmatrix} = P \text{ say}$$

Clearly P is non singular and $P^{-1} = \begin{pmatrix} -5 & 2 \\ 3 & -1 \end{pmatrix}$

Also, $$P^{-1}M_1(T)P = \begin{pmatrix} -5 & 2 \\ 3 & -1 \end{pmatrix}\begin{pmatrix} 0 & 2 \\ 3 & -1 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 3 & 5 \end{pmatrix}$$

$$= \begin{pmatrix} 6 & -12 \\ -3 & 7 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 3 & 5 \end{pmatrix}$$

$$= \begin{pmatrix} -30 & -48 \\ 18 & 29 \end{pmatrix}.$$

$$= M_2(T)$$

Definition :

Let A and B be n×n matrix over F. Then B is said to be similar to A if there exists a non-singular n×n matrix P such that $B=P^{-1}AP$.

Theorem 6.3 :

Similarity of matrices is an equivalence relation on the set S of all n×n matrices.

Proof :

(i) The identity matrix I of order n is non-singular and $I^{-1} = I$.

Since $I^{-1}AI = A$, A is similar to A for all $A \in S$.

∴ Similarity is a reflexive relation.

(ii) Let $A, B \in S$ and let A be similar to B. Then there exists a non-singular matrix P such that $A = P^{-1}AP$.

∴ $B = PAP^{-1} = (P^{-1})^{-1}AP^{-1}$

Since P^{-1} is non-singular, B is similar to A.

∴ Similarity is a symmetric relation.

(iii) Let $A, B, C \in S$ and let A be similar to B and B similar to C. Then there exists non-singular matrices P and Q such that $A = P^{-1}BP$ and $B=Q^{-1}CQ$.

∴ $A = P^{-1}BP = P^{-1}(Q^{-1}CQ)P = (QP)^{-1}C(QP)$

Since Q, P are non-singular, QP is also non-singular.

∴ A is similar to C.

∴ Similarity is a transitive relation.

Hence similarity of matrices is an equivalence relation.

Remark :

If A is similar to B, we can say that A and B are similar matrices.

Theorem 6.4 :

Similar matrices have the same characteristic equation.

Proof :

Let A be an n×n matrix and P, a non-singular matrix of order n. Then A and $P^{-1}AP$ are similar.

If P and A are n×n matrices and P is non-singular then A & $P^{-1}AP$ have the same characteristic roots. From this A and $P^{-1}AP$ have the same characteristic equation.

Note :

Theorem 6.2 says that the matrices associated with the same linear transformation T:V→V with respect to different bases are similar.

Definition :

A matrix B of order n is said to be **congruent** to a matrix A of order n if there exists a non-singular matrix P such that $B=P^TAP$.

It is easy to prove that congruence of matrices is an equivalence relation.

If A and B are congruent matrices, then $B = P^TAP$, where P is non-singular. Also, P^T is non-singular.

∴ A and B have the same rank.

(i.e.,) congruent matrices have the same rank.

Definition :

The **trace** of a square matrix A over F, written as tr A is the sum of the elements on the leading diagonal of A.

If $A = (a_{ij})$ is of type n×n, then

$$\text{tr } A = a_{11}+a_{22}+....+a_{nn}$$

Lemma 6.5 :

Let A, B be n×n matrices over F. Let $\lambda \in F$. Then

(i) $\text{tr } (\lambda A) = \lambda \text{ tr } A$

(ii) $\text{tr } (A+B) = \text{tr } A + \text{tr } B$

(iii) tr (AB) = tr (BA)

(iv) If A is a similar to B, then tr A = tr B.

Proof :

Let $A = (a_{ij})$, $B = (b_{ij})$

(i) $\text{tr}(\lambda A) = \lambda a_{11}+\lambda a_{22}+.....+\lambda a_{nn}$

$= \lambda(a_{11}+.....+a_{nn})$

$= \lambda\ (a_{11}+.....+a_{nn})$

$= \lambda\ \text{tr}\ (A)$

(ii) $\text{tr}(A+B) = (a_{11}+b_{11})+(a_{22}+b_{22})+.....+(a_{nn}+b_{nn})$

$= (a_{11}+a_{22}+....+a_{nn})+(b_{11}+b_{22}+.....+b_{nn})$

$= \text{tr}\ A + \text{tr}\ B$

(iii) $AB = (C_{ij})$ where $C_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}$

$\therefore\ \text{tr}(AB) = \sum_{i=1}^{n} C_{ii} = \sum_{i=1}^{n}\sum_{k=1}^{n} a_{ik}b_{ki}$

$= \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}b_{ji}$

$BA = (d_{ij})$ where $d_{ij} = \sum_{k=1}^{n} b_{ik}a_{kj}$

$\therefore\ \text{tr}\ (BA) = \sum_{j=1}^{n} d_{ij} = \sum_{j=1}^{n}\sum_{k=1}^{n} b_{jk}a_{kj}$

$= \sum_{j=1}^{n}\sum_{i=1}^{n} b_{ji}a_{ij}$

$= \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}b_{ji} = \text{tr}\ (AB)$

(iv) If A is similar to B, then there exists a non-singular matrix P such that $A = P^{-1}BP$.

$\therefore\ \text{tr}\ A = \text{tr}\ (P^{-1}BP) = \text{tr}\ (PP^{-1}B)$ by (iii)

$= \text{tr}\ (B)$

Definition :

If $J \in A(V)$, 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.

The above definition is meaningful, it depends only on T and not on any particular basis of V. For, if $M_1(T)$ and $M_2(T)$ are the matrices of T in two different basis of V, then by theorem 6.2 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 tr T does not depend upon any particular basis of V.

For example, let $T:R^2 \to R^2$ be the linear transformation defined by $T(x,y) = (2y, 3x-y)$.

Then in the basis $\{(1,0), (0, 1)\}$ matrix of T is $\begin{pmatrix} 0 & 2 \\ 3 & -1 \end{pmatrix}$.

$\therefore$ tr T = 0–1 = –1.

Also, in the basis $\{(1, 3), (2, 5)\}$ matrix of T is $\begin{pmatrix} -30 & -48 \\ 18 & 29 \end{pmatrix}$.

$\therefore$ tr T = –30 + 29 = –1.

Definition :

Let V be an n dimensional vector space over a field F. Then $T \in A(V)$ is said to be **similar** to $S \in A(V)$ if these exists a non-singular linear transformation $P \in A(V)$ such that $T=P^{-1}SP$.

It is easy to check that the relation of similarity is an equivalence relation on A(V). The equivalence class of $T \in A(V)$ is called its **similarity class**. To find whether two linear transformations are similar we calculate a particular canonical form for each and see if these are the same.

Exercises :

1. If A and B are similar matrices, show that their determinants are equal.

2. Let V be the vector space of all polynomials in x over F of degree$\leq$3. Let D be the differential operator d/dx.

Find (i) the matrix $M_1(D)$ in the basis $\{1, x, x^2, x^3\}$ and (ii) the matrix $M_2(D)$ in the basis $\{1, 1+x, 1+x^2, 1+x^3\}$. Also find a non-singular matrix P such that $M_2(D) = P^{-1}M_1(D)P$.

3. Let $V=R^2$. Show that it is impossible to find a matrix P such that $P^{-1}\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}P = \begin{pmatrix}a & 0\\ 0 & b\end{pmatrix}$ for any $a, b \in R$.

4. Let $V = R^3$ and let $\begin{bmatrix}1 & 1 & 2\\ -1 & 2 & 1\\ 0 & 1 & 3\end{bmatrix}$ be the matrix of $T \in A(V)$ in the basis $\{(1,0,0), (0,1,0), (0,0,1)\}$. Find the matrix of T in the basis $\{(1,1,0), (1,2,0), (1,2,1)\}$.

5. Prove that the relation of congruence in matrices is an equivalence relation.

6. If A is non-singular, show that every matrix congruent to A is also non-singular.

7. Let $T:R^3 \to R^3$ be the linear map defined by $T(x,y,z) = (2y+z, x-4y, 3x)$. Find tr T.

6.3. SOLUTION OF SYSTEMS OF LINEAR EQUATIONS USING MATRICES AND DETERMINANTS

Matrix form of a set of linear equations :

Consider a system of m linear equations in n unknowns $x_1, x_2, \ldots, x_n$ given by

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

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

.....

.....

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

Using the concept of matrix multiplication and equality of matrices this system can be written as $AX = B$ where

$$A = \begin{bmatrix}a_{11} & a_{12} & \ldots & a_{1n}\\ a_{21} & a_{22} & \ldots & a_{2n}\\ \vdots & \vdots & \vdots & \vdots\\ a_{m1} & a_{m2} & \ldots & a_{mn}\end{bmatrix}$$

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

The m×n matrix A is called the **co-efficient matrix.**

Definition :

A set of values of $x_1, x_2,.....,x_n$ which satisfy the above system of equations is called a **solution** of the system. The system of equations is said to be **consistent** if it has atleast one solution. Otherwise the system is said to be **inconsistent.**

The m×(n+1) matrix given by

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} & b_1 \\ a_{21} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{bmatrix}$$

is called the **augmented matrix** of the system and is denoted by (A,B). Thus the augmented matrix (A, B) is obtained by annexing to A the column matrix B, which becomes the $(n+1)^{th}$ column in (A, B).

Note : Since every column in A appears in (A, B) the column space of the matrix A is a subspace of the column space of the matrix (A, B).

Hence the rank of A ≤ rank of (A, B).

Theorem 6.6 :

The system of linear equations AX=B is consistent iff rank of A = rank of (A, B).

Proof :

Let the system be consistent.

Let $u_1, u_2,.....,u_n$ be a solution of the system.

Then $B = u_1C_1+u_2C_2+.....+u_nC_n$ where $C_1, C_2,.....,C_n$ denote the columns of A.

Hence the column space of the augmented matrix (A, B) namely $<C_1,C_2,....,C_n, B>$ is the same as the column space $<C_1,C_2,.....C_n>$ of A.

Hence the rank of A = rank of (A, B).

Conversely let rank of A = rank of (A, B)

Then the column rank of A = column rank of (A, B)

$\therefore \dim\langle C_1, C_2, \ldots C_n\rangle = \dim\langle C_1, C_2, \ldots C_n, B\rangle$

But $\langle C_1, C_2, \ldots, C_n\rangle$ is a subspace of $\langle C_1, C_2, \ldots, C_n, B\rangle$

$\therefore$ B is a linear combination of $C_1, C_2, \ldots, C_n$.

If $B = u_1C_1 + \ldots + u_nC_n$ then $u_1, u_2, \ldots, u_n$ is a solution of the system.

Hence the theorem.

Remark :

The solution of a given of simultaneous equations is not altered by interchanging any two equations or by multiplying any equation by a non-zero constant or by adding a multiple of one equation to another. Hence we can reduce the given system of equations to an equivalent system by applying elementary row operations to the augmented matrix. This reduced form will enable us to test for the consistency and to find the solution if it exists. This is illustrated in the following problems.

Solved Problems :

Problem 1 :

Show that the equations

$$\begin{aligned} x+y+z &= 6 \\ x+2y+3z &= 14 \\ x+4y+7z &= 30 \end{aligned}$$

are consistent and solve them.

Solution :

The given system of equations can be put in the matrix form

$$AX = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 4 & 7 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 6 \\ 14 \\ 30 \end{bmatrix} = B$$

The augmented matrix is given by

$$(A, B) = \begin{bmatrix} 1 & 1 & 1 & 6 \\ 1 & 2 & 3 & 14 \\ 1 & 4 & 7 & 30 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 1 & 1 & 6 \\ 0 & 1 & 2 & 8 \\ 0 & 3 & 6 & 24 \end{bmatrix} \begin{matrix} R_2 \to R_2 - R_1 \\ R_3 \to R_3 - R_1 \end{matrix}$$

$$\sim \begin{bmatrix} 1 & 1 & 1 & 6 \\ 0 & 1 & 2 & 8 \\ 0 & 0 & 0 & 0 \end{bmatrix} R_3 \to R_3 - 3R_2$$

Hence rank of A = Rank of (A, B) = 2.

Hence the given system is consistent.

Also the given system of equations reduces to

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 6 \\ 8 \\ 0 \end{bmatrix}$$

$$x+y+z = 6$$

$$y+2z = 8$$

Putting z = C, we obtain the general solution of the system as x = C–2, y=8–2C, z=C.

Problem 2 :

Verify whether the following system of equations is consistent. If it is consistent, find the solution.

$$x-4y-3z = -16$$

$$4x-y+6z = 16$$

$$2x+7y+12z = 48$$

$$5x-5y+3z = 0$$

Solution :

The matrix form of the system is given by

$$\begin{bmatrix} 1 & -4 & -3 \\ 4 & -1 & 6 \\ 2 & 7 & 12 \\ 5 & -5 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -16 \\ 16 \\ 48 \\ 0 \end{bmatrix}$$

∴ The augmented matrix is given by

$$(A, B) = \begin{bmatrix} 1 & -4 & -3 & -16 \\ 4 & -1 & 6 & 16 \\ 2 & 7 & 12 & 48 \\ 5 & -5 & 3 & 0 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & -4 & -3 & -16 \\ 0 & 15 & 18 & 80 \\ 0 & 15 & 18 & 80 \\ 0 & 15 & 18 & 80 \end{bmatrix} \begin{matrix} \\ R_2 \to R_2 - 4R_1 \\ R_3 \to R_3 - 2R_1 \\ R_4 \to R_4 - 5R_1 \end{matrix}$$

$$\sim \begin{bmatrix} 1 & -4 & -3 & -16 \\ 0 & 15 & 18 & 80 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{matrix} \\ R_3 \to R_3 - R_2 \\ R_4 \to R_4 - R_2 \\ \\ \end{matrix}$$

∴ Rank of A = Rank of (A, B) = 2 and hence the system is consistent. Also the system of equations reduces to

$$\begin{bmatrix} 1 & -4 & -3 \\ 0 & 15 & 18 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -16 \\ 80 \\ 0 \\ 0 \end{bmatrix}$$

∴ $x - 4y - 3z = -16$ and

$15y + 18z = 80$

Putting z=C we obtain the general solution of the systems as

$$x = -\left(\frac{9C}{5}\right) + \left(\frac{16}{3}\right)$$

$$y = -\left(\frac{6C}{5}\right) + \left(\frac{16}{3}\right)$$

$$z = C$$

Problem 3 :

Show that the equations

$$\begin{aligned} x+2y-z &= 3 \\ 3x-y+2z &= 1 \\ 2x-2y+3z &= 2 \\ x-y+z &= -1 \end{aligned}$$

are consistent and solve the same.

Solution :

The matrix form of the system is given by

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

Here the coefficient matrix

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

$$\sim \begin{bmatrix} 1 & 2 & -1 \\ 0 & -7 & 5 \\ 0 & -6 & 5 \\ 0 & -3 & 2 \end{bmatrix} \begin{matrix} R_2 \to R_2 - 3R_1 \\ R_3 \to R_3 - 2R_1 \\ R_4 \to R_4 - R_1 \end{matrix}$$

In this form, the third order minor

$$\begin{vmatrix} 1 & 2 & -1 \\ 0 & -7 & -5 \\ 0 & -6 & 5 \end{vmatrix} = -35-30 \neq 0$$

Hence the rank of A is 3.

The augmented matrix.

$$(A, B) = \begin{bmatrix} 1 & 2 & -1 & 3 \\ 3 & -1 & 2 & 1 \\ 2 & -2 & 3 & 2 \\ 1 & -1 & 1 & -1 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 2 & -1 & 3 \\ 0 & -7 & 5 & -8 \\ 0 & -6 & 5 & -4 \\ 0 & -3 & 2 & -4 \end{bmatrix} \begin{matrix} R_2 \to R_2 - 3R_1 \\ R_3 \to R_3 - 2R_1 \\ R_4 \to R_4 - R_1 \end{matrix}$$

$$\sim \begin{bmatrix} 1 & 2 & -1 & 3 \\ 0 & -7 & 5 & -8 \\ 0 & 0 & 1 & 4 \\ 0 & -3 & 2 & -4 \end{bmatrix} R_3 \to R_3 - 2R_4$$

$$\sim \begin{bmatrix} 1 & 2 & -1 & 3 \\ 0 & -7 & 5 & -8 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & -1 & -4 \end{bmatrix} R_4 \to 7R_4 - 3R_2$$

$$\sim \begin{bmatrix} 1 & 2 & -1 & 3 \\ 0 & -7 & 5 & -8 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix} R_4 \to R_4 + R_3$$

In this form, the fourth order determinant is zero.

The third order determinant

$$\begin{vmatrix} 1 & 2 & -1 \\ 0 & -7 & 5 \\ 0 & 0 & 0 \end{vmatrix} = -7 \neq 0$$

Hence the rank of [A, B] is 3.

The two matrices have the same rank. So the equations are consistent.

From the final form of the augmented matrix, the given system is equivalent to the equations

$$x+2y-z = 3 \quad \text{--------(1)}$$

$$-7y+5z = -8 \quad \text{--------(2)}$$

$$z = 4 \quad \text{--------(3)}$$

putting (3) in (2),

$$-7y = -8-20 = -28$$

$\therefore$ $y = 4$

From (1), $x = 3-8+4 = -1$

$\therefore$ $x = -1$

$\therefore$ $x = -1,\ y = 4\ \&\ z = 4$

Problem 4 :

Examine for consistency the following equations.

$$2x+6y+11 = 0$$

$$6x+20y-6z+3 = 0$$

$$6y-18z+1 = 0$$

The equation can be written as

$$2x+6y+0z = -11$$

$$6x+20y-6z = -3$$

$$0x+6y-18z = -1$$

Here the coefficient matrix

$$A = \begin{bmatrix} 2 & 6 & 0 \\ 6 & 20 & -6 \\ 0 & 6 & -18 \end{bmatrix}$$

$$\sim \begin{bmatrix} 2 & 6 & 0 \\ 0 & 2 & 6 \\ 0 & 6 & -18 \end{bmatrix} R_2 \to R_2 - 3R_1$$

In this form, the third order determinant

$$= 2\begin{vmatrix} 2 & -6 \\ 6 & -18 \end{vmatrix} = 2(-36+36) = 0$$

The second order minor $\begin{vmatrix} 2 & 6 \\ 0 & 2 \end{vmatrix} = 4 \neq 0$

Hence A is of rank 2.

The augmented matrix

$$[A, B] = \begin{bmatrix} 2 & 6 & 0 & -11 \\ 6 & 20 & -6 & -3 \\ 0 & 6 & -18 & -1 \end{bmatrix}$$

$$\sim \begin{bmatrix} 2 & 6 & 0 & -11 \\ 0 & 2 & -6 & 30 \\ 0 & 6 & -18 & -1 \end{bmatrix} R_2 \to R_2 - 3R_1$$

$$\sim \begin{bmatrix} 2 & 6 & 0 & -11 \\ 0 & 2 & -6 & 30 \\ 0 & 0 & 0 & -91 \end{bmatrix} R_3 - 3R_2$$

In this form, the third order minor

$$\begin{vmatrix} 6 & 0 & -11 \\ 2 & -6 & 30 \\ 0 & 0 & -91 \end{vmatrix} = -91 \begin{vmatrix} 6 & 0 \\ 2 & -6 \end{vmatrix}$$

$$= 91 \times 36 \neq 0$$

Hence $R[A, B] \geq 3$

But [A, B] has 3 rows and 4 columns.

$\therefore$ $R[A, B] \leq 3$

$\therefore$ $R[A, B] = 3$

(i.e.,) the coefficient matrix and the augmented matrix are not of the same rank.

Hence the given system of equation is inconsistent.

Problem 5 :

For what values of η the equations

$$x+y+z = 1$$

$$x+2y+4z = \eta$$

$$x+4y+10z = \eta^2 \text{ are consistent?}$$

Solution :

The matrix form of the system is given by

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 4 & 10 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ \eta \\ \eta^2 \end{bmatrix}$$

∴ The augmented matrix is given by

$$(A, B) = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & \eta \\ 1 & 4 & 10 & \eta^2 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 3 & \eta-1 \\ 0 & 3 & 9 & \eta^2-1 \end{bmatrix} \begin{matrix} R_2 \to R_2 - R_1 \\ R_3 \to R_3 - R_1 \end{matrix}$$

$$\sim \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 3 & \eta-1 \\ 0 & 0 & 0 & \eta^2-3\eta+2 \end{bmatrix} R_3 \to R_3 - 3R_2$$

∴ The given system is consistent iff $\eta^2-3\eta+2 = 0$

∴ $\eta = 2$ (or) 1.

Problem 6:

Show that the system of equations

$$\begin{aligned} x+2y+z &= 11 \\ 4x+6y+5z &= 8 \\ 2x+2y+3z &= 19 \text{ is inconsistent.} \end{aligned}$$

Solution :

The matrix form of the system is given by

$$\begin{bmatrix} 1 & 2 & 1 \\ 4 & 6 & 5 \\ 2 & 2 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 11 \\ 8 \\ 19 \end{bmatrix}$$

∴ The augmented matrix is given by

$$(A, B) = \begin{bmatrix} 1 & 2 & 1 & 11 \\ 4 & 6 & 5 & 8 \\ 2 & 2 & 3 & 19 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 2 & 1 & 11 \\ 0 & -2 & 1 & -36 \\ 0 & -2 & 1 & -3 \end{bmatrix} \begin{matrix} R_2 \to R_2 - 4R_1 \\ R_3 \to R_3 - 2R_1 \end{matrix}$$

$$\sim \begin{bmatrix} 1 & 2 & 1 & 11 \\ 0 & -2 & 1 & -36 \\ 0 & 0 & 0 & 33 \end{bmatrix} R_3 \to R_3 - R_2$$

∴ Rank of A = 2 and Rank of (A, B) = 3.

∴ The given system is inconsistent.

Problem 7 :

Investigate for what values of a, b the simultaneous equations x+y+2z = 2, 2x–y+3z = 2, 5x–y+az = b have (i) no solution (ii) a unique solution and (iii) an infinite number of solutions.

Solution :

Here the coefficient matrix

$$A = \begin{bmatrix} 1 & 1 & 2 \\ 2 & -1 & 3 \\ 5 & -1 & a \end{bmatrix}$$

$$|A| = 1(-a+3)-1(2a-15)+2(-2+5)$$

$$= -a+3-2a+15+6$$

$$= -3a+24 = 3\,(8-a)$$

If $a \neq 8$, $|A| \neq 0$. Hence A is of rank 3 and the augmented matrix

$$[A, B] = \begin{bmatrix} 1 & 1 & 2 & 2 \\ 2 & -1 & 3 & 2 \\ 5 & -1 & a & b \end{bmatrix}$$

will also be of rank 3 as the leading third order determinant $\neq 0$. Hence if $a \neq 8$ and whatever be the value of b, the matrices A and [A, B] will have the same rank. So the given equations will be consistent.

Also in the case, the common rank = 3 = the number of unknowns.

So the system will have a unique solution.

If $a = 8$, $|A| = 0$.

Leading minor of order 2 is $\begin{vmatrix} 1 & 1 \\ 2 & -1 \end{vmatrix} = -1-2 \neq 0$

Hence A is of rank 2.

$$[A, B] = \begin{bmatrix} 1 & 1 & 2 & 2 \\ 2 & -1 & 3 & 2 \\ 5 & -1 & 8 & 6 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 1 & 2 & 2 \\ 0 & -3 & -1 & -2 \\ 0 & -6 & -2 & b-10 \end{bmatrix} \begin{matrix} R_2 \to R_2 - 2R_1 \\ R_3 \to R_3 - 5R_1 \end{matrix}$$

$$\sim \begin{bmatrix} 1 & 1 & 2 & 2 \\ 0 & -3 & -1 & -2 \\ 0 & 0 & 0 & b-6 \end{bmatrix} R_3 \to R_3 - 2R_2$$

If b = 6, in the above final form, the last row will contain zeros.

Hence all the four third order determinants will be zero.

The leading minor of order 2 is $\begin{vmatrix} 1 & 1 \\ 0 & -3 \end{vmatrix} = -3 \neq 0$

Hence R[A, B] = 2

The two matrices A and [A, B] have the same rank and the equations will be consistent.

In this case, common rank = 2 and this is < 3, the number of unknowns. So the system will have an infinite number of solutions.

If $b \neq 6$, the third order determinant in the final form of [A, B].

$$\begin{vmatrix} 1 & 2 & 2 \\ -3 & -1 & -2 \\ 0 & 0 & b-6 \end{vmatrix} = (b-6)(-1+6) = 5(b-6)$$ and this is $\neq 0$.

Hence [A,B] will be of rank 3 while A is of rank 2. As the ranks are different, the equations will be inconsistent.

Summing up,

(i) If $a \neq 8$ and b has any value, the equations will be consistent and have a unique solution.

(ii) If $a = 8$ and $b = 6$, the equation will be consistent and have infinitely many solutions.

(iii) If $a = 8$ and $b \neq 6$, the equation will be inconsistent.

Exercises :

1. Solve or prove the inconsistency of the following systems of equation

(i)
$$\begin{aligned} x+2y+z &= 3 \\ 2x+3y+2z &= 5 \\ 3x-5y+5z &= 2 \\ 3x+9y-z &= 4 \end{aligned}$$

(ii)
$$\begin{aligned} x+2y-z &= 2 \\ 2x-3y+7z &= -1 \\ -x+y+3z &= 6 \\ 5z+y+3z &= 0 \end{aligned}$$

(iii)
$$\begin{aligned} 3x+y+z &= 8 \\ x-y+2z &= 5 \\ x+y+z &= 6 \\ 2x-2y+3z &= 7 \end{aligned}$$

(iv)
$$\begin{aligned} x+2y+2z &= 2 \\ 3x-2y-z &= 5 \\ 2x-5y+3z &= -4 \\ x+4y+6z &= 0 \end{aligned}$$

(v)
$$\begin{aligned} x+y+z &= 6 \\ 3x+y+z &= 8 \\ -x+y-2z &= -5 \\ -2x+2y-3z &= -7 \end{aligned}$$

(vi)
$$\begin{aligned} x+2y-z &= 3 \\ 3x-y+2z &= 1 \\ 2x-2y+3z &= 2 \\ x-y+z &= -1 \end{aligned}$$

(vii)
$$\begin{aligned} x+2y-5z &= -9 \\ 3x-y+2z &= 5 \\ 2x+3y-z &= 3 \\ 4x-5y+z &= -3 \end{aligned}$$

(viii)
$$\begin{aligned} x+y+z &= 1 \\ x+2y+3z &= 1 \\ x+3y+5z &= 7 \\ x+4y+7z &= 10 \end{aligned}$$

(ix)
$$\begin{aligned} x-2y-z-t &= -1 \\ 3x-2z+3t &= -4 \\ 5x-4y+t &= -3 \end{aligned}$$

(x)
$$\begin{aligned} x+y+z &= 7 \\ x+2y+3z &= 8 \\ y+2z &= 6 \end{aligned}$$

2. For what values of λ and μ the system of equations

$$\begin{aligned} x+y+z &= 6 \\ x+2y+3z &= 10 \\ x+2y+\lambda z &= \mu \end{aligned}$$

is (a) inconsistent (b) consistent (c) consistent and the solution in unique.

3. Investigate for what values of λ, μ the equations

$$\begin{aligned} x+y+z &= 6 \\ x+2y+3z &= 10 \\ x+2y+\lambda z &= \mu \end{aligned}$$

have (i) no solution (ii) a unique solution (iii) infinite number of solutions.

4. Discuss the solution of the equations

$$\begin{aligned} ax-2y+z &= 1 \\ x-2ay+z &= -2 \\ x-2y+az &= 1 \end{aligned}$$

determining when the system has no solution, one solution and infinity of solutions.

Answers

1. (i) consistent; $x = -1, y = 1, z = 2$

(ii) consistent; $x = -1, y = 2, z = 1$

(iii) consistent; $x = 1, y = 2, z = 3$

(iv) consistent; $x = 2, y = 1, z = -1$

(v) consistent; $x = 1, y = 2, z = 3$

(vi) consistent; $x = -1, y = 4, z = 4$

(vii) consistent; $x = 1/2, y = 3/2, z = 5/2$

(viii) consistent; $x = C-2, y = 3-2C, z = C$

(ix) inconsistent

(x) inconsistent.

2. If $\lambda = 3$ and $\mu \neq 10$, inconsistent

If $\lambda = 3$ and $\mu = 10$, consistant

If $\lambda \neq 3$, consistent and the solution is unique

3. (i) If $\lambda = 3$ and $\mu \neq 10$, equations will be inconsistent and hence no solution.

(ii) If $\lambda \neq 3$ and μ takes any value, equations will be consistent and have a unique solution.

(iii) If $\lambda \neq 3$ and $\mu = 10$, equations will be consistent and have an infinite number of solutions.

4. (i) No solution if $a = 1$

(ii) One solution if a does not take the values 1 and -2.

(iii) Infinitely of solutions if $a = -2$.

THEORY OF NUMBERS

7.1. PRIME AND COMPOSITE NUMBERS :

Prime number is an integer greater than one which has no divisors except itself and unity. Thus, 2, 3, 5, 7, 11, 13, 17, 19,..... are prime numbers.

Composite numbers which can be expressed as the product of two smaller integers. (i.e.,) A natural number which is neither a unit nor a prime is called a composite number.

Examples of composite numbers are 4, 12, 18, 15,.....

Two numbers which have no common divisor other than one are said tobe prime to one another. Thus 12 and 17, 32 and 63, 28 and 45 are prime to one another.

7.2. THE SIEVE OF ERATOSTHENES :

Prime numbers were a subject of great interest from early days. Eratosthenes who lived about 200 B.C. deviced a simple method to find the primes below a given number.

The method consists inwriting down all the integers up to the given number in their natural succession and then striking out all the multiples of 2, then the multiplesof 3, then those of 5 and so on.

If we want to determine the primes less than 500, it is not necessary to go beyond multiples of 23. This scheme wtih a little modification is used even today for the construction of tables of prime numbers.

The Sieve of Eratosthenes for prime less than 100 is given below.

1	2	3	~~4~~	5	~~6~~	7	~~8~~	~~9~~	~~10~~
11	~~12~~	13	~~14~~	~~15~~	~~16~~	17	~~18~~	19	~~20~~
~~21~~	~~22~~	23	~~24~~	~~25~~	~~26~~	~~27~~	~~28~~	29	~~30~~
31	~~32~~	~~33~~	~~34~~	~~35~~	~~36~~	37	~~38~~	~~39~~	~~40~~
41	~~42~~	43	~~44~~	~~45~~	~~46~~	47	~~48~~	~~49~~	~~50~~
~~51~~	~~52~~	53	~~54~~	55	~~56~~	~~57~~	~~58~~	59	~~60~~
61	~~62~~	~~63~~	~~64~~	~~65~~	~~66~~	67	~~68~~	~~69~~	~~70~~
71	~~72~~	73	~~74~~	~~75~~	~~76~~	~~77~~	~~78~~	79	~~80~~
~~81~~	~~82~~	83	~~84~~	~~85~~	~~86~~	~~87~~	~~88~~	89	~~90~~
~~91~~	~~92~~	~~93~~	~~94~~	~~95~~	~~96~~	97	~~98~~	99	~~100~~

The prime numbers below 100 are 1, 2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, 53, 59, 61, 67, 71, 73, 79, 83, 89 and 97.

7.1. Theorem :

The number of primes is infinite.

Proof :

Let the number of prime be finite, say n, and let those primes be p_1, p_2, p_3,....p_n. Then all the other numbers are composite and therefore should be exactly divisible by atleast one of the prime numbers $p_1,p_2,p_3,.....p_n$. Let us consider the number

$$A = p_1.p_2.p_3.....p_n+1$$

When this number is divided by p_1 or by p_2, or by p_3,..... or by p_n the reminder is 1. Hence the number A is not exactly divisible by any of the prime numbers $p_1,p_2,....p_n$.

(i.e.,) A is a prime number which is contrary to our assumption.

Hence our assumption that the number of primes, is finite, is wrong.

Hence the number of primes is infinite.

7.2.(a) The previous article gives us a method of constructing an infinite sequence of primes. We know that 2 and 3 are primes.

Hence $2\times3 + 1$ (i.e.) 7 is a prime

$2\times3\times5 + 1$ (i.e.) 31 is a prime

$2\times3\times5\times7 + 1$ (i.e.) 211 is a prime

By this method it is not possible to find all the primes. So attempts have been made to find some simple arithmetical formulae that give only primes even though they may not give all of them.

The following are some formulae which give prime numbers for certain value of n :

1) n^2+n+41 is a prime number if $n < 40$.

2) n^2+n+17 is a prime number if $n < 16$.

3) $2n^2+29$ is a prime number if $n < 29$.

4) $n^2-79n+1601$ is a prime number if $n < 80$.

5) $2^{2^n}+1$ is a prime number if $n < 5$.

7.2. Prime Number theorem :

The number of prime numbers less than or equal to a number N is usualy denoted by $\pi(N)$. Thus $\pi(1) = 1$; $\pi(2) = 2$; $\pi(3) = 3$; $\pi(4) = 3$; $\pi(5) = 4$; $\pi(6) = 4$.....

The distribution of prime numbers is very irregular and no exact formula has been discovered for $\pi(N)$, but it has been shown that $\lim_{N\to\infty} \frac{\pi(N)}{N/\log N} = 1$.

This result is known as `Prime Number theorem'.

7.2.(b) Every composite number can be resolved into prime factors and this can be done only in one way.

Let N be the composite number. Since the numbr is composite, it has a factor other than N and 1. Let it be a and the quotient when N is divided by a be b.

Then N = ab.

If a and b are not primes, we can find the divisors of a and b and express a and b in the form a = cd and b = ef.

$\therefore$ N = cdef.

Here a and b are less than N.

c and d are less than a

e and f are less than b.

$\therefore$ c, d, e and f are less than N.

Proceeding in this way we must come finally to factors which are prime numbers since the factors diminish at every stage.

Hence N can be expressed in the form N = pqr.... where p,q,r..... are all prime numbers, not necessarily different.

$\therefore$ N can be expressed as $N = p^a q^b r^c$..... where p, q, r.... are all primes and a, b, c..... integers.

Let N be resolved into prime factors in another way and let that be $P^AQ^BR^C$.... where P, Q, R..... are all primes and A, B, C,..... integers.

$\therefore N = p^a q^b r^c ..... = P^A Q^B R^C .....$

Since the prime P is a divisor of the product $p^a q^b r^c .....$ it is a divisor of one of the factors p, q, r.....

Since p, q, r..... are all primes p must be equal to one of them.

Similarly each one of P, Q, R,..... is equal to one of p, q, r.....

$\therefore p^a, q^b, r^c ..... = p^A q^B r^C .....$

If $A \neq a$, then let A be equal to a+k.

Since A and a are integers k is also an integer.

$\therefore \quad p^a q^b r^c ..... = p^{a+k} q^B r^C .....$

$\therefore \quad q^b r^c ..... = p^k q^B r^C .....$

p is a factor of the expression in the right side of the equation.

$\therefore$ p is a factor of the expression in the left side but this is impossible since the expression in the left side is prime to p.

$\therefore k = 0$. Hence A = a.

Similarly B = b, C = c,.....

Hence the factorisation of a composite number into product of primes is unique.

7.3. DIVISORS OF A GIVEN NUMBER N :

N can be expressed as the product of primes and let N be $p^a q^b r^c ....$, where p, q, r... are primes.

Let n be the number of divisors.

The divisors of N are the terms in the expansion of

$(1+p+p^2+.....+p^a)(1+q+q^2+.....+q^b)(1+r+r^2+....+r^c)....$

Hence the number of terms in the product will be the number of divisors and we can easily see that the number of divisors is (a+1)(b+1)(c+1).... The divisors include 1

and the number N itself. The sum of al the divisors is the sum of all the terms in the continued product.

$$\therefore S = \frac{p^{a+1}-1}{p-1}.\frac{q^{b+1}-1}{q-1}.\frac{r^{c+1}-1}{r-1}.....$$

Example 1 :

Find the number and sum of all the divisors of 360.

Solution :

$360 = 2^3.3^2.5^1$

The number of divisors = $(3+1)(2+1)(1+1) = 24$

Sum of the divisors = $\frac{2^4-1}{2-1}.\frac{3^3-1}{3-1}.\frac{5^2-1}{5-1}$

$= \frac{15}{1}.\frac{26}{2}.\frac{24}{4}$

$= 1170$

Example 2 :

Find the smallest number with 18 divisors.

Solution :

Let the number be N which is equal to $p^a\ q^b\ r^c$ .... where p, q, r are primes and a, b, c..... are integers.

Since we have to find the smallest numbers p, q, r..... must be as small as possible and a, b, c should be in the decending order of magnitude.

$\therefore N = 2^a\ 3^b\ 5^c$ ....

Number of divisors = (a+1) (b+1) (c+1).....

$\therefore 18 = (a+1)\ (b+1)\ (c+1).....$

but $18 = 2\times3\times3$

$\therefore$ We can take $c+1 = 2$, $b+1 = 3$, $a+1 = 3$

(i.e.,) $c = 1$, $b = 2$, $a = 2$

$\therefore N = 2^2\ 3^2\ 5^1 = 180$

Example 3 :

Find the product of all the divisors of N.

Solution :

Let N be expressed in the form $p^a q^b r^c$.....

The number of divisors is (a+1)(b+1)(c+1).....

If x is a divisor of N, then $\frac{N}{x}$ is also a divisor of N.

∴ All the divisors of N can be grouped into pairs whose product is N.

∴ (a+1)(b+1)(c+1)..... divisors can be grouped into $\frac{1}{2}(a+1)(b+1)(c+1)$..... pairs, the product of each pair being N.

∴ Product of the divisors = $N^{\frac{1}{2}(a+1)(b+1)(c+1)...}$

Exercises :

1. Find the number of divisors of 480 excluding 1 and 480.
2. Find the number of divisors of (i) 840 (ii) 1458 (iii) 288, excluding the number itself.
3. Verify that (i) 220 and 284, (ii) 17296 and 18416 are `amicable numbers' (i.e.,) that each is the sum of the divisors of the other (including 1 but excluding the number itself).
4. Show that the number of divisors of an integer is odd if and only if this integer is a square.
5. If N has n divisors including itself and 1, prove that their continued product is $\sqrt{N^n}$.
6. Find the smallest number (i) with 24 divisors, (ii) with 10 divisors.
7. Show that if 2^n-1 is a prime, then $2^{n-1}(2^n-1)$ is a perfect number and find the three least numbers given by the formula.

   (N is a perfect number if the sum of all its divisors excluding N is N).
8. Prove that the sum of reciprocals of the divisors of the perfect number $2^{n-1}(2^n-1)$ is 2.

8.1. EULER'S FUNCTION $\phi(N)$

The number of positive integers lessthan N and prime to it is denoted by $\phi(N)$. From the definition we get $\phi(2) = 1$, $\phi(3) = 2$, $\phi(4) = 2$, $\phi(5) = 4$, $\phi(6) = 2$. Even though $\phi(1)$ has no meaning as per this definition, we define it equal to 1.

Value of $\phi(N)$:

We have shown that N can be expressed as $N = p^a q^b r^c \ldots$ where p, q, r.... are all primes and a, b, c,.... integers. An integer will be prime to N iff it is not divisible by any of the primes p, q, r..... Therefore $\phi(N)$ is the number of itnegers in the series 1,2,3.....N which are not divisible by any of these primes. If we can find the numbers which are divisible by these primes in the series 1,2,.....N, we can get $\phi(N)$ by subtracting the numbers of such numbers from N. Numbers which are divisible by p in the series 1,2,....N are p, 2p, 3p,..... N/p and therefore there are N/p numbers which are divisible by p.

So also there are $\frac{N}{q}$ numbers which are divisible by q.

Similarly there are $\frac{N}{r}$ numbers which are divisible by r.

,, ,, ,, ,,

,, ,, ,, ,,

,, $\frac{N}{pq}$,, by pq

,, $\frac{N}{qr}$,, by qr.

,, $\frac{N}{pqr}$,, by pqr

and so on.

Consider the series $= \Sigma\frac{N}{p} - \Sigma\frac{N}{pq} + \Sigma\frac{N}{pqr} \ldots\ldots$ --------(1)

Consider any integer not greater than N.

Suppose that it is divisible by exactly k of the primes p, q, r.....

This number occurs kC_1 times in $\Sigma\frac{N}{p}$

,, kC_2 times in $\Sigma\frac{N}{pq}$

,, kC_3 times in $\Sigma\frac{N}{pqr}$

and so on.

Therefore the number of times it is counted in the expression (1) is $kC_1-kC_2+kC_3....$ which is equal to $1-(1-1)^k$ (i.e.,) 1.

Every integer (excluding unity) not greater than N and not prime to it is counted exactly once in (1). In evaluating $\phi(N)$ we count 1 also.

$$\therefore\quad \phi(N) = N-\Sigma\frac{N}{p}+\Sigma\frac{N}{pq}-\Sigma\frac{N}{pqr}+.....$$

$$= N\left(1-\Sigma\frac{1}{p}+\Sigma\frac{1}{pq}-\Sigma\frac{1}{pqr}+.....\right)$$

$$= N\left(1-\frac{1}{p}\right)\left(1-\frac{1}{q}\right)\left(1-\frac{1}{r}\right).....$$

Corollarly 1 :

If N = ab where a and b are prime to one another, then $\phi(N) = \phi(a).\phi(b)$.

Let $a = p_1^{a1}.p_2^{a2}.....p_k^{ak}$ and $b = q_1^{b1}.q_2^{b2}....q_r^{kr}$ where $p_1, p_2,.....p_k, q_1,q_2,....q_r$ are primes.

Since a and b are prime to one another $p_1,p_2,....p_k, q_1,q_2,.....q_r$ are different.

$$\phi(a) = a\left(1-\frac{1}{p_1}\right)\left(1-\frac{1}{p_2}\right).....\left(1-\frac{1}{p_k}\right)$$

$$\phi(b) = b\left(1-\frac{1}{q_1}\right)\left(1-\frac{1}{q_2}\right).....\left(1-\frac{1}{q_r}\right)$$

$$N = ab = p_1^{a1}p_2^{a2}.....p_k^{ak} \bullet q_1^{b1}.q_2^{b2}....q_r^{br}$$

$$\therefore\quad \phi(N) = N\left(1-\frac{1}{p_1}\right).....\left(1-\frac{1}{p_k}\right)\left(1-\frac{1}{q_1}\right)....\left(1-\frac{1}{q_r}\right)$$

$$= ab\left(1-\frac{1}{p_1}\right)\left(1-\frac{1}{p_2}\right)....\left(1-\frac{1}{p_k}\right)\left(1-\frac{1}{q_1}\right)....\left(1-\frac{1}{q_r}\right)$$

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

Corollary 2 :

If a, b, c, d,.....k are prime to one another, $\phi(abcd.....k) = \phi(a).\phi(b).....\phi(k)$.

Corollarly 3 :

If p is prime, then $\phi(p^r) = p^r\left(1-\frac{1}{p}\right)$

Example 1 :

Find the number of integers less than n and prime to it when n = 729 and 720.

We have $729 = 3^6$

$\therefore\ \phi(729) = 729\left(1-\frac{1}{3}\right) = 486$

$720 = 2^4.3^2.5$

$\therefore\ \phi(720) = 720\left(1-\frac{1}{2}\right)\left(1-\frac{1}{3}\right)\left(1-\frac{1}{5}\right)$

$= 192$

Example 2 :

Prove that the sum of the integers less than N and prime to it including unity is $\frac{1}{2}N\phi(N)$.

Let x be one of the integers less than N and prime to it. Then N–x is also prime to it.

$\therefore$ All the numbers less than N, prime to it can be grouped into pairs whose sum is N.

$\therefore\ \phi(N)$ can be grouped into $\frac{\phi(N)}{2}$ pairs the sum of each pair being N.

$\therefore$ Sum of the numbers $= \frac{N\phi(N)}{2}$.

Example 3 :

If $d_1, d_2, d_3, \ldots\ldots d_r$ (including 1 and N) are the divisors of N, then show that

$\phi(d_1)+\phi(d_2)+\ldots\ldots+\phi(d_r) = N$

Let $N = p_1^{a_1}p_2^{a_2}\ldots\ldots p_n^{a_n}$ where $p_1, p_2, \ldots p_n$ are primes and $a_1, a_2, a_3 \ldots$ are integers.

Every divisors is of the form $p_1^x, p_2^y, p_3^z, \ldots$ where x, y, z,.... take integral values from 0 to a_1, 0 to a_2, 0 to a_3,.... respectively. We know that,

$$\phi(p_1^x.p_2^y.p_3^z\ldots\ldots) = \phi(p_1^x)\phi(p_2^y)\phi(p_3^z)\ldots\ldots$$

Hence $\phi(d_1), \phi(d_2), \phi(d_3)\ldots\ldots\phi(d_r)$ are the terms in the expansion of

$[1+\phi(p_1)+\phi(p_1^2)+\ldots\ldots+\phi(p_1^{a_1})]$

$\times [1+\phi(p_2)+\phi(p_2^2)+\ldots\ldots+\phi(p_2^{a_2})]$

$\times \ldots\ldots$

$\times [1+\phi(p_n)+\phi(p_n^2)+\ldots\ldots+\phi(p_n^{a_n})]$ ---------(1)

$\therefore \phi(d_1)+\phi(d_2)+\ldots\ldots+\phi(d_r)$

= Continued product of the expression --------(1)

$$1+\phi(p_1)+\phi(p_1^2)+\ldots\ldots+\phi(p^{a_1}) = 1+p_1\left(1-\frac{1}{p_1}\right)+p_1^2\left(1-\frac{1}{p_1}\right)+\ldots.+p_1^{a_1}\left(1-\frac{1}{p_1}\right)$$

$$= 1+p_1-1+p_1^2-p_1+p_1^3-p_1^2+\ldots.+p_1^{a_1}-p_1^{a_1-1}$$

$$= p_1^{a_1}$$

Similarly,

$1+\phi(p_2)+\phi(p_2^2)+\ldots\ldots+\phi(p_2a^2) = p_2^{a_2}$ and so on.

$\therefore \phi(d_1)+\phi(d_2)+\ldots\ldots+\phi(d_r) = p_1^{a_1}.p_2^{a_2}\ldots\ldots p_n^{a_n} = N$

Exercises :

1. How many numbers including unity are less than 210 and prime to it?
2. Find the sum of the positive integers including unity which are less than 600 and prime to it.
3. How many numbers are there less than 500 which are not divisible by 2, 3 or 5?

4. If a is prime to N, show that the number of terms of the arithmetical progression x, x+a, x+2a,....., x+(n–1)a which are prime to N is $\phi(N)$.

5. Show that if N be any number and a, b, c.....be its different prime factors, then the sum of all the numbers less than N and prime to N is

$$\frac{N^2}{2}\left(1-\frac{1}{a}\right)\left(1-\frac{1}{b}\right)\left(1-\frac{1}{c}\right)....$$

and the sum of the squares of all such numbers is

$$\frac{N^3}{3}\left(1-\frac{1}{a}\right)\left(1-\frac{1}{b}\right)\left(1-\frac{1}{c}\right)....+\frac{N}{6}(1-a)(1-b)(1-c).....$$

6. Show that the arithmetic mean of all numbers less than N and prime to it (including unity) is N/2.

8.2. INTEGRAL PART OF A REAL NUMBER :

The integral part of a number x is denoted by the symbol [x].

For example,

$$\left[4\frac{2}{3}\right] = 4;\ [6] = 6,\ \left[\frac{2}{5}\right] = 0,\ \left[\sqrt{2}\right] = 1,\ \left[-\sqrt{2}\right] = -2$$

The "fractional part" is considered to be positive. From the definition, the following properties are easily deduced.

(1) $[x] \le x < [x]+1$

(2) $[x+a] = [x] + a$ if a is an integer

(3) $[x+y] \ge [x] + [y]$

8.3. THE HIGHEST POWER OF A PRIME p CONTAINED IN N! :

n! = 1.2.3.....n

If n<p, there is no number in n! which is divisible by p.

If n≥p, then n! contains numbers which are divisible by p.

The factors in n! which will be divisible by p are p, 2p 3p,.....$\left[\frac{n}{p}\right]$p.

Hence the highest power of p in n! is the highest power of p in the product.

(i.e.,) in $p^{\left[\frac{n}{p}\right]}$ 1.2.3.....$\left[\frac{n}{p}\right]$

But in 1.2.3.$\left[\frac{n}{p}\right]$ the prime p is a factor in the numbers p, 2p, 3p.....$\left[\frac{n}{p^2}\right]$p.

∴ Power of p in the product 1.2.3.....$\left[\frac{n}{p}\right]$ is the power of p in the product p.2p.3p....$\left[\frac{n}{p^2}\right]$p.

(i.e.,) $p^{\left[\frac{n}{p^2}\right]}$ 1.2.3.....$\left[\frac{n}{p^2}\right]$

Hence the highest power of p in n! is the highest power of p in the product $p^{\left[\frac{n}{p}\right]}.p^{\left[\frac{n}{p^2}\right]}$ 1.2.....$\left[\frac{n}{p^2}\right]$

(i.e.,) in $p^{\left[\frac{n}{p}\right]+\left[\frac{n}{p^2}\right]}$ 1.2.3.....$\left[\frac{n}{p^2}\right]$

In the same way we find the highest power of p in n! is the highest power of p in the product $p^{\left[\frac{n}{p}\right]+\left[\frac{n}{p^2}\right]+\left[\frac{n}{p^3}\right]}$.1.2.3.....$\left[\frac{n}{p^3}\right]$

(i.e.,) in the product $p^{\left[\frac{n}{p}\right]+\left[\frac{n}{p^2}\right]+\left[\frac{n}{p^3}\right]+\left[\frac{n}{p^4}\right]}$.1.2.3.....$\left[\frac{n}{p^4}\right]$

$\left[\frac{n}{p}\right],\left[\frac{n}{p^2}\right],\left[\frac{n}{p^3}\right]\cdots$ from a decreasing sequence and hence there exists a positive integer k where $\left[\frac{n}{p^k}\right]=0$.

Hence if we continue this process we will get the highest power of

$$p = \left[\frac{n}{p}\right] + \left[\frac{n}{p^2}\right] + \left[\frac{n}{p^3}\right] + \ldots + \left[\frac{n}{p^{k-1}}\right] \text{ where} \left[\frac{n}{pk}\right] = 0$$

Example 1 :

Find the highest power of 3 dividing 1000!

$$\left[\frac{1000}{3}\right] = 333$$

$$\left[\frac{1000}{3^2}\right] = \left[\frac{333}{3}\right] = 111$$

$$\left[\frac{1000}{3^3}\right] = \left[\frac{111}{3}\right] = 37$$

$$\left[\frac{1000}{3^4}\right] = \left[\frac{37}{3}\right] = 12$$

$$\left[\frac{1000}{3^5}\right] = \left[\frac{12}{3}\right] = 4$$

$$\left[\frac{1000}{3^6}\right] = \left[\frac{4}{3}\right] = 1$$

The highest power of 3 in 1000! is 333+111+37+12+4+1 = 498

Thus 3^{498} is the highest power of 3 dividing 1000!

Example 2 :

With how many zeros does 79! end?

Let us find the highest power of 2 and 5 in 79!

$$\left[\frac{79}{2}\right] = 39, \quad \left[\frac{79}{2^2}\right] = 19, \quad \left[\frac{79}{2^3}\right] = 9,$$

$$\left[\frac{79}{2^4}\right] = 4 \quad \left[\frac{79}{2^5}\right] = 2, \quad \left[\frac{79}{2^6}\right] = 1$$

Thus 2^{74} is the highest power of 2 in 79!, $\left[\frac{79}{5}\right] = 15, \quad \left[\frac{79}{5^2}\right] = 3.$

Thus 5^{18} is the highest power of 5 in 79!

∴ The highest power of 10 in 79! is 18.

∴ 79! will end in 18 zeros.

8.4. THE PRODUCT OF r CONSECUTIVE INTEGERS IS DIVISIBLE BY r! :

Let n+1, n+2,.....,n+r be the consecutive integers.

Product of these integers = (n+1)(n+2).....(n+r)

$$= \frac{(n+r)!}{n!}$$

We have to prove that $= \frac{(n+r)!}{n!}$ is divisible by r!

(i.e.) $\frac{(n+r)!}{n!r!}$ is an integer.

(i.e.) We have to show that the highest power of a prime p occuring in (n+r)! is greater than the highest power of p in n!r!. The highest power of any prime p occurring in (n+r)!, n!, r! are respectively.

$$\left[\frac{n+r}{p}\right]+\left[\frac{n+r}{p^2}\right]+.....$$

$$\left[\frac{n}{p}\right]+\left[\frac{n}{p^2}\right]+....$$

$$\left[\frac{r}{p}\right]+\left[\frac{r}{p^2}\right]+.....$$

∴ Highest power of p occuring in n!r! is

$$\left[\frac{n}{p}\right]+\left[\frac{r}{p}\right]+\left[\frac{n}{p^2}\right]+\left[\frac{r}{p^2}\right]+....$$

We know that $[x+y] \geq [x]+[y]$

(i.e.) $$\left[\frac{x+y}{p^k}\right] \geq \left[\frac{x}{p^k}\right]+\left[\frac{y}{p^k}\right]$$

By putting x = n, y = r, k = 1,2,.....

We get $\left[\frac{n+r}{p}\right] \geq \left[\frac{n}{p}\right]+\left[\frac{r}{p}\right]$

$\left[\frac{n+r}{p^2}\right] \geq \left[\frac{n}{p^2}\right]+\left[\frac{r}{p^2}\right]$

....

....

$\therefore \left[\frac{n+r}{p}\right]+\left[\frac{n+r}{p^2}\right]+... \geq \left[\frac{n}{p}\right]+\left[\frac{r}{p}\right]+\left[\frac{n}{p^2}\right]+\left[\frac{r}{p^2}\right]+.....$

(i.e.) the prime p enters in the numerator (n+r)! in powers not lower than in the denominators n! r!.

$\therefore$ The numerator is divisible by the denominator.

$\therefore$ (n+1)(n+2).....(n+r) is divisible by r!.

Corollary 1 :

$\frac{n!}{a!b!c!...}$ is an integer when a+b+c+..... = n.

Corollary 2 :

If n is a prime, n_{c_r} is divisible by n.

$$n_{c_r} = \frac{n(n-1)(n-2)....(n-r+1)}{r!}$$

The numerator is a product of r consecutive integers.

$\therefore$ It is divisible by r!

n is prime to r.

$\therefore$ It is prime to r!

$\therefore$ (n–1)(n–2).... (n–r+1) is divisible by r!

$\therefore$ n_{c_r} = n × an integer since $n_{c_r} = \frac{n(n-1)....(n-r+1)}{r!}$

$\therefore$ n_{c_r} is divisible by n.

Example :

Show that $n(n+1)(2n+1)$ is divisible by 6.

$$n(n+1)(2n+1) = n(n+1)(2n+4-3)$$
$$= 2n(n+1)(n+2)-3n(n+1)$$

$n(n+1)$ is divisible by 2.

$n(n+1)(n+2)$ is divisible by 3! (i.e.) 6.

Hence the expression is divisible by 6.

Note :

$n(n+1)(2n+1)$ is a multiple of 6.

This is usually written as $n(n+1)(2n+1) = M(6)$.

Exercises :

1. Find the highest powers of 2, 5, 7, 11, 13 contained in 1000!
2. With how many zeros does (i) 61! (ii) 257! and (iii) 82! end?
3. If n is any odd number, show that $n(n^2-1)$ is divisible by 24.
4. Show that $n(n^2-1)(29n^2+4) = M(120)$
5. Show that $n^5-n = M(30)$
6. Show that $n(n-1)(n+25)(n+50) = M(24)$
7. Show that the greatest power of n in $(n^r-1)!$ is $\frac{n^r - nr + r - 1}{n-1}$.
8. Show that if n is odd (i) $(n^2+3)(n^2+7) = M(32)$ (ii) $n^4+4n^2+11 = M(16)$
9. Show that if n is a positive integer $(n+1)(n+2).....(n+n)$ is divisible by 2^n.
10. If n be an odd prime, show that $(a+1)^n-(a^n+1) = M(2n)$
11. If $n = 2^a+2^b+2^c+....$ m terms where $a<b<c<....$ show that the greatest power of that must divide n! is (n–m).
12. Show that if n be prime greater than 3. $n(n^2-1)(n^2-4)(n^2-9) = M(2^7.3^2.5.7)$

Answers :

1. 994; 249; 164; 98; 81
2. (i) 14; (ii) 63; (iii) 19.

9.1. CONGRUENCES

Two integers a and b are called congruent with respect to the modulus m if an integer k exists such that a–b = km.

k may be positive, zero or negative. The congruence is denoted by $a \equiv b \pmod m$ (or) by $a-b \equiv 0 \pmod m$.

For example, $18 \equiv 4 \pmod 7$

$13 \equiv 28 \pmod 5$

$14-4 \equiv 0 \pmod 5$

If two numbers are congruent with respect to the modulus m, each is called a residue of the other to the modulus m.

Every residue of a to the modulus is of the form a+km where k may be positive or negative.

Congruences with the same moduli possess many properties of equialities. Some of them are given below.

1. If $a \equiv b \pmod m$ and $a_1 \equiv b_1 \pmod m$ and if q, r are integers, then $qa+ra_1 = qb+rb_1 \pmod m$

$a \equiv b \pmod m$ ∴ $a-b = km$

$a_1 \equiv b_1 \pmod m$ ∴ $a_1-b_1 = k_1m$

$qa+ra_1 = q(b+km)+r(b_1+k_1m)$

$= qb+rb_1+m(qk+rk_1)$

$= qb+rb_1+M(m)$

∴ $qa+ra_1 \equiv qb+rb_1 \pmod m$

Corollary :

If $a \equiv b \pmod m$; $a_1 = b_1 \pmod m$

$a+a_1 \equiv b+b_1 \pmod m$

$a-a_1 \equiv b-b_1 \pmod m$

In general if $a \equiv b \pmod m$

$a_1 \equiv b_1 \pmod m$, $a_2 \equiv b_2 \pmod m$

then $a+a_1+a_2+\ldots = b+b_1+b_2+\ldots \pmod m$

2. If $a \equiv b \pmod m$; $a_1 = b_1 \pmod m$, then $aa_1 = bb_1 \pmod m$

$$a = b \pmod m$$

$\therefore$ $$a = b+km$$

$$a_1 = b_1 \pmod m$$

$\therefore$ $$a_1 = b_1+k_1m$$

$$aa_1 = (b+km)(b_1+k_1m)$$

$$= bb_1+m(kb_1+k_1b+kk_1m)$$

$\therefore$ $$aa_1 = bb_1 \pmod m$$

Corollary 1 :

If $a \equiv b \pmod m$, $a_1 \equiv b_1 \pmod m$, $a_2 \equiv b_2 \pmod m$ then $aa_1a_2\ldots \equiv bb_1b_2\ldots (\text{mod } m)$.

Corollary 2 :

If $a \equiv b \pmod m$, then $a^n = b^n \pmod m$

Corollary 3 :

If $a \equiv b \pmod m$, then $f(a) \equiv f(b) \pmod m$, if $f(x)$ is a polynomial in x.

3. These result show that congruences may be manipulated as regards addition, subtraction and multiplication with integral numbers, just like equations. As regards division a modification is necessary.

If $ax \equiv bx \pmod m$ and if h is H.C.F. of x, m then $a \equiv b \bmod \left(\frac{m}{h}\right)$

$x = ph$, $m = qh$ where p, q are co-prime.

$ax \equiv bx \pmod m$

$\therefore$ $ax-bx = km$

(i.e.,) $aph-bph = kqh$

(i.e.) $a-b = k.\frac{q}{p}$

q is prime to p.

$\therefore$ a–b has q as a factor.

$a-b = M(q) = M\left(\frac{m}{h}\right)$

$\therefore a \equiv b \bmod \left(\frac{m}{h}\right)$

Corollary :

If $h = 1$, $a \equiv b \pmod m$

Thus the rule of concellation holds for congruences on the condition that the cancelled factor is relatively prime to the modulus.

4. If $a \equiv b \pmod{m_1}$, $a \equiv b \pmod{m_2}$, $a \equiv b \pmod{m_3}$..... $a \equiv b \pmod{m_n}$, then $a \equiv b \pmod m$, where m is the least multiple of $m_1, m_2, \ldots, m_n$.

$a \equiv b \pmod{m_1}$

$\therefore$ $a-b$ = a multiple of m_1

|||ly $a-b$ = a multiple of m_2

$a-b$ = a multiple of m_3

....

$a-b$ = a multiple of m_n

To satisfy these equations a–b = a multiple of m where m is the least common multiple of $m_1, m_2, \ldots m_n$.

$\therefore a \equiv b \pmod m$.

9.2. CRITERIA OF DIVISIBILITY OF NUMBER :

We can derive the criteria of divisibility of a number by 3, 9, 11 from the properties of congruences.

Let N be the number and let the digits in the units, tens, hundreds....place be a,b,c,d,.....

Then N = a+10b+100c+1000d+.....

1) $10 \equiv 1 \pmod 3$

$100 \equiv 1 \pmod 3$

$1000 \equiv 1 \pmod 3$

....

$\therefore N = a+b+c+d+..... \pmod 3$

$\therefore$ N is divisible by 3 iff the sum of its digits is divisible by 3.

2) $10 \equiv 1 \pmod 9$

$100 \equiv 1 \pmod 9$

$1000 \equiv 1 \pmod 9$

....

$\therefore N = a+b+c+d.... \pmod 9$

Hence the number N is divisible by 9 iff the sum of its digit is divisible by 9.

3) We have $10 \equiv -1 \pmod{11}$

$100 \equiv 1 \pmod{11}$

$1000 = -1 \pmod{11}$

....

$\therefore N = a-b+c-d+..... \pmod{11}$

Hence N is divisible by 11 iff the alternate sum a–b+c–d+.... of the digits is divisible by 11 or if a–b+c–d+.... = 0 (i.e.,) a+c+..... = b+d+.... (i.e.) the sum of the odd digits is equal to sum of the even digits.

Corollary :

Since congruent numbers leave the same remainder when divided by the modulus, the preceding congruences can be used in finding remainders in divisions by 3, 9, 11.

Example 1 :

Find a number having the remainders 5, 4, 3, 2 when divided by 6, 5, 4, 3 respectively.

Solution :

Let N be that number.

$N \equiv 5 \pmod 6$ (i.e.,) $N \equiv -1 \pmod 6$

$N \equiv 4 \pmod 5$ (i.e.,) $N \equiv -1 \pmod 5$

$N \equiv 3 \pmod 4$ (i.e.) $N \equiv -1 \pmod 4$

$N \equiv 2 \pmod 3$ (i.e.) $N \equiv -1 \pmod 3$

$\therefore N \equiv -1 \pmod{\text{L.C.M. of } (6, 5, 4, 3)}$

$\equiv -1 \pmod{60}$

$\therefore$ The least value of N is 59.

N can take any value $-1+60k$ where k is a positive integer.

Example 2 :

Find the remainder when 9^{10} is divided by 11.

Solution :

We have $9^2 \equiv 4 \pmod{11}$ -------(1)

$\therefore$ $9^8 \equiv 256 \pmod{11}$

Also $256 \equiv 3 \pmod{11}$

$\therefore$ $9 \equiv 3^8 \pmod{11}$ -------(2)

From (1) & (2) $9^{10} \equiv 4.3 \pmod{11}$

$9^{10} \equiv 1 \pmod{11}$

Hence the remainder is 1.

Example 3 :

Show that $13^{2n+1}+9^{2n+1}$ is divisible by 22.

Solution :

$$13^{2n+1} = 13.(13)^{2n} = 13.(169)^n$$

$$= 13.(22\times7+15)^n$$

$$\equiv 13.(15)^n \pmod{22}$$

$$9^{2n+1} = 9.81^n = 9(3\times22+15)^n$$

$$\equiv 9.15^n \pmod{22}$$

$\therefore$ $13^{2n+1}+9^{2n+1} \equiv [13.15^n+9.15^n] \pmod{22}$

$$\equiv 22.15^n \pmod{22}$$

$$\equiv 0 \pmod{22}.$$

$\therefore$ $13^{2n+1}+9^{2n+1}$ is divisible by 22.

Example 4 :

Find the remainder when 2^{1000} is divisible by 17.

Solution :

$$2^4 = 16$$

$$\equiv -1 \pmod{17}$$

$$(2^4)^{250} \equiv (-1)^{250} \pmod{17}$$

$$\equiv 1 \pmod{17}$$

$\therefore$ The remainder when 2^{1000} is divided by 17 is 1.

Example 5 :

Find the remainder obtained by dividing 2^{46} by 47.

$$2^5 \equiv 32$$

$$\equiv -15 \pmod{47}$$

$$2^{10} \equiv (-15)^2 \pmod{47}$$

$$\equiv 225 \pmod{47}$$

$$\equiv -10 \pmod{47}$$

$$2^{20} \equiv (-10)^2 \pmod{47}$$

$$\equiv 100 \pmod{47}$$

$$\equiv 6 \pmod{47}$$

$$2^{40} \equiv 36 \pmod{47}$$

$$\equiv -11 \pmod{47}$$

$$2^{45} \equiv 2^{40}.2^5$$

$$\equiv -11.-15 \pmod{47}$$

$$\equiv 165 \pmod{47}$$

$$\equiv 24 \pmod{47}$$

$$2\times2^{45} \equiv 48 \pmod{47}$$

$$\equiv 1 \pmod{47}$$

∴ The remainder is 1.

Example 6 :

If p is a prime number and p>0, then $(a+b)^p \equiv (a^p+b^p) \pmod{p}$.

Solution :

$$(a+b)^p = a^p+p_{c_1}a^{p-1}b+.....+p_{c_{p-1}}ab^{p-1}+b^p$$

Now, $$p_{c_r} = \frac{p!}{r!(p-r)!} = \frac{p(p-1)!}{r!(p-r)!}$$ is an integer.

Since r<p and p is prime, r! and (p–r)! do not divide p.

Hence $\frac{(p-1)!}{r!(p-r)!}$ is an integer.

∴ p divides p_{c_r}

∴ $(a+b)^p = a^p+p^k+b^p$ where $k \in z$

∴ $(a+b)^p \equiv (a^p+b^p) \pmod{p}$

Example 7 :

A natural number n is divisible by 3 iff the sum of its digits is divisble by 3.

Solution :

Let a, b, c, d, ... be the digits in the units, tens, hundreds, thousands,... places.

Then $n = a+10b+100c+1000d+.....$

Now

$1 \equiv 1 \pmod 3$ $\therefore a \equiv a \pmod 3$

$10 \equiv 1 \pmod 3$ $\therefore 10b \equiv b \pmod 3$

$100 \equiv 1 \pmod 3$ $\therefore 100c \equiv c \pmod 3$

$1000 \equiv 1 \pmod 3$ $\therefore 1000d \equiv d \pmod 3$

....

$n \equiv (a+b+c+d.....) \pmod 3$

$\therefore$ n is divisible by 3 iff a+b+c+d is divisible by 3.

Exercises :

1.a. Find the least two positive integers having the remainders 2,3,2 when divided by 3,5,7 respectively.

b. Find the least positive number which when divided by 7, 8, 9 will leave remainders 1,2,3 respectively. Find the general formula for such numbers.

2. Find a multiple of 7 which has remainder 1 when divided by 2,3,4,5 or 6.

3. Prove that $3^{4n+2}+5^{2n+1}$ is divisible by 14.

4. Show that $3^{2n+1}+2^{n+2}$ is divisible by 7.

5. Show that $7^{2n+1}+1 = M(8)$

6. Show that $19^{2n}-1 = M(360)$

7. Show that $23^{2n}-1 = M(528)$

8. Show that $17^{2n}-1= M(288)$

9. Show that $3^{2n+4}-2^{2n} = M(5)$

10. Show that $3^{2n-1}+2^{n+1}$ is divisible by 7.

Answers :

1. (a) 128, 233 (b) 498; (504)k–6.

2. 301

9.3. NUMBERS IN ARITHMETIC PROGRESSION :

If x, x+a, x+2a,..... x+(n–1)a, n terms of an arithmetical progression, are divided by n, where n is prime to a, the remainders are numbers 0,1,2,.....n–1 taken in certain order.

Let us assume that x+pa and x+qa when divided by n leave equal remainder r.

Then $x+pa = k_1n+r$

$x+qa = k_2n+r$

$(p-q)a = (k_1-k_2)n$

∴ (p–q) is less than n.

(p–q) is less than n.

∴ a is divisible by n which is contrary to the hypothesis that n is prime to a.

The remainders are all different and as each is less than n, they must be the numbers 0,1,2,.....n–1, taken in some order or other.

Corollary 1 :

One of the numbers x, x+a, ...,x+(n–1)a is divisible by n.

Corollary 2 :

If the progression is continued beyond the n^{th} term, the remainders recur inthe same order.

Corollary 3 :

If p is prime to a, then when a, 2a, 3a,.....(p–1)a are divided by p, the remainders are 1,2,.....,(p–1) in some order or other.

Example :

Show that every integer which is a perfect cube is of the form 7p or 7p±1.

An integer, N when it is divided by 7 has one of the remainders 0, 1,2,3,4,5,6. Every integer has one of the forms.

7m, 7m+1, 7m+2, 7m+3, 7m+4, 7m+5, 7m+6

(i.e.,) $7m, 7m+1, 7m+2, 7m+3, 7m-3, 7m-2, 7m-1$

(i.e.,) $7m, 7m\pm1, 7m\pm2, 7m\pm3$

$$(7m\pm r)^3 = (7m)^3\pm3(7m)^2r+3(7m)r^2\pm r^3$$
$$= M(7)\pm r^3$$

Hence in the four possible cases we have

$$N^3 = (7m)^3 = M(7)$$
$$N^3 = (7m\pm1)^3 = M(7)\pm1$$
$$N^3 = (7m\pm2)^3 = M(7)\pm8 = M(7)\pm1$$
$$N^3 = (7m\pm3)^3 = M(7)\pm27 = M(7)\pm28\mp1$$
$$= M(7)\pm1$$

In every case therefore the cube has one or other of the form $7p$ or $7p\pm1$.

Theorem 9.1 :

If x be congruent with r with respect to the modulus m, f(x) will be congruent with f(r) with respect to modulus m where f(x) is a polynomial in x.

Let f(x) be $p_0+p_1x+p_2x^2+.....+p_nx^n$

x is congruent with r with respect to modulus m.

∴ x is of the form qm+r

By the binomial expansion, we have

$$(qm+r)^n = (qm)^n+n_{c_1}(qm)^{n-1}r+.....n_{c_{n-1}}(qm)r^{n-1}+r^n$$
$$= (q^nm^{n-1}+n_{c_1}q^{n-1}m^{n-2}r+.....n_{c_{n-1}}qr^{n-1})m+r^n$$
$$= M(m)+r^n$$

Similarly,

$$(qm+r)^{n-1} = M(m)+r^{n-1} \text{ and so on.}$$

Hence if $x = qm+r$

$$f(x) = f(qm+r)$$

$$= p_0+p_1(qm+r)+p_2(qm+r)^2+.....+p_n(qm+r)^n$$

$$= p_0+p_1r+p_2r^2+.....+p_nr^n+M(m)$$

$$= f(r)+M(m)$$

Hence f(x) is congruent with f(r) with respect to modulus m.

Corollary :

Since all integers are congruent (with respect to modulus m) with one or other of the series 0,1,2,.....m–1, it follows that to test the divisibility of f(x) by m for all integral values of x we need only test the divisibility by m of f(0), f(1), f(2),.....f(m–1).

Example 1 :

Show that $x^5–x$ is divisible by 30.

Solution :

$$f(x) = x^5–x = x(x^4–1)$$

$$f(0) = 0$$

$$f(1) = 0$$

$$f(2) = 2\times15 = 30$$

$$f(3) = 3\times80 = 240$$

$$f(4) = 4\times255 = 1020$$

$$f(5) = 5\times624 = 3120$$

f(0), f(1), f(2), f(3), f(4) are divisibly by 5.

∴ f(x) is divisible by 5.

f(0), f(1), f(2), f(3), f(4), f(5) are divisible by 6.

∴ f(x) is divisible by 6.

∴ f(x) is divisible by both 5 and 6.

(i.e.) f(x) is divisible by 30.

Example 2 :

If x, y, z be three consecutive integers, show that $(\Sigma x)^3-3\Sigma x^3$ is divisible by 108.

Since x, y, z are consecutive integers y = x+1, z = x+2

$\therefore$ $(\Sigma x)^3-3\Sigma x^3 = (3x+3)^3-3\{x^3+(x+1)^3+(x+2)^3\}$

$= 18x^3+54x^2+36x$

$= 18x(x+1)(x+2)$

Let f(x) be x(x+1)(x+2)

$f(0) = 0$

$f(1) = 6$

$f(2) = 24$

f(0), f(1) are divisible by 2.

$\therefore$ f(x) is divisible by 2.

f(0), f(1), f(2) are divisible by 3.

$\therefore$ f(x) is divisible by 3.

$\therefore$ f(x) is divisible by 6.

$\therefore$ $(\Sigma x)^3-3\Sigma x^3$ is divisible by 108.

Exercises :

1. Show that every square is of the form 3m or 3m+1
2. Show that every square is of the form 5m or 5m±1
3. Show that every cube is of the form 9m (or) 9m±1
4. Show that every forth power is of the form 5m (or) 5m+1
5. Show that $2^{2x+1}+1$ is divisible by 3.
6. Show that in order that x^4+1 may be divisible by 17, x should be of the form 17m±2 (or) 17m±8.
7. If n is a prime number greater than 3. Show that n^2-1 is divisible by 24.
8. If n is a prime number greater than 7, show that n^6-1 is divisible by 504. [Show that n^6-1 is divisible by 7, 8 and 9]

9.4. FERMAT'S THEOREM :

If p is a prime and a is any number prime to p then $a^{p-1}-1$ is divisible by p.

We have proved that if n is a prime number, then n_{c_r} is divisible by n.

$$(a+1)^p = a^p+p_{c_1}.a^{p-1}+p_{c_2}.a^{p-2}+.....+p_{c_r}.a^{p-r}+....p_{c_{p-1}}.a+1$$

(i.e.,) $$(a+1)^p-(a^p+1) = p_{c_1}.a^{p-1}+p_{c_2}.a^{p-2}+.....+p_{c_r}.a^{p-r}+.....+p_{c_{p-1}}.a$$

Since p isprime $p_{c_1}, p_{c_2},.....p_{c_{p-1}}$ are divisible by p.

$\therefore$ $(a+1)^p-(a^p+1)$ = a multiple of p

$\therefore$ $(a+1)^p \equiv (a^p+1) \pmod p$

Since this result is true for all values of a, replacing a by a–1, a–2, a–3,.....3,2,1 in succession we get

$$a^p \equiv [(a-1)^p+1] \pmod p \quad \text{--------(1)}$$

$$(a-1)^p \equiv [(a-2)^p+1] \pmod p \quad \text{--------(2)}$$

$$(a-2)^p \equiv [(a-3)^p+1] \pmod p \quad \text{--------(3)}$$

....

....

$$3^p \equiv [2^p+1] \pmod p \quad \text{--------(a–2)}$$

$$2^p \equiv [1^p+1] \pmod p \quad \text{--------(a–1)}$$

Adding all the equations (1), (2), (3).... (a–1) we get

$$a^p+(a-1)^p+(a-2)^p+.....+3^p+2^p \equiv [(a-1)^p+1+(a-2)^p+1+(a-3)^p+1+..... +2^p+1+1^p+1] \pmod p$$

Since we are adding a–1 equations, we have

$a^p+(a-1)^p+(a-2)^p+.....+3^p+2^p \equiv [(a-1)^p+(a-2)^p+....+2^p+1^p+a-1] \pmod p$

(i.e.,) $a^p \equiv [1+(a-1)] \pmod p$

(i.e.,) $a^p \equiv a \pmod p$

(i.e.,) $a^p - a$ is divisible by p

(i.e.,) $a(a^{p-1}-1)$ is divisible by p.

Since a is prime to p, $a^{p-1}-1$ is divisible by p.

II Method :

$$(x+y)^p = x^p+p_{c_1}x^{p-1}y+p_{c_2}x^{p-2}y^2+....+p_{c_{p-1}}xy^{p-1}+y^p$$

$p_{c_1}, p_{c_2},.....p_{c_{p-1}}$ are divisible by p.

$$\therefore (x+y)^p \equiv (x^p+y^p) \pmod p$$

$$(x+y+z)^p \equiv (x+y)^p+p_{c_1}(x+y)^{p-1}z+....+z^p$$

$$\equiv [(x+y)^p+z^p] \pmod p$$

$$\equiv (x^p+y^p+z^p) \pmod p$$

So in general

$$(x+y+z+....+w)^p \equiv (x^p+y^p+z^p+.....+w^p) \pmod p$$

where x, y, z,.....w are any integers.

Let there be a integers x, y, z,.....w.

Put each equal to 1.

Then $(1+1+.....\text{a terms})^p = (1^p+1^p+...+1^p) \pmod p$

$$\therefore a^p \equiv a \pmod p$$

(i.e.,) a^p-a is divisible by p.

$\therefore a^{p-1}-1$ is divisible by p.

III Method :

When a, 2a, 3a,..... (p–1)a are divided by p, the remainders are 1, 2,....p–1 in a certain order since p is prime to a.

$$a \equiv r_1 \pmod p$$

$$2a \equiv r_2 \pmod p$$

$$3a \equiv r_3 \pmod p$$

....

....

$$(p-1)a \equiv r_{p-1} \pmod p$$

Here $r_1, r_2, r_3, \ldots r_{p-1}$ are 1, 2, 3,.....p–1 in a certain order.

$\therefore$ $a.\ 2a.\ 3a \ldots (p-1)a \equiv r_1\ r_2\ r_3 \ldots\ldots r_{p-1} \pmod p$

(i.e.,) $(p-1)!\ a^{p-1} \equiv 1.2.3 \ldots (p-1) \pmod p$

(i.e.,) $(p-1)!\ a^{p-1} \equiv (p-1)! \pmod p$

(i.e.,) $(p-1)!\ a^{p-1} - (p-1)! \equiv 0 \pmod p$

$\therefore$ $(p-1)!\ (a^{p-1}-1)$ is divisible by p.

But (p–1)! is not divisible by p, since p is prime.

$\therefore$ $a^{p-1}-1$ is divisible by p.

Corollary 1 :

$a^p - a$ is divisible by p if p is prime and a is prime to p.

Corollary 2 :

If p is an odd prime and a is prime to p, then $a^{\frac{1}{2}(p-1)} \pm 1$ is divisible by p.

$$a^{p-1}-1 = \left\{a^{\frac{1}{2}(p-1)} - 1\right\}\left\{a^{\frac{1}{2}(p-1)} + 1\right\}$$

$a^{p-1}-1$ is divisible by p.

$a^{\frac{1}{2}(p-1)} - 1$ or $a^{\frac{1}{2}(p-1)} + 1$ is divisible by p.

(i.e.,) $a^{\frac{1}{2}(p-1)} \pm 1$ is divisible by p.

Example :

1. Show that if x and y are both prime to the prime number n, then $x^{n-1}-y^{n-1}$ is divisible by n. Deduce that $x^{12}-y^{12}$ is divisible by 1365.

Solution :

$x^{n-1}-1 \equiv 0 \pmod n$ since x is prime to n and n is prime.

Similarly $y^{n-1}-1 \equiv 0 \pmod n$

subtracting we get $x^{n-1}-y^{n-1} \equiv 0 \pmod n$

$$x^{12}-y^{12} = x^{13-1}-y^{13-1}$$

$$\equiv 0 \pmod{13}$$

$\therefore$ $x^{12}-y^{12} = (x^6-y^6)(x^6+y^6)$

but $x^6-y^6 \equiv 0 \pmod{7}$

x^6-y^6 is divisible by 7.

(i.e.,) $x^{12}-y^{12}$ is divisible by 7.

$$x^{12}-y^{12} = (x^4-y^4)(x^8+x^4y^4+y^8)$$

by $x^4-y^4 \equiv 0 \pmod{5}$

$\therefore$ $x^{12}-y^{12}$ is divisible by 5.

$$x^{12}-y^{12} = (x^2-y^2)(x^4+x^2y^2+y^4)(x^6+y^6)$$

but $x^2-y^2 \equiv 0 \pmod{3}$

$\therefore$ $x^{12}-y^{12}$ is divisible by 3.

$\therefore$ $x^{12}-y^{12}$ is divisible by 13, 7, 5 and 3.

(i.e.,) it is divisible by $13\times7\times5\times3$

(i.e.,) by 1365

Example 2 :

Show that the 8th power of any number is of the form 17m or $17m\pm1$

Solution :

Let the number be N.

N may be prime to 17 or may not be prime to 17.

If N is not prime to 17, it must be a multiple of 17, since 17 is a prime number.

In that case N is a multiple of 17.

$\therefore$ N^8 is of the form 17m.

If N is prime to 17,

$N^{17-1}-1$ is divisible by 17.

(i.e.) N^{16-1} is divisible by 17.

(i.e.) $(N^8+1)(N^8-1)$ is divisible by 17.

$\therefore$ N^8+1 or N^8-1 is divisible by 17.

$\therefore$ N^8+1 or N^8-1 is a multiple of 17.

$\therefore$ $N^8+1 = 17m$ (or) $N^8-1 = 17m$

(i.e.,) $N^8 = 17m \pm 1$

Hence N^8 is one of the forms $17m$ (or) $17m \pm 1$.

Another Method :

Since 17 is a prime number.

$a^{17} \equiv a \pmod{17}$

$\therefore \quad \frac{17}{a^{17}-a}$

$\therefore \quad \frac{17}{a\left(a^8-1\right)\left(a^8+1\right)}$

$\therefore \quad \frac{17}{a}$ (or) $\frac{17}{a^8-1}$ (or) $\frac{17}{a^8+1}$

Hence a^8 is of the form $17m$ (or) $17m \pm 1$.

Example 3 :

Prove that the 5th power of any integer N has the same units digit as N.

By Fermat's theorem, N^5-N is divsible by 5.

$\therefore \quad N^5-N = M(5)$

(i.e.) $\quad N(N^4-1) = M(5)$

(i.e.)$N(N^2+1)(N+1)(N-1) = M(5)$

Since N is any integer, either N or N–1 is divisible by 2.

$\therefore \quad N^5-N =$ a multiple of 10.

(i.e.,) $\quad N^5 = M(10)+N$

N can be put in the form $10p+q$ where $q<10$

$\therefore \quad N^5 = M(10)+10p+q$

$= M(10)+q$

$\therefore$ N^5 has the same units digit q as that of N.

Example 4 :

Find the remainder when 2^{460} is divided by 47.

Solution :

$$2^{46} \equiv 1 \pmod{47} \text{ (Fermat's theorem)}$$

$\therefore$ $(2^{46})^{10} \equiv 1^{10} \pmod{47}$

$\therefore$ $2^{460} \equiv 1 \pmod{47}$

Hence the remainder is 1.

Example 5 :

Find the remainder when 2^{1000} is divided by 13.

Solution :

$$2^{12} \equiv 1 \pmod{13} \text{ Fermat's theorem)}$$

$\therefore$ $(2^{12})^{83} \equiv 1 \pmod{13}$

$\therefore$ $(2^{996}) \equiv 1 \pmod{13}$

Also, $2^4 \equiv 3 \pmod{13}$

$\therefore$ $2^{1000} \equiv 3 \pmod{13}$

Hence the remainder is 3.

Example 6 :

If a and b are prime to n, show that $\frac{n}{\left(a^{n-1}-b^{n-1}\right)}$

Solution :

Since (a, n) = 1, by Fermat's theorem

$$a^{n-1} \equiv 1 \pmod{n}$$

Similarly $b^{n-1} \equiv 1 \pmod{n}$

Hence $a^{n-1}-b^{n-1} \equiv 0 \pmod{n}$

$\therefore$ $\frac{n}{\left(a^{n-1}-b^{n-1}\right)}$

Exercise :

1. Show that n^5-n is divisible by 30.
2. Show that n^7-n is divisible by 42.
3. Show that $n^{13}-n$ is divisible by 2730.
4. Show that if n is any prime number greater than 19, then $n^{18}-1$ is divisible by 9576.
5. Show that the 4^{th} power of any number is of the form 5m (or) 5m+1.
6. Show that the 12^{th} power of any number is of the form 13m (or) 13m+1.
7. Show that the 9^{th} power of any number is one of the forms i9m or 19m±1.
8. Show that if n be a prime number $1^{n-1}+2^{n-1}+3^{n-1}+....+(n-1)^{n-1}+1 = M(n)$
9. Show that if m and n are primes, then $m^{n-1}+n^{m-1}-1 \equiv 0 \pmod{mn}$
10. Show that if m, n and p are prime, then $(np)^{m-1}+(pm)^{n-1}+(mn)^{p-1}-1 \equiv 0 \pmod{mnp}$
11. If n be prime and > x show that $x^{n-2}+x^{n-3}+x^{n-4}+....+x+1 \equiv 0 \pmod{n}$
12. If n be an odd prime, show that $1+2(n+1)+2^2(n+1)^2+...+2^{n-2}(n+1)^{n-2} \equiv 0 \pmod{n}$
13. If n be odd, show that $1^n+2^n+....+(n-1)^n \equiv 0 \pmod{n}$
14. If n is of the form 4m+1, show that a^n ends with the same digit as a for all values of a.

    [Hint : $f(m) = a(a^{4m}-1)$, $f(0) = 0$; $f(1) = a(a^4-1)$; $f(2 = a(a^8-1)$; $f(3) = a(a^{12}-1)$; $f(4) = a(a^{16}-1)$

    $a(a^4-1)$ is a factor of f(0), f(1), f(2), f(3), f(4). a^4-1 is divisible by 5.

    $a(a^4-1) = a(a+1)(a-1)(a^2+1)$ is so divisible by 2.

    ∴ $a(a^4-1)$ is divisible by 10.

    ∴ f(m) is divisible by 10].

10.1. GENERALIZATION OF FERMAT'S THEOREM

If n is any number and a is prime to n then $a^{\phi(n)} \equiv 1 \pmod n$

Let $a_1, a_2, \ldots . a_{\phi(n)}$ be the $\phi(n)$ integers less than n and prime to it.

Consider the products $aa_1, aa_2, \ldots . aa_{\phi(n)}$

Let k be the remainder when aa_r is divided by n.

Then $aa_r = M(n) + k$ -----(1)

Since a and a_r are prime to n and k is also prime to n. Hence the remainder is prime to n.

Suppose the product aa_s gives the same remainder k when it is divided by n.

Then $aa_s = M(n)+k$ -----(2)

From (1)&(2), we get $a(a_r - a_s) = M(n)$

(i.e.) $a(a_r - a_s)$ is a multiple of n which cannot be the case since $a_r - a_s < n$ and a is prime to n.

∴ aa_r and aa_s will not give the same remainder when they are divided by n. Hence when the products $aa_1, aa_2, \ldots . aa_{\phi(n)}$ are divided by n the remainders $k_1, k_2, \ldots k_n$ are all different and prime to n.

∴ They are $a_1, a_2, \ldots . a_{\phi(n)}$ is some order.

$$aa_1 \equiv k_1 \pmod n$$

$$aa_2 \equiv k_2 \pmod n$$

. . . . . .

. . . . . .

$$aa_{\phi(n)} \equiv k_{\phi(n)} \pmod n.$$

∴ $aa_1, aa_2, \ldots aa_{\phi(n)} \equiv k_1 k_2 \ldots k_{\phi(n)} \pmod n$

(ie) $a^{\phi(n)} a_1 a_2 \ldots . a_{\phi(n)} \equiv a_1 a_2 \ldots . . a_{\phi(n)} \pmod n$

Dividing by $a_1, a_2, \ldots a_{\phi(n)}$ which is prime to n.

we get

$$a^{\phi(n)} \equiv 1 \pmod n$$

Cor :

If n a is prime number, then $\phi(n) = n-1$. Then this theorem reduces to Fermat's theorem.

Exercises :

1. If $\alpha\beta,\gamma.....$be primes and $N=\alpha\beta,\gamma.....$then show that $\Sigma\left(\frac{N}{\alpha}\right)^{\alpha-1} \equiv 1(\text{mod}\alpha\beta,\gamma.....)$

2. Show that $16^{99} \equiv 1(\text{mod } 437)$

   [Hint $16^{99} = 2^{4\times 99} =$ and $\phi(437) = 437\left(1-\frac{1}{19}\right)\left(1-\frac{1}{23}\right)$ and use extension of Fermat's theorem]

3. Show that $x^{pq-q} \equiv 1 \text{mod}(pq)$ it x is prime to P, P is prime and $q = p^n$

4. Find the remainder obtained in dividing 2^{460} by 47.

5. When p+1 and 2p+1 are both prime numbers, show that $x^{2P}-1$ is divisible by 8(p+1)(2p+1) where x is prime to 2, p+1 and 2p+1.

10.2. WILSON'S THEOREM :

If P is a prime number, then (P–1)!+1 is divisible by p.

If a is any number of the series 1,2,....,(P-1) where p is a prime then when a,2a,3a,....(P -1) a are divided by p the remainders are 1,2,3....P–1 in some order or other

Hence there is only one number say, a_1 among the numbers 1,2,...P–1 such that when aa_1 is divided by p, the remainder is 1.

$\therefore aa_1 \equiv 1(\text{mod } p)$

Such two numbers are called associate residues.

Suppose $a = a_1$ then $a^2 \equiv 1(\text{mod } p)$

(ie) $a^2-1 \equiv 0 \text{ (mod p)}$

(ie) $(a+1)(a-1) \equiv 0 \text{ (mod p)}$

$\therefore$ Either a+1 is divisible by p(or) a = 1.

Since a is less than p, a+1 = p (or) a = 1.

(ie) a = p–1 (or) 1.

Hence numbers which are identical with their associate residues are 1 and p–1. Excluding these 2 numbers 1 and p–1, the remaining numbers 2,3,4.....p–2 can be grouped into $\frac{p-3}{2}$ pairs of associate residues such that the product of each pair is congruent with 1.

$\therefore$ 2..3..(p–2) $\equiv$1(mod p) —1

we also have 1. (p–1) $\equiv -1$(mod p) —2

Multiplying (1)&(2) we get,

1,2,3.....(p–2) (p–1) $\equiv -1$ (mod p)

(ie) (p–1)!+1 $\equiv$ 0(mod p)

(ie) (p–1)!+1 is divisible by p.

10.3. LAGRANGE'S THEOREM

If $(x+1)(x+2).....(x+p-1) = x^{P-1}+A_1x^{P-2}+....A_{P-2}x+A_{P-1}$

and P be prime, then $A_1,A_2,.....A_{P-2}$ are all divisible by p.

Changing x into x+1..we get

$$(x+2)(x+3)....(x+p) = (x+1)^{P-1}+A_1(x+1)^{P-2}+.....+A_{P-2}(x+1)+A_{P-1}.$$

$\therefore$ $(x+1)(x+2)(x+3)....(x+p)$

$$= (x+1)\{(x+1)^{P-1}+A_1(x+1)^{P-2}+..+A_{P-2}(x+1)+A_{P-1}\}$$

$$= (x+1)^P+A_1(x+1)^{P-1}+...+A_{P-2}(x+1)^2+A_{P-1}(x+1)$$

$\therefore$ $(x+p)\{x^{P-1}+A_1x^{P-2}+A_2x^{P-3}+....+A_{P-2}x+A_{P-1}\}$

$$= (x+1)^P+A_1(x+1)^{P-1}+...+A_{P-2}(x+1)^2+A_{P-1}(x+1)$$

(ie) $x^P+A_1x^{P-1}+A_2x^{P-2}+...+A_{P-2}x^2+A_{P-1}x+p.x^{P-1}+A_1px^{P-2}+....+A_{P-2}p.x+pA_{P-1}$

$$= (x+1)^P+A_1(x+1)^{P-1}+....+A_{P-2}(x+1)^2+A_{P-1}(x+1)$$

$\therefore$ $\{(x+1)^P-x^P\}+A_1\{(x+1)^{P-1}-x^{P-1}\}+A_2\{(x+1)^{P-2}x^{P-2}\}+...+A_{P-1}\{(x+1)-x\}$

$$= px^{P-1}+A_1px^{P-2}+.....+A_{P-2}x+A_{P-1}$$

Equating the coefficient of $x^{P-2},x^{P-3},.....$

we get

$$pA_1 = pC_2+(p-1)C_1A_1$$

$$pA_2 = pC_3 + (p-1)C_2 + A_1 (p-2)C_1.A_2$$

$$pA_3 = pC_4+(p-1)C_3 A_1+(p-2)C_2 A_2+(p-3)C_1A_3$$

. . . . . .

. . . . . .

$$PA_{P-1} = 1+A_1+A_2+.....+A_{P-2}+A_{P-1}$$

Since $(p-1)C_1$, $(p-2)C_2$, $(p-3)C_3$,....are not divisible by p, if p is prime, we get by successive steps that $A_1,A_2,A_3,....A_{P-2}$ are all divisible by P.

Cor 1 :

$$(x+1)\ (x+2).....(x+p-1) = x^{P-1}+A_1x^{P-2}+....+A_{P-2}x+A_{P-1}$$

Put x = 0, we get $A_{P-1} = (p-1)!$

Put x = 1, we get $2,3....p = 1+A_1+.....+A_{P-2}+A_{P-1}$

$\therefore$ $A_{P-1}+1 = p!-(A_1+.....+A_{P-2})$

(ie) $(p-1)! + 1 = p!- (A_1+.....+A_{P-2})$

The left side is divisible by p.

$\therefore$ (p–1)! + 1 is divisible by p.

This is **wilson's theorem.**

Cor 2 :

$$x(x+1)\ (x+2)....(x+p-1) = x(x^{P-1}+A_1x^{P-2}+A_2x^{P-3}+....+A_{P-2}x+A_{P-1})$$

$$= x^P+A_1x^{P-1}+A_2x^{P-2}+.....+A_{P-2}x^2+A_{P-1}x$$

$$= (x^P-x)+A_1x^{P-1}+A_2x^{P-2}+....+A_{P-2}x^2+(A_{P-1}+1)x$$

$\therefore$ $x^P-x = x(x+1)(x+2).....(x+p-1)-\{A_1x^{P-1}+A_2x^{P-3}+.....$

$$+A_{P-2}x^2\} - (A_{P-1}+1)x$$

$x(x+1)\ (x+2).....(x+p-1)$ being the product of p consecutive integers, must be divisible by P. Also if p be prime $A_{P-1}+1,A_1,A_2,...A_{P-2}$ are divisible by p.

$\therefore$ $x^P - x$ is divisible by p if p be prime.

This is **Fermat's theorem:**

Cor 3 :

If p is a prime greater than 3, then A_{P-2} is a multiple of p^2.

In the lagrange's theorem substitute the values of p and $-2p$ instead of x

we get

$$(p+1)\,(p+2).....(2p-1) = A_{P-1}+A_{P-2}p+.....+p^{P-1} \qquad \text{-------(1)}$$

$$(-2p+1)\,(-2p+2).....(-2p+p-1)$$

$$= (-2p)^{P-1}+A_1(-2p)^{P-2}+....+A_{P-2}(-2p)+A_{P-1}$$

$$\text{(ie) } (2p-1)\,(2p-2).....(p+2)\,(p+1)\,(-1)^{P-1}$$

$$= (-1)^{P-1}\{(2p)^{P-1} - A_1(2p)^{P-2}+.....+A_{P-1}\}$$

Since there are P terms on the right side and P is odd.

$$\therefore (p+1)\,(p+2).....(2_{P-1}) = A_{P-1}-2pA_{P-2}.....+(2p)^{P-1} \qquad \text{-------(2)}$$

Subtracting (2) from (1) we get

$0 = 3p.\ A_{P-2} - 3P^2.\ A_{P-3}$ + a multiple of p^4.

Since A_{P-1}, A_{P-2}, A_{P-3},.....are divisible by p.

$$\therefore \qquad 3P\ A_{P-2} = 3P^2A_{P-3} - M(P^4).$$

$$\therefore \qquad A_{P-2} = pA_{P-3} - \frac{1}{3}M(P^3)$$

A_{P-3} is a multiple of p and $p>3$.

A_{P-2} is a multiple of p^2.

Cor : 4

$$(p-1)!\left(1+\frac{1}{2}+\frac{1}{3}+....+\frac{1}{P-1}\right) \text{ is a multiple of } P^2.$$

In cor. 3 we have learnt that $A_{P-2} = M(p^2)$

$$A_{P-2} = \text{Coeff. of x in } (x+1)\,(x+2)....(x+p-1)$$

$$= (2.3.....p-1) + (1.3.4...p-1) + ...(1+2....p-2)$$

$= (p-1)!\ (1+\frac{1}{2}+\frac{1}{3}+...+\frac{1}{P-1})$

Hence the result.

Example 1 :

Show that (18)! + 1 is divisible by 437

Solution :

19 is a prime number.

$\therefore$ (19–1)!+1 is divisible by 19

(ie) (18)!+1 is divisible by 19.

23 is a prime number.

$\therefore$ (22)!+1 is divisible by 23.

(ie) $22.21.20.19\ (18)!+1 = M(23)$

(ie) $(23-1)\ (23-2)\ (23-3)\ (23-4)\ (18)!+1 = M(23)$

(ie) $\{M(23) + 1.2.3.4\}\ (18)!+1 = M(23)$

(ie) $\{M(23) +23+1\}\ (18)!+1 = M(23)$

(ie) $\{M(23)+1\}\ (18)!+1 = M(23)$

(ie) $\{M(23)\ (18)! + (18)!\ \} + 1 = M(23)$

$\therefore$ (18)!+1 is divisible by 23

$\therefore$ (18)!+1 is divisible by 19×23

(ie) (18)!+1 is divisible by 437.

Example 2 :

If P is a Prime number and p = 4m+1 where m is a positive integer, prove that $\{(2m)!\}^2+1$ is divisible by p.

Since p is a prime number,

$(p-1)!+1 \equiv 0 \pmod{P}$

(ie) $(4m+1-1)!+1 = 0 \pmod{P}$

(ie) $(4m)!+1 \equiv 0 \pmod{P}$

(ie) $(2m)!(2m+1)(2m+2)....(4m)+1 \equiv 0 \pmod{P}$

(ie) $(2m)!\{(p-2m)(p-2m-1).....(p-1)\}+1 \equiv 0 \pmod p$

(ie) $(2m)!\{M(p)+(2m)!\}+1 \equiv 0 \pmod p$

(ie) $M(p)+\{(2m)!\}^2+1 \equiv 0 \pmod p$

$\therefore$ $\{(2m)!\}^2+1 \equiv 0 \pmod p$

Example 3 :

If $M = 1.3.5....(p-2)$ where p is an odd prime, show that $M^2 \equiv (-1)^{\frac{P+1}{2}} \pmod p$.

P is a prime number.

$\therefore$ $(p-1)!+1 \equiv 0 \pmod p$

(ie) $1.2.3....(p-1)+1 \equiv 0 \pmod p$

(ie) $1.3.5....(p-2).2.4.6.....(p-1)+1 \equiv 0 \pmod p$

(ie) $1.3.5....(p-2)\{p-p-2)(p-p-4)...(p-1)\}+1 \equiv 0 \pmod p$

(ie) $1.3.5....(p-2)\{M(p)+(-1)^{\frac{P-1}{2}}(p-2)(p-4)...3.1\}+1 \equiv 0 \pmod p$

(ie) $1.3.5....(p-2)M(p)+(-1)^{\frac{P-1}{2}}1^2.3^2.5^2....(p-2)^2+1 \equiv 0 \pmod p$

$\therefore$ $(-1)^{\frac{P-1}{2}}1^2.3^2.5^2.....(p-2)^2+1 \equiv 0 \pmod P$.

Multiplying throughout by $(-1)^{\frac{p-1}{2}}$

we get,

$$(-1)^{P-1}1^2.3^2.5^2.....(p-2)^2+(-1)^{\frac{p-1}{2}} \equiv 0 \pmod P$$

p is an odd prime

$\therefore$ p–1 is even

$\therefore$ $M^2+(-1)^{\frac{p-1}{2}} \equiv 0 \pmod p$

(ie) $M^2-(-1)^{\frac{p+1}{2}} \equiv 0 \pmod p$

(ie) $M^2 \equiv (-1)^{\frac{(p+1)}{2}} \pmod p$

Example 4 :

Show that if n is a prime number and r<n, $(n-r)!(r-1)!+(-1)^{r-1} \equiv 0 \pmod n$

Deduce that $\left\{\left[\frac{1}{2}(n-1)\right]!\right\}^2 + (-1)^{\frac{n-1}{2}} \equiv 0 \pmod n$

n is a prime number

By Wilson's theorem, $(n-1)!+1 \equiv 0 \pmod n$

(ie) $(n-r)!(n-r+1)(n-r+2)....(n-1)+1 \equiv 0 \pmod n$

(ie) $(n-r)!\left(n-\overline{r-1}\right)\left(n-\overline{r-2}\right).....(n-1)+1 \equiv 0 \pmod n$

(ie) $(n-r)!\{M(n)-(-1)^{r-1}(r-1)!\}+1 \equiv 0 \pmod n$

where M(n) is a multiple of n.

(ie) $M(n)\,(n-r)!-(-1)^{r-1}\,(n-r)!(r-1)!+1 \equiv 0 \pmod n$

$\therefore (-1)^{r-1}\,(n-r)!\,(r-1)!+1 \equiv 0 \pmod n$

Multiplying throughout by $(-1)^{r+1}$, we get

$(-1)^{2r}\,(n-r)!\,(r-1)! + (-1)^{r+1} \equiv 0 \pmod n$

(ie) $(n-r)!(r-1)!+(-1)^{r+1} \equiv 0 \pmod n$

Put $r = \frac{n+1}{2}$ in the above result.

we get, $\left(n-\frac{n+1}{2}\right)!\left(\frac{n+1}{2}-1\right)!+(-1)^{\frac{(n-1)}{2}} \equiv 0 \pmod n$

(ie) $\left(\frac{n-1}{2}\right)!\left(\frac{n-1}{2}\right)!+(-1)^{\frac{(n-1)}{2}} \equiv 0 \pmod n$

(ie) $\left\{\left(\frac{n-1}{2}\right)!\right\}^2 +(-1)^{\frac{(n-1)}{2}} \equiv 0 \pmod n$

Example 5 :

Show that $28! + 233 \equiv 0 \pmod{899}$

Solution :

$899 = 29.31$

Now $28! + 1 \equiv 0 \pmod{29}$ (Wilson's theorem)

$\therefore$ $28!+1+29(8) \equiv 0 \pmod{29}$

$\therefore$ $28!+233 \equiv 0 \pmod{29}$ -------(1)

Now $30!+1 \equiv 0 \pmod{31}$ (Wilson's theorem)

$\therefore$ $30.29.28!+1 \equiv 0 \pmod{31}$

$\therefore (31-1)(31-2)28!+1+31 \equiv 0 \pmod{31}$

$\therefore$ $(-1)(-2)\,28!+32 \equiv 0 \pmod{31}$

$\therefore$ $28!+16 \equiv 0 \pmod{31}$ (since (2,31) = 1)

$\therefore$ $28!+16+7.31 \equiv 0 \pmod{31}$

$28!+233 \equiv 0 \pmod{31}$ -------(2)

From (1)&(2) we get $28!+233 \equiv 0 \pmod{899}$

Since (31,29) = 1.

Example 6 :

Show that $7^{2n}+16n-1 \equiv 0 \pmod{64}$

Solution :

$$7^{2n}+16n-1 = (1-8)^{2n}+16n-1$$
$$= (1-2n_{c_1}.8+2n_{c_2}.8^2-.....+8^{2n})+16n-1$$
$$= 1-16n + (\text{a multiple of } 64)+16n-1$$
$$= \text{a multiple of } 64.$$

Hence $7^{2n}+16n-1 \equiv 0 \pmod{64}$.

Example 7 :

Show that $3^{2n+2}+2^{n+1} \equiv 0 \pmod{7}$

Solution :

The result is true for n = 1

Let it be true when n = m. Then

$$3^{2m+1}+2^{m+2} \equiv 0 \pmod{7}$$

Now, $3^{2(m+1)+1}+2^{(m+1)+2} = 9.3^{2m+1}+2.2^{m+2}$

$= 2[3^{2m+1}+2^{m+2}]+7.3.^{2m+1}$

$=$ a multiple of 7.

Hence the result follows by induction.

Example 8:

Prove that for any integer n, n^5-n is divisible by 30.

Solution :

$n^5-n = n(n^4-1) = (n-1)n(n+1)(n^2+1)$

Now n–1, n, n+1 are three consecutive integers and hence the product (n–1)(n+1)n is divisible by 6.

Therefore n^5-n is divisible by 6 -----(1)

Also by Fermat's theorem, $n^4-1 \equiv 0 \pmod 5$

Hence n^4-1 is divisible by 5.

$\therefore$ n^5-n is divisible by 5 -----(2)

Now, since 5 and 6 are relatively prime, n^5-n is divisible by 30.

Exercises :

1. If p is a prime number, show that 2(p–3)!+1 is divisible by p.
2. Prove that $712!+1 = 0 \pmod{719}$
3. Show that (1) 28!+233 is divisible by 899, (2) $28! \equiv 666 \pmod{899}$
4. Show that $18!-22 \equiv M(46)$
5. Show that any prime of the form 4n+1 is a divisor of a number of the form $1+k^2$.
6. If P is a prime of the form 4m–1, show that $\{(2m-1)!\}^2-1 \equiv 0 \pmod P$.
7. If P is an odd prime, show that $(1,2,3,.....\frac{P-1}{2})^2 + (-1)^{p-1/2} \equiv 0 \pmod P$

[Hint : If P is an odd prime.

$$\frac{P+1}{2} = P-\frac{P-1}{2} \equiv -\frac{P-1}{2} \pmod P$$

$$\frac{P+3}{2} \equiv P - \frac{P-3}{2} \cdot \frac{p-3}{2} \pmod{P}$$

. . . . . .

. . . . . .

$$p-1 \equiv p\ 1 \equiv -1 \pmod{P}$$

$$\therefore \quad \frac{p+1}{2}\ \frac{p+3}{2}\ldots.(p\ 1) \equiv (-1)^{\frac{(P-1)}{2}}\ 1,2,3\ldots\ldots(\text{mod } P)$$

$$(p\ 1)!+1 \equiv 0 \pmod{P}.$$

(ie) $1,2,3,\ldots\ldots\frac{p-1}{2}.\frac{p+1}{2}.\frac{p+3}{2}\ldots\ldots(p\ 1)+1 \equiv 0 \pmod{P}$].