Institute of Distance and Open Learning (IDOL)
University of Mumbai.

M.Sc. (Mathematics), SEM- I

Paper - I

ALGEBRA – I

PSMT101

CONTENT

***

SYLLABUS

Unit I. Dual spaces

1. Vector spaces over a field, linear independence, basis for finite dimensional and infinite dimensional vector spaces and dimension.

2. Kernel and image, rank and nullity of a linear transformation, rank-nullity theorem (for finite dimensional vector spaces), relationship of linear transformations with matrices, invertible linear transformations. The following are equivalent for a linear map T : V → V of a finite dimensional vector space V:

(a) T is an isomorphism.

(b) ker T = {0}.

(c) Im(T) = V.

3. Linear functionals, dual spaces of a vector space, dual basis (for finite dimensional vector spaces), annihilator in the dual space of a subspace W of a vector space V and dimension formula, a k-dimensional subspace of an n-dimensional vector space is intersection of n−k many hyperspaces. Double dual of a Vector space V and canonical embedding of V into is isomorphic to V when V is of finite dimension.

(ref:[1] Hoffman K. and Kunze R.).

4. Transpose of a linear transformation T. For finite dimensional vector spaces: rank ( ) =rank T, range( ) is the annihilator of kernel (T), matrix representing . (ref:[1] Hoffman K and Kunze R)

Unit II. Determinants & Characteristic Polynomial

Rank of a matrix. Matrix of a linear transformation, change of basis, similar matrices. Determinants as alternating -forms, existence and uniqueness, Laplace expansion of determinant, determinants of products and transposes, adjoint of a matrics. determinants and invertible linear transformations, determinant of a linear transformation. Solution of system of linear equations using Cramer's rule. Eigen values and Eigen vectors of a linear transformation, Annihilating polynomial, Characteristic polynomial, minimal polynomial, Cayley-Hamilton theorem.

(Reference for Unit II: [1] Hoffman K and Kunze R, Linear Algebra).

Unit III. Triangulation of matrices

Triangulable and diagonalizable linear operators, invariant subspaces and simple matrix representation (for finite dimension).

(ref: [5] N.S. Gopalkrishnan & [3] Serge Lang)

Nilpotent linear transformations on finite dimensional vector spaces, index of a Nilpotent linear transformation. Linear independence of
where N is a nilpotent linear transformation of index k ≥ 2 of a vector space V and u ∈ V with

(Ref: [2] I.N.Herstein).

For a nilpotent linear transformation N of a finite dimensional vector space V and for any subspace W of V which is invariant under N, there exists a subspace of V such that .

(Ref:[2] I.N.Herstein).

Computations of Minimum polynomials and Jordan Canonical Forms for 3×3-matrices through examples.

(Ref:[6] Morris W. Hirsch and Stephen Smale).

Unit IV. Bilinear forms

1. Inner product spaces, orthonormal basis, Gram-Schmidt process.
2. Adjoint of a linear operator on an inner product space, unitary operators, self adjoint operators, normal operators.

   (ref:[1] Hoffman K and Kunze R).

Spectral theorem for a normal operator on a finite dimensional complex inner product space. (ref:[4] Michael Artin, Ch. 8).

Spectral resolution (examples only). (ref:[1] Hoffman K and Kunze R, sec 9.5).

3. Bilinear form, rank of a bilinear form, non-degenerate bilinear form and equivalent statements.
   (ref:[1] Hoffman K and Kunze R).
4. Symmetric bilinear forms, orthogonal basis and Sylvester's Law, signature of a Symmetric bilinear form. (ref:[4] Michael Artin).

1

LINEAR EQUATIONS

Unit Structure :

1.0 INTRODUCTION

Linear word comes from line. You know the equation of a straight line in two dimensions has the form $ax + by = c$. This is a linear equation in two variables x any y. solving this equation means to find x and y in $\mathbb{R}$. which satisfied $ax + by = c$. The geometric interpretation of the equation is that the set of all points satisfying the equation forms a straight line in the plane through the point $(c/c, o)$ and with slope $-a/bc$. In this chapter, we shall review the theory of such equations in n variables and interpret the solution geometrically.

1.1 OBJECTIVES

After going through this chapter, you will be able to :

- Understand the characteristic of the solutions.
- Solve if the equations are solvable.
- Interpret the system geometrically.

1.2 SYSTEMS OF LINEAR EQUATIONS

The collection of linear equations :

$$\begin{aligned} a_{11}x_1 + \ldots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + \ldots + a_{2n}x_n &= b_2 \\ &\vdots \\ am_1x_1 + \ldots + am_nx_n &= b_m \end{aligned}$$

is called a system of m linear equations in n unknowns $x_1,...,x_n$. Here $a_{ij}, b_i \in R$ are given. We shall write this in a short form as

$$\sum_{j=1}^{n} a_{ij} x_j = b_i, 1 \le i \le m \text{ ………………….. (1.2.1)}$$

Solving this system means to find real numbers $x_1,...,x_n$ which satisfy the system. Any n-tuple $(x_1,...,x_n)$ which satisfies the system is called a solution of the system. If $b_1 = b_2 = - = b_m = 0$, we say that the system is homogeneous which can be written is short form, in

$$\sum_{j=1}^{n} a_{ij} x_j = 0, 1 \le i \le m \text{ …………………… (1.2.2)}$$

Note that $\mathbf{0} = (0, \ldots, \mathbf{0})$ always satisfies (1.2.2). This solution is called the trivial solution. We say $(x_1,...,x_n)$ is a nontrivial if $(x_1,...,x_n) \neq (0,..,0)$. That is if there exits it least one such that $x_i \neq 0$.

Perhaps the most fundamental technique for finding the solutions of a system of linear equations is the technique of elimination. We can illustrate this technique on the homogenous system

$$\begin{aligned} 2x_1 - x_2 + x_3 &= 0 \\ x_1 + 3x_2 + 4x_3 &= 0 \end{aligned}.$$

If we added (-2) times the second equation to the first equation we obtain $-7x_2 - 7x_3 = 0$ i.e. $x_2 = -x_3$. Similarly eliminating x_2 from the above equation, we obtain $7x_1 + 7x_3 = 0$ i.e. $x_1 = -x_3$. So we conclude that if (x_1, x_2, x_3) is a solution then $x_1 = x_2 = -x_3$. Thus the set of solutions consists of all triples *(a, a, -a)*. Let $(a_1,..,a_n)$ be a solution of the homogeneous system *(1.2.2)*. Then we see that $(\alpha a_1,...,\alpha a_n)$ is again a solution of *(1.2.2)* for any $\alpha \in \mathbb{R}$. This has the following geometric interpretation in the case of the three dimensional space $\mathbb{R}^3$.

Let *(r, s, t)* be a solution of the system $ax + by + cz = 0$. That is, $ar + bs + ct = 0$. Then the solution set is a plane through the origin. So the plane contains *(0, 0, 0)* and *(r, s, t)*. The line joining these two points is $\frac{x-0}{0-r} = \frac{t-0}{0-s} = \frac{z-0}{0-t} = (-\alpha)$ (say) i.e.

$x = \alpha r, y = \alpha s, z = \alpha t$ is again a solution of the system *ax + by + cz = 0.*

Also if $(b_1,....,b_n)$ is another solution of *(1.2.2),* then $(a_1 + b_1,..,a_n + b_n)$ is again a solution of *(1.2.2).* These two together can be described as the set of solutions of a homogeneous system of linear equations closed under addition and scalar multiplication.

However, the set of solutions of a non-homogeneous system of linear equations need not be closed under addition and scalar multiplication. For example, consider the equation *4x - 3y = 1,* ***a*** = *(1, 1)* is a solution but $\alpha(1,1) = (\alpha,\alpha)$ is not a solution if $\alpha \neq 1$. Also, $b = \left(\frac{1}{4},0\right)$ is another solution but $a + b = \left(\frac{5}{4},1\right)$ is not a solution.

The homogeneous system given by *(1.2.2)* is called the associated homogeneous system of *(1.2.1).*

Let S be the set of solutions of the non-homogenous system and S_h be the set of solution of the associated homogenous system of equations. Assume $S \neq \varnothing$. S_h is always non-empty, as the privial solution $(0,...,0) \in S_h$. Let $x \in S$ and $y \in S_n$. We will show that for any $\alpha \in \mathbb{R}$, $x + \alpha y \in S$.

Since $x \in S$ we have $\sum_{j=1}^{n} a_{ij}x_j = b_i$ similarly, $\sum_{j=1}^{n} a_{ij}y_j = 0$ for $1 \leq i \leq m$. For $\alpha \in \mathbb{R}$ and $1 \leq i \leq m$, we have

$$\sum_{j=1}^{n} a_{ij}\left(x_j + \alpha y_j\right) = \sum_{j=1}^{n} a_{ij}x_j + \alpha\sum_{j=1}^{n} a_{ij}y_j$$

$$= \sum_{j=1}^{n} a_{ij}x_j$$

$$= b_i \; for 1 \leq i \leq m$$

So $x + \alpha y$ is also a solution of (1.2.1).

Now if $\boldsymbol{z} = (z_1,...,z_n) \in^s$ and $= (x_1,...,x_n) \in^s$ thus $\sum_{j=1}^{n} a_{ij}z_i = b_i$ and $\sum_{j=1}^{n} a_{ij}x_i = b_i$ .Therefore

$$\sum_{j=1}^{n} a_{ij}\left(x_j - z_j\right) = \sum_{j=1}^{n} a_{ij}x_j - \sum_{j=1}^{n} a_{ij}z_j \; = b_i - b_i = 0 \; .$$

That is, if **x** and **z** are any two solutions of the non-homogeneous system then **x - z** is a solution of the homogeneous system. That is, $x - z \in S_h$. So by the above two observations we can conclude a single fact by following way.

Let us fix $x \in S$. Then if we define $x + S_h = \{x + y | y \in S_h\}$. The first observation that $x + \alpha y$ is also a solution if (1.2.1) implies $x + S_n \mathbb{C} S$. Also for all $z \in S, z = x + (z - x) \in x + S$. This implies $S \subset x + S$. So $S = x + S_h$. This x is called *a* particular solution of (1.2.1). So we have the fact :

To find all the solution of (1.2.1) it is enough to find all the solutions of the associated homogeneous system and any particular solution (1.2.1).

These are mainly for the purpose of reviewing the so-called Gaussian elimination method of solving linear equations. Here we eliminate one variable and reduce the system to another set of linear equations with fewero number of variables. We repeat the above process with the system so obtained by deleting again one equation till we are finally left with a single equation. In this last equation, except for the first xi terms, the rest of the variables are treated as "free" and assigned arbitrary peal numbers. Let us clarify this by the following examples.

Example 1.2.1 :

$$E_1 : x + y + z = 1$$
$$E_2 = 2x - y + z = 2$$

To eliminate y we do $E_1 + E_2$ and get the equation $3x + 2z = 3$. We treat z as the free variable and assign the value t to z. i.e. z = t. so $x = 1 - \frac{2}{3}t$. Substituting y and z in E_1 we get $x = -t/3$. Thus the solution set S in given by

$$S = \left\{\left(1 - \frac{2}{3}t, -\frac{1}{3}t, t\right) / t \in R\right\}$$
$$= (1,0,0) + \left\{t\left(-\frac{2}{3}, -\frac{1}{3}, 1\right) / t \in R\right\}.$$

So (1,0,0) is a particular solution of the given system which satisfies both E_1 and E_2, hence lies on both the planes defined by the equations E_1 and E_2. And $\left(-\frac{2}{3},-\frac{1}{3},1\right)$ is a point of $\mathbb{R}^3$ which lies on the plane through the origin corresponding to the associated homogeneous system. Hence all the point on the line joining $\left(-\frac{2}{3},-\frac{1}{3},1\right)$ and the origin also lie on the plane through the origin.

Example 1.2.2 :

Consider the system -

$$E_1 := x_1 + x_2 + x_3 + x_4 = 1$$
$$E_2 := x_1 + x_2 + x_3 - x_4 = -1$$
$$E_3 := x_1 + x_2 + x_3 + 5x_4 = 5$$

Here E_1 - E_2 gives us $2x_4 = 2$. So $x_4=1$ substituting this value in above equation we get $x_1 + x_2 + x_3 = 0$. This is a linear equation in three variables and we can think x_2 and x_3 as free variables. So we let $x_2=s$ and $x_3 = t$ so that $x_1 = - s - t$. Hence the solution set is

$$\begin{aligned} S : &= \{(-s-t,s,t,1) | s,t \in R\} \\ &= \{s(-1,1,0,0) + t(-1,0,1,0) + (0,0,0,1) | s,t \in R\} \\ &= (0,0,0,1) + \{s(-1,1,0,0) + t(-1,0,1,0) | s,t \in R\} \end{aligned}$$

Check your Progress :

Solve the following systems :

1) $3x + 4y + z = 0$
$x + y + z = 0$ $\quad [Ans. t(-3,2,1) | t \in R]$

2) $x - y + 4z = 4$
$2x + 6z = -2$ $\quad [Ans. s(-1,-5,0) + t(-3,1,1) | s\ t \in R]$

3) $3x + 4y = 0$
$x + y = 0$ $\quad [Ans(0,0)]$

Observation :

By the above discussion we see that a homogeneous, system need not always have non-trivial solutions. We also observe that if the number of unknowns is more than the number of equations then the system always has a non-trivial solution.

This can be geometrically interpreted as follows :

Let $ax + b = 0$, this single equation has two variables and its solutions are all points lying on a line given by the equation.

Again

$$a_1 x + b_1 y + c_1 z = 0$$

$$a_2 x + b_2 y + c_2 z = 0$$

always has non-trivial solutions which lie on the line of intersection of the above two planes.

Theorem 1.2.1 :

The system $\sum_{j=1}^{n} a_{ij} x_j = 0$ for $1 \le i \le m$ always has non-trivial solution if $m < n$.

Proof : Let $m = 1$ *and* $n > 1$.

$\therefore a_{11} x_1 + ... + a_{1n} x_n = 0$.

If each $a_i = 0$ then any value of the variables will be a solution and a non-trivial solution ceptainly exists. Suppose some co-efficient, say $a_{ij} \ne 0$. Then we write

$x_j = -a_{1j}^{-1} \left(a_{11} x_1 + ..1 + a_{1j-1} x_{j-1} + a_{ij+1} x_{j+1} + ... + a_{1n} x_n \right)$.

Hence if we choose $\alpha_i \in \mathbb{R}$ arbitrarily for all $i \ne j$ and take $\alpha_j = -a_{1j}^{-1} \left(a_{11} \alpha_1 + .. + a_{1j-1} \alpha_{j-1} + a_{1j+1} \alpha_{j+1} + .. + a_{1n} \alpha_n \right)$. then $(\alpha_1, .., \alpha_n)$ is a solution of $\sum_{j=1}^{n} a_{ij}^{xj} = 0$. Thus for m = 1 and n > 1 we get a non trivial solution.

We prove the result by induction on m. As induction hypothesis, let the system (m-1) equation in k variables where (m-1)<k, has a non-trivial solution. We prove it for m and n with m < n.

Let $\sum_{j=1}^{n} a_{ij}^{xj} = 0$ for $1 \le i \le m$ be a system of m equations in n unknowns with m < n. If each $a_{ij} = 0$, only n tuple $(x_1, \ldots, x_n)$ is a solution. Hence non-trivial solutions exist. If not let there exist (i, j) such that $a_{ij} \ne 0$. Let

$$E_1 := a_{11}x_1 + ... + a_{1n}x_n = 0$$
$$E_2 := a_{21}x_1 + ... + a_{2n}x_n = 0$$
$$\vdots$$
$$E_i := a_{i1}x_1 + ... + a_{ih}x_n = 0$$
$$\vdots$$
$$E_m := a_{m1}x_1 + ... + a_{mn}x_n = 0$$

Since $a_{ij} \neq 0$, from E_i we have $x_j = -a_{jj}^{-1}\left(a_{ij}x_1 + .. + a_{i,j-1}x_{j-1} + a_{i,j+1}x_{j+1} + .. + a_{in}x_n\right)$.

If we substitute this value of x_j in other equations we will get a new system of *(m-1)* equations in *(n-1)* variables $x_1, ..x_{j-1}, x_{j+1}, ..x_n$ as follows :

For $1 \le k \le m, k \neq i$ $E_k := \sum_{r \neq j}\left[a_{kr} + a_{kj}(-a_{ij}^{-1})\ a_{ir}\right]x_r = 0$

because by E_1 we get

$a_{11}x_1 + ..ta_{1,j-1}x_{j-1} + a_{1j}\left[-a_{ij}^{-1}\left(a_{i1}x_1 + .. + a_{i,j-1}x_{j-1}ta_{i,j+1}x_{j+1} + .. + a_{in}x_n\right)\right] + a_{1,j+1}x_{j+1} + .. + a_{1n}x_n = 0$ which implies.

$$\left[a_{11} + a_{1j}\left(-a_{ij}^{-1}\right)a_{i1}\right]x_1 + ----- +$$
$$\left[a_{1j-1} + a_{1j}\left(-a_{ij}\right)^{-1} a_{i,j-1}\right]x_j + ----- +$$
$$\left[a_{1n} + a_{1j}\left(-a_{ij}^{-1}\right)a_{in}\right]x_n = 0$$

i.e. $E_i := \sum_{\substack{r=1 \\ r \neq j}}^{n}\left[a_{1r} + a_{ij}\left(-a_{ij}^{-1}\right)a_{ir}\right]x_r = 0$ so by induction hypothesis *(m-1)* equation E_k has a non-trivial solution.

$x_1, ..., x_{j-1}, x_{j+1}, .., x_n$ as $m-1 < n-1$. In particular, $x_k \neq 0$ for some $k \neq j$. We take $x_j = \alpha_j$ so, $\alpha_j = -a_{ij}^{-1}\sum_{r \neq j} a_{ir}\alpha_p$. We claim $\alpha_1, .., \alpha_{j-1}, \alpha_j, .., \alpha_n$ is a non-trivial solution.

For $1 \le k \le m$,

$$E_k = \sum_{r \neq j} a_{kr}x_r + a_{kj}\alpha_j$$
$$= \sum_{r \neq j} a_{kr}\alpha_r + a_{kj}\left(-a_{ij}^{-1}\right)\sum_{r \neq j} a_{is}\alpha_s$$
$$= \sum_{r \neq j}\left[a_{kr} + a_{kj}\left(-a_{ij}^{-1}\right)a_{ir}\right]\alpha_p$$
$$= E_k = \text{O for } k \neq i$$

As $(\alpha_1,...\alpha_n)$ is a solution of $E_i, \sum_r a_{ir}\alpha_p = \sum a_{ir}\alpha_r + (a_{ij})$

$$\left\|\left(-a_{ij}^{-1}\sum_{r\neq j} a_{ir}\alpha_r\right) = \sum_{r\neq j}(a_{ir} - \alpha_{ir})\alpha_r = 0\right.$$

$(\alpha_1, -, \alpha_n)$ is non-trivial since $\alpha_k \neq 0$ for some $k \neq j$ by the induction hypothesis.

Thus $\alpha_1,....,\alpha_n$ is a non-trivial solution of the original system.

Exercise 1.1 :

1) Find one non-trivial solution for each one of the following systems of equations.

a) x + 2y + z = 0

b) 3x + y + z = 0
x + y + z = 0

c) 2x - 3y + 4z = 0
3x + y + z = 0

d) 2x + y + 4z + 19 = 0
-3x + 2y - 3z + 19 = 0
x + y + z = 0

2) Show that the only solution of the following systems of equations are trivial.

a) 2x + 3y = 0
x - y = 0

b) 4x + 5y = 0
-6x + 7y = 0

c) 3x + 4y - 2z = 0
x + y + z = 0
-x - 3y + 5z = 0

d) 4x - 7y + 3z = 0
x + y = 0
y - 6z = 0

1.3 SOLUTION OF THE SYSTEM OF LINEAR EQUATIONS BY GAUSS ELIMINATION METHOD

Let the system of m linear equations in n unknown be $\sum_{j=1}^{n} a_{ij}x_j = b_i, 1 \le i \le m$.

Then the co-efficient matrix is -

$$\begin{pmatrix} a_{11} & a_{12} & .... & a_{1n} \\ \vdots & \vdots & .... & \vdots \\ a_{m1} & a_{m2} & .... & a_{mn} \end{pmatrix}$$

Also we define the augmented matrix b_y

$$\begin{pmatrix} a_{11} & a_{12} & .... & a_{1n} & b_1 \\ \vdots & \vdots & .... & \vdots & \vdots \\ a_{m1} & a_{m2} & .... & a_{mn} & b_m \end{pmatrix}$$

We will perform the following operations on the system of linear equations, called elementary row operations :

Multiply one equation by a non-row number.

Add are equation to another.

Interchange two equations.

These operations are reflected in operations on the augmented matrix, which are also called elementary row operations.

Suppose that a system of linear equations is changed by an elementary row operation. Then the solutions of new system are exactly the same as the solutions of the old systems. By making row operations, we will try to simplify the shape of the system so that it is earier to find the solutions.

Let us define two matrices to be row equivalent if one can be obtained from the other by a succession of elementary row operations. If A is the matrix of co-efficients of a system of linear equations, and B the column vector as above, so that (A, B) in the augmented matrix and if (A^1,B^1) in row-equivalent of the system. AX = B are the same as the solutions of the system A^1 X = B^1.

Example 1.3.1 :

Consider the system of linear equations

$$3x - 2y + z + 2\upsilon = 1$$
$$x + y - z - \upsilon = -2$$
$$2x - y + 3z = 4$$

The augmented matrix is :

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

Subtract 3 times second row from first row :

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

Subtract 2 times second row from third row :

$$\begin{pmatrix} 0 & -5 & 4 & 5 & 7 \\ 1 & 1 & -1 & -1 & -2 \\ 0 & -3 & 5 & 2 & 8 \end{pmatrix}$$

Interchange first and second row.

$$\begin{pmatrix} 1 & 1 & -1 & -1 & 2 \\ 0 & -5 & 4 & 5 & 7 \\ 0 & -3 & 5 & 2 & 8 \end{pmatrix}$$

Multiply second row by -1

$$\begin{pmatrix} 1 & 1 & -1 & -1 & -2 \\ 0 & 5 & -4 & -5 & -7 \\ 0 & -3 & 5 & 2 & 8 \end{pmatrix}$$

Multiply second row by 3 and third row by 5

$$\begin{pmatrix} 1 & 1 & -1 & -1 & -2 \\ 0 & 15 & 12 & 15 & -21 \\ 0 & -15 & 25 & 10 & 40 \end{pmatrix}$$

Add second row to third row.

$$\begin{pmatrix} 1 & 1 & -1 & -1 & -2 \\ 0 & 15 & -12 & -15 & -21 \\ 0 & 0 & 13 & -5 & 17 \end{pmatrix}$$

The new system whose augmented matrix is the last matrix coin be written as :

$$\begin{aligned} &x + y - z - \upsilon = -2 \\ &15y - 12z - 15\upsilon = -21 \\ &13z - 5\upsilon = 19 \end{aligned}$$

Now it we consider $\begin{aligned} &\upsilon = t, \\ &z = \frac{19+5t}{13} \end{aligned}$

$$\begin{aligned} 15y &= \frac{12}{13}(19+5t) + 15t - 21 \\ y &= \frac{255t-51}{195} \\ x &= -\frac{255t-51}{195} + \frac{19+5t}{13} + t - 2 \\ &= \frac{15t-54}{194} \end{aligned}$$

This method is known as Gauss elimination method.

Example 1.3.2 :

Consider the system of linear equations.

$$\begin{aligned} x_1 + x_2 + x_3 + x_4 + x_5 &= 1 \\ -x_1 - x_2 \qquad\qquad + x_5 &= -1 \\ -2x_1 - 2x_2 \qquad\qquad + x_5 &= 1 \\ x_3 + x_4 + x_5 &= -1 \\ x_1 + x_2 + 2x_3 + 2x_4 + 2x_5 &= 1 \end{aligned}$$

The augmented matrix is :

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ -1 & -1 & 0 & 0 & 1 & -1 \\ -2 & -2 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & -1 \\ 1 & 1 & 2 & 2 & 2 & 1 \end{pmatrix}$$

Adding 2nd row to 1st row, two times 3rd row to 1st row and subtracting last row from 1st row we get.

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

Subtracting twice the 2nd row from 3rd row, 4th row from 2nd row and 5th row from 2nd row we get

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

The equations represented by the last two rows are :

$$ox_1 + ox_2 + ox_3 + ox_4 + ox_5 = -4$$
$$ox_1 + ox_2 + ox_3 + ox_4 + ox_5 = -3$$

Which implies that the system is inconsistent.

Exercise 1.2 :

For each of the following system of equations, use Gaussian elimination to solve them.

i) $$3x_1 + 2x_2 - x_3 = 4$$
$$x_1 - 2x_2 + 2x_3 = 1$$
$$11x_1 + 2x_2 + x_3 = 14$$

ii) $$x_1 - x_2 + 2x_3 = 4$$
$$2x_1 + 3x_2 - x_3 = 1$$
$$7x_1 + 3x_2 + 4x_3 = 7$$

iii) $$x_1 + x_2 + x_3 + x_4 = 0$$
$$2x_1 + 3x_2 - x_3 - x_4 = 2$$
$$x_1 - x_2 + 2x_3 + 2x_4 = 3$$
$$2x_1 + 5x_2 - 2x_3 - 2x_4 = 4$$

iv) $$x_1 + 3x_2 + x_3 + x_4 = 3$$
$$2x_1 - 2x_2 + x_3 + 2x_4 = 8$$
$$3x_1 + x_2 + 2x_3 - x_4 = -1$$

Answer

Exercise 1.2 :

i) $\left\{\left(\frac{5-x}{4}, \frac{1+7x}{8}, x\right) : x \; rual\right\}$

ii) inconsistent

iii) inconsistent

iv) $\left\{\left(\frac{15}{4}, -\frac{5x}{8} - \beta, -\frac{1}{4} - \frac{1}{8}x, x, \beta\right)\right\}$

❖❖❖❖

2

VECTOR SPACE

Unit Structure :

2.0 INTRODUCTION

The concept of a vector is basic for the study of functions of several variables. It provides geometric motivation for everything that follows. We know that a number can be used to represent a point on a line, once a unit length is selected. A pair of numbers *(x,y)* can be used to represent a point in the plane where as a triple of numbers (x,y,z) can be used to represent a point in 3 dimensional space denoted by $\mathbb{R}^3$. The line can be called 1- dimensional space denoted by $\mathbb{R}$ or plane. 2-) dimensional space denoted by $\mathbb{R}^2$ Continuing this way we can define a point in n-space as (x_1, x_2, -z, x_n). Here $\mathbb{R}$ is a set of real numbers and x is an element in $\mathbb{R}$ which we write as $x \in \mathbb{R}$. $\mathbb{R}^2$ is a set of ordered pair and (x,y)$\in \mathbb{R}^2$. Thus X is an element of $\mathbb{R}^n$ or $X \in \mathbb{R}^n$ means X = (x_1, x_2, …, x_n). These elements as a special case are called vectors from respective spaces. The vectors from a same set or space can be added and multiplied by a number. It is now convenient to define in general a notion which includes these as a special case.

2.1 OBJECTIVES

After going through this chapter you will be able to :

- Verify that a given set is a vector space or not over a field.
- Get concept of vector subspace.
- Get concept of basis and dimension of vector space.

2.2 DEFINITION AND EXAMPLES

We define a vector space to be a set on which "addition" and "readar multiplication" are defined. More precisely, we can tell.

2.2.1 Definition : Vector Space

Let (F, + , .) be a field. The elements of F will be called scalars. Let V be a non-empty set whose elements will be called vectors. Then V is a vector space over the field F, if

1. There is defined an internal composition in V called addition of vectors and denoted by '+' in such a way that :

i) $\alpha+\beta \in V$ for all $\alpha, \beta \in V$ (closer property)
ii) $\alpha+\beta=\beta+\alpha$ for all $\alpha, \beta \in V$ (commutative property)
iii) $\alpha+(\beta+\gamma)=(\alpha+\beta)+\gamma$ for all $\alpha, \beta, \gamma \in V$ (Associate property)
iv) $\exists$ an element $O \in V$ such that $\alpha+O=\alpha$ for all $\alpha \in V$ (Existence of Identity)
v) To every vector $\alpha \in V \exists$ a vector $-\alpha \in V$ such that $\alpha+(-\alpha)=\varnothing$ (Existence of inverse)

2. There is an external composition in V over F called scalar multiplication and denoted multiplicatively in such a way that :

i) $a\alpha \in V$ for all $a \in F$ and $\alpha \in V$ (Closer property)
ii) $a(\alpha+\beta)=a\alpha+a\beta$ for all $a \in F$ and $\alpha, \beta \in V$ (Distributive property)
iii) $(a+b)\alpha=a\alpha+b\alpha$ for all $a, b \in F$, $\alpha \in V$ (distributive property)
iv) $(ab)\alpha=a(b\alpha) \forall a, b \in F$ and $\alpha \in V$.
v) $1\alpha=\alpha$ for all $\alpha \in V$ and 1 in the unity element of F.

When V is a vector space over the field F, we shall say that V(F) is a vector space or sometimes simply V is a vector space. If F is the field $\mathbb{R}$ or real numbers, V is called a real vector space; similarly if F is Q or C, we call V as a rational vector space or complex vector space.

Note 1 : There should not be any confusion about the use of the word vector. Here by vector we do not mean the vector quantity which we have defined in vector algebra as a directed line segment. Here we shall call the elements of the set V as vectors.

Note 2 : The symbol '+' is used for addition of vectors which is also used to denote the addition of two scalars in F. There should be no confusion about the two compositions. Similarly for scalar multiplication, we mean multiplication of an element of V by an element of F.

Note 3 : In a vector space we shall be dealing with two types of zero elements. One is the zero elemen of F which is owl well known 0. Another is the zero vector in V i.e. if $V=\mathbb{R}^3, O=(0,0,0)$.

Example 1 : The n-triple space, F^n.

Let F be any field and let V be the set of all n-tuples $\alpha=(x_1, x_2, ..., x_n)$ of scalars x_i in F. If $\beta=(y_1, y_2, ..., y_n)$ with y_i in F, the sum of α and β is defined by $\alpha+\beta=(x_1+y_1, ...x_n+y_n)$. The product of scalar c and vector α is defined by $c\alpha=(cx_1, cx_2, ..., cx_n)$. This vector addition and scalar multiplication satisfy all conditions of vector space. (Verfication is left for the students).

For n = 1, 2 or 3, F = $\mathbb{R}$, $\mathbb{R}^2$ or $\mathbb{R}^3$ are basic examples of vector space.

Example 2 : The space of mxn matrices, M_{mxn}(F).

Let F be any field and let m and n be positive integers. Let M_{mxn}(F) be the set of all mxn matrices over the field F. The sum of two vectors A and B in M_{mxn}(F) is defined by $(A+B)_{ij}=A_{ij}+B_{ij}$. The product of a scalar C and the matrix A is defined by $(CA)_{ij}=CA_{ij}$.

Example 3 : The space of functions from a set to a field.

Let F be any field and let S be any non-empty set. Let V be the set of all function from the set S into F. The sum of two vectors f and g in V is the vector f + g i.e. the function from S into F, defined by $(f+g)(f_n)=f(f_n)+g(f_n)$.

The product of the scalar c and the function f is the function e f defined by $(cf)(f_n)=cf(f_n)$.

For this example we shall indicate how one verifies that V is a vector space over F. Here $V=\{f : f : S \to F\}$. We have,

$(f+g)(s)=f(s)+g(s)\forall s\in S$. Since f (s) and g(s) are in F and F is a field, therefore f (x) +g(x) in also in F. Thus f+g is also a function from S to F. Therefore $f+g\in V\forall f,g\in V$. Therefore V is closed under addition.

Associatvity of addition :

We have $[(f+g)+h](x)=(f+g)(x)+h(x)$ (by def.)

$=[f(x)+g(x)]+h(x)$ (by def)

$=f(x)+[g(x)+h(x)]$

[$\because f(x),g(x),h(x)$ are elements of F and addition in F is associative]

$=f(x)+(g+h)(x)$

$=[f+(g+h)](x)$

$\therefore (f+g)+h=f+(g+h)$

Commutativity of addition :

We have $(f+g)(x)=f(x)+g(x)$

$=g(x)+f(x)$ [$\because$ addition is commutative in F]

$=(g+f)(x)$

$\therefore f+g=g+f$

Existence of Additive identity :

Let us define a function $\hat{O}:S\to F$ such that $\hat{O}(x)=O\forall x\in S$. Then $\hat{O}\in V$ and it is called zero function.

We have $(f+\hat{O})(x)=f(x)+\hat{O}(x)=f(x)+O=f(x)$

$\therefore f+\hat{O}=f$

$\therefore$ The function $\hat{O}$ is the additive identity.

Existence of additive inverse :

Let $f\in V$. Let us define a function $-f:S\to F$ by

$-f(x)=-[f(x)]\forall x\in S$. Then $-f\in V$ and we have

$[f+(-f)](x)=f(x)+[(-f)(x)]$

$$= f(x) + [-f(x)]$$
$$= f(x) - f(x)$$
$$= O = \widehat{O}(x)$$

$\therefore f + (-f) = \widehat{O}$

$\therefore$ the function **-** f in the additive inverse of f.

Now for scalar multiplication $if \in F$ and $f \in V$, then $\forall x \in S$,

$(cf)(x) = cf(x)$.

Now $f(x) \in F$ and $c \in F$. Therefore c f (x) is in F. Thus V is closed with respect to scalar multiplication.

Next we observe that

i) If $c \in F$ and $f, g \in V$ then

$$[c(f+g)](x) = c[(f+g)(x)] = c[f(x) + g(x)]$$
$$= cF(x) + cg(x)$$
$$= (cf)(x) + (cg)(x)$$

$\therefore c(f+g) = cf + cg$

ii) If c_1, c_2 $\in$F and $f \in V$, then

$$[(c_{1=} + c_2) f](x) = (c_1 + c_2) f(x)$$
$$= c_1 f(x) + c_2 f(x)$$
$$= (c_1 f)(x) + (c_2 f)(x)$$

$\therefore (c_1 + c_2) f = c_1 f + c_2 f$

iii) If $c_1, c_2 \in F$ and $f \in V$ then

$$[(c_1 c_2) f](x) = (c_1 c_2) f(x) = c_1 [c_2 + (x)]$$
$$= c_1 [(c_2 f)(x)]$$
$$= [c_1 (c_2 f)](x)$$

$\therefore (c_1 c_2) f = c_1 (c_2 f)$

iv) If 1 is the unity element of F and $f \in V$, then

$(1f)(x) = 1f(x) = f(x)$

$\therefore 1f = f$

Hence V is a vector space over F.

Example 4 : The set of all convergent sequences over the field of real numbers.

Let V denote the set of all convergent sequences over the field of real numbers.

Let $\alpha=\{\alpha_1,\alpha_2,...,\alpha_{n,}.....\}=\{\alpha_n\}$, $\beta=\{\beta_1,\beta_2,...,\beta_{n,}.....\}=\{\beta_n\}$ and $\gamma=\{\gamma_1,\gamma_2,...,\gamma_{n,}.....\}=\{\gamma_n\}$ be any three convergent sequence.

1. Properties of vector addition.

i) We have $\alpha+\beta=\{\alpha_n\}+\{\beta_n\}=\{\alpha_n+\beta_n\}$ which is also a convergent sequence. Therefore V is closed for addition of sequences.

ii) Commutativity of addition : We have $\alpha+\beta=\{\alpha_n\}+\{\beta_n\}=$ $\{\alpha_n+\beta_n\}=\{\beta_n+\alpha_n\}=\{\beta_n\}+\{\alpha_n\}=\beta+\alpha$

iii) Associativity of addition : We have

$$
\begin{aligned}
\alpha+(\beta+\gamma)&=\{\alpha_n\}+[\{\beta_n\}+\{\gamma_n\}]\\
&=\{\alpha_n\}+\{\beta_n+\gamma_n\}\\
&=\{\alpha_n+(\beta_n+\gamma_n)\}\\
&=\{(\alpha_n+\beta_n)+\gamma_n\}\\
&=[\{\alpha_n\}+\{\beta_n\}]+\{\gamma_n\}\\
&=(\alpha+\beta)+\gamma
\end{aligned}
$$

iv) Existence of additive identity : The zero sequence $\{0\}=\{0,0,..,0,..\}$ is the additive identity.

v) Existence of additive inverse : for every sequence $\{\alpha_n\}\exists$ a sequence $\{-\alpha_n\}$ such that $\{\alpha_n\}+\{-\alpha_n\}=\{\alpha_n-\alpha_n\}=\{0\}$.

2) Properties of scalar multiplication :

i) Let $a\in\mathbb{R}$. Then $a\alpha=a\{\alpha_n\}=\{a\alpha_n\}$ which is also a convergent sequence because $\lim_{h\leftarrow\alpha} a\alpha_n=a\lim_{h\leftarrow\alpha}\alpha_n$.

Thus V is closed for scalar multiplication.

ii) Let $a \in \mathbb{R}$ and $\alpha, \beta \in V$, then we have

$$a(\alpha+\beta) = a[\{\alpha_n\}+\{\beta_n\}] = a\{\alpha_n+\beta_n\}$$
$$= \{a(\alpha_n+\beta_n)\}$$
$$= \{a\alpha_n + a\beta_n\} = \{a\alpha_n\}+\{a\beta_n\}$$
$$= a\{\alpha_n\}+a\{\beta_n\} = a\alpha + a\beta$$

iii) Let $a,b \in \mathbb{R}$ and $\alpha \in V$,

$$(a+b)\alpha = (a+b)\{\alpha_n\} = \{(a+b)\alpha_n\}$$
$$= \{a\alpha_n + b\alpha_n\} = \{a\alpha_n\}+\{b\alpha_n\}$$
$$= a\{\alpha_n\}+b\{\alpha_n\} = a\alpha + b\alpha$$

iv) $(ab)\alpha = (ab)\{\alpha_n\} = \{(ab)\alpha_n\} = \{a(b\alpha_n)\}$

$$= a\{b\alpha_n\} = a[b\{\alpha_n\}] = a(b\alpha)$$

v) $1\alpha = 1\{\alpha_n\} = \{1\alpha_n\} = \{\alpha_n\} = \alpha$

Thus all the postulates of a vector space are satisfied. Hence V in a vector space over the field of real numbers.

Check your Progress :

1. Show that the following are vector spaces over the field $\mathbb{R}$.
   i) The set of all real valued functions defined in some interval [0,1].
   ii) The set of all polynominals of degree at most n.

Example 5 :

Let V be the set of all pairs (*x*, *y*) of real numbers and let F be the field of real numbers. Let us define $(x,y)+(x_1,y_1)=(x+x_1,0)$ $c=(x,y) \; =(cx,0)$.

Is V with these operations, a vector space over the field $\mathbb{R}$?

Solution :

If any of the postulates of the vector space in not satisfied, then V will not be a vector space. We shall show that for the operation of addition of vectors as defined in this problem the identity element does not exist. Suppose (x_1, y_1) is additive identity element.

Then we must have $(x,y)+(x_1,y_1)=(x,y)\forall x,y\in\mathbb{R}\Rightarrow(x+x_1,0)=(x,y)$ which is not possible unless y = 0. Thus $\exists$no element $(x_1,y_1)$of V s.t. $(x,y)+(x_1,y_1)=(x,y)\forall(x,y)\in V$.

As the additive identity element does not exist in V, it is not a vector space.

Exercise : 2.1

1) What is the zero vector in the vector space $\mathbb{R}^4$?
2) Is the set of all polynomials in x of degree ≤ 2 a vector space?.
3) Show that the complex field ₵ is a vector space over the real field $\mathbb{R}$.
4) Prove that the set $V=\{(a,b):a,b\in\mathbb{R}\}$ is a vector space over the field $\mathbb{R}$ for the compositions of addition and scalar multiplication defined as $(a,b)+(c,d)=(a+c,b+d)k(a,b)=(ka,kb)$.
5) Let V be the set of all pairs *(x, y)* of real numbers and let F be the field of real numbers. Define $(x,y)+(x_1,y_1)=(x_1+x_1,y+y_1)c(x,y)=\ (ex,y)$. Show that with these operations V is not a vector space over $\mathbb{R}$.

2.2.2 Definition : Vector subspace

Let V be a vector space over the field F and let $W\underline{C}V$. Then W is called a subspace of V if W itself is a vector space over F with respect to the operations of vector addition and scalar multiplication in V.

Theorem 1 :

The necessary and sufficient condition for a non-empty subset W of a vector space V(F) to be a subspace of V is $a,b\in F$ and $\alpha,\beta\in W\Rightarrow a\alpha+b\beta\in W$

Proof :

The condition necessary :

If W is a subspace of V, by the definition it is also a vector space and hence it must be closed under scalar multiplication and vector addition. Therefore $a\in F,\alpha\in W\Rightarrow a\alpha\in W$ and $b\in F,\beta\in W\Rightarrow b,\beta\in W$ and $a\alpha\in W,b\beta\in W\Rightarrow a\alpha+b\beta\in W$. Hence the condition is necessary.

The condition sufficient :

Now W is a non-empty subset of V satisfying the given condition i.e. $a, b \in F$ and $\alpha, \beta \in W \Rightarrow a\alpha + b\beta \in W$. Taking $a = 1, b = 1$ we have $\alpha + \beta \in W \forall \alpha, \beta \in W$. Thus W is closed wider addition taking $a = -1, b = 0$ we have $-\alpha \in W \forall \alpha \in W$. Thus additive inverse of each element of W is also in W.

Taking $a = \mathbf{0}, b = \mathbf{0}$, we have that if $\alpha \in W$, $\mathbf{0}\alpha + \mathbf{0}\alpha \in W$ $\Rightarrow \mathbf{0} + \mathbf{0} \in W$ $\Rightarrow \mathbf{0} \in W$.

Thus the zero vector of V belongs to W which is also the zero vector in W.

Since the elements of W are also the elements of V, therefore vector addition will be associatine as well as commutative in V.

Now taking $\beta = \mathbf{0}$, we see that if $a, b \in F$ and $\alpha \in W$, then $a\alpha + l\mathbf{0} \in W$ i.e. $a\alpha + \mathbf{0} \in W$ i.e. $\mathbf{0}\alpha \in W$ So W is closed under scalar multiplication.

The remaining postulates of a vector space will hold in W since they hold in V of which W is a subset. Thus W(F) is a vector space. Hence W(F) is a subspace of V(F).

Example 5 :

a) If V is any vector space V is a subspace of V. The subset consisting of the zero vector alone is a subspace of V, called the zero subspace of V.

b) The space of polynomial functions over the field F is a subspace of the space of all functions from F into F.

c) The symmetric matrices form a subspace of the space of all *nxn* matrices over F.

d) An nxn matrix A over the field ₵ of complex numbers is Hermitian if $A_{jk} = \overline{A}_{kj}$ for each j, k, the bar denoting complex conujugation. A *zxz* matrix is Hermitian if and only if it has the form $\begin{bmatrix} z & x+iy \\ x-iy & \omega \end{bmatrix}$ where *x, y, z* and *w* are real numbers.

The set of all Hermitian matrices is not a subspace of the space of all *nxn* matrices over ₵. For if A is Hermitian, its diagonal entries A_{11}, A_{22} are all real number but the diagonal entries of iA are in general not real. On the other hand, it is easily verified that the set of *nxn* complex Hermitian matrices is a vector space over the field $\mathbb{R}$ with the usual operations.

Theorem 2 :

Let V be a vector space over the field F. The intersection of any collection of subspaces of V is a subspace of V.

Proof :

Let $\{Wa\}$ be a collection of subspaces of V and let $W = \bigcap Wa$ be their intersection. By definition of W, it is the set of all elements belonging to every W*a*. Since each W*a* is a subspace, each contains the zero vector. Thus the zero vector is in the intersection W and so W is non-empty. Let $b\beta \in W$ and $a, b \in F$. So, both α, β belong to each W*a*. But W*a* is a subspace of V and hence $a\alpha + b\beta \in Wa$. Thus $a\alpha + b\beta \in W$. So W is *a* subspace of V.

The above theorem follows that if S is any collection of vectors in V, then there is a smallest subspace of V which contains S, that is, a subspace which contains S and which is continued in every other subspace containing S.

Definition : Let S be a set of vectors in a vector space V. The **subspace spanned** by S is defined to be the intersection W of all subspace of V which contain S when S is a finite set of vectors, $S = \{\alpha_1, \alpha_2, ...\alpha_n\}$, we shall simply call W the subspace spanned by the vectors $\alpha_1, \alpha_2, ..., \alpha_n$.

Definition : Linear Combination :

Let V(F) be a vector space. If $\alpha_1, \alpha_2, ..., \alpha_n \in V$, then any vector $\alpha = a_1\alpha_1 + ... + a_n\alpha_n$ when $a_1, a_2, ...a_n \in F$ is called a linear combination of the vectors $\alpha_1, \alpha_2, ...\alpha_n$

Definition : Linear Span :

Let V(F) be a vector space and S be any non-empty subset of V. Then the linear span of S is the set of all linear combinations of finite sets of elements of S and is denoted by L(S). Thus we have $L(S) = \{a_1\alpha_1 + a_2\alpha_2 + ... + a_n\alpha_n : \alpha_1, \alpha_2, .., \alpha_n \in S \text{ and } a_1, a_2, ...a_n \in F\}$.

Theorem 3 :

The linear span L(S) of any subset S of a vector space V(F) is a subspace of V generated by S i.e. L(S) = {S}.

Proof :

Let α, β be any two elements of L(S). Then $\alpha = a_1\alpha_1 + ... + a_n\alpha_n$ and $\beta = b_1\beta_1 + ... + b_n\beta_n$ where $a_i, b_i \in F, \alpha_i, \beta_i \in S, i = 1,...m, j = 1,....n$.

If a, l be any two elements of F, then $a\alpha + b\beta = a(a_1\alpha_1 + .. + a_m\alpha_m) + b(b_1\beta_1 + ... + b_n\beta_n) = a(a_1\alpha_1) + .. + a(a_m\alpha_m) + b(b_1\beta_1) + .. + b(b_n\beta_n) = (aa_1)\alpha_1 + .. + (aa_m)\alpha m + (bb_1)\beta_1 + ... + (bb_n)\beta_n$.

Thus $a\alpha + b\beta$ has been expressed as a linear combination of a finite set $\alpha_1, ..., \alpha_m, \beta_1, ..., \beta_n$ of the element of S. Consequently $a\alpha + b\beta \in L(S)$. Thus $a, b \in F$ and $b, \beta L(S) \Rightarrow a\alpha + b\beta \in L(S)$.

Hence L(S) is a subspace of V(F). Also each element of S belongs to L(S) as if $\alpha_r \in S$, then $\alpha_r = 1\alpha r$ and this implies that $\alpha_r \in L(S)$. Thus L(S) in a subspace ov V and S in contained in L(S).

Now if W is any subspace of V containing S, then each element of L(S) must be in W because W is to be closed under vector addition and scalar multiplication. Therefore L(S) will be contained in W. Hence L(S) = {S} i.e. L(S) in the smallest subspace of V containing S.

Check your progress :

1) Let $W = \{(a, a_2, 0) : a_1, a_2 \in \mathbb{R}\}$. Show that W is a subspaces of $\mathbb{R}^3$.
2) Show that the set W of the elements of the vector space $\mathbb{R}^3$ of the form (x + 2y, y, -x + 3y) where $x, y \in \mathbb{R}$ is a subspaces of $\mathbb{R}^3$.
3) Which of the following are subspaces of $\mathbb{R}^3$.
   i) $\{(x, 2y, 3z) : x, y, z \in \mathbb{R}\}$
   ii) $\{(x, x, x) : x \in \mathbb{R}\}$
   iii) $\{(x, y, z) : x, y, z \text{ are rational numbers}\}$

Definition : Linear Dependence :

Let V(F) be a vector space, A finite set $\{\alpha_1,...,\alpha_n\}$ of vectors of V is said to be linearly dependent if there exist scalars $a_1,...,a_n \in$ F not all of them O such that $a_1\alpha_1 + .. + a_n\alpha_n = 0$.

Definition : Linear independence :

Let V(F) be a vector space. A finite set $\{\alpha_1,...\alpha_n\}$ of vectors of V is said to be linearly independent if every relation of the form $a_1\alpha_1 + .. + a_n\alpha_n = 0, a_i \in F, 1 \le i \le n \Rightarrow a_i = 0$ for each $1 \le i \le n$.

Any infinite set of vectors of V is said to be linearly independent if its every finite subset is linearly independent, otherwise it is linearly dependent.

Exercises : 2.2

1) Which of the following sets of vectors $\alpha = (a_1,..a_n)$ in $\mathbb{R}^n$ are subspaces of $\mathbb{R}^n$? $(n \ge 3)$.
   i) all α such that $a_1 \ge 0$.
   ii) all α such that $a_1 + 3a_2 = a_3$.
   iii) all α such that $a_2 = a_1^2$.
   iv) all α such that $a_1 a_2 = 0$.
   v) all α such that a_2 is national..

2) State whether the following statements are true or false.
   i) A subspace of $\mathbb{R}^3$ must always contain the origin.
   ii) The set of vectors $\alpha = (x, y) \in \mathbb{R}^2$ for which $x^2 = y^2$ is a subspace of $\mathbb{R}^2$
   iii) The set of ordered triads (x,y,z) of real numbers with *x > 0* is a subspace of $\mathbb{R}^3$.
   iv) The set of ordered trials (x, y, z) of real numbers with *x + y = 0* is a subspaces of $\mathbb{R}^3$.

3) In $\mathbb{R}^3$, examine each of the following sets of vectors for linear dependence.
   i) $\{(2,1,2),(8,4,8)\}$
   ii) $\{(1,2,0),(0,3,1),(-1,0,1)\}$
   iii) $\{(2,3,5),(4,9,25)\}$
   iv) $\{(1,2,1),(3,1,5),(3,-4,7)\}$

4) Is the vector (2,-5,3) in the subspace of $\mathbb{R}^3$ spanned by the vectors (1,-3,2), (2, -4, -1), (1, -5, 7)?

5) Show that the set {1, x, x(1-x)} is a linearly independent set of vectors in the space of all polynomials over $\mathbb{R}$.

2.2.3 Basis and dimension :

In this section we will assign a task to give dimension to certain vector spaces. We usually associate 'dimension' with something geometrical. But after developing the concept of a basis for a vector space we can give a suitable algebraic definition of the dimension of a vector space.

Definition : Basis of a vector space

A subset S of a vector space V(F) is said to be a basis of V(F), if

i) S consists of linearly independent vectors.

ii) S generates V(F) i.e. L(S) = V i.e. each vector in V is a linear combination of a finite number of elements of S.

Example 1 :

Let $V = \mathbb{R}^n$, If $x = (x_1, ..., x_n) \in \mathbb{R}^n$ we call x_j, the ith co-ordinate of *x*.. Let e_i : (0, …, 0, 1, 0, .. 0) be the vector whose it is co-ordinate in 1 and others are 0. It is easy to show that $\{e_i | 1 \leq i \leq n\}$ is a basis of V. This called the standard basis of $\mathbb{R}^n$.

Example 2 :

The infinite set $S = \{1, x, x^2, ..., x^n, ..\}$ is a basis of the vector space F[x] of polynomials over the field F.

Definition : Finite Dimensional Vectors Spaces. The vector space V(F) is said to be finite dimensional or finitely generated if there exists a finite subset S of V such that V = L(S).

The vector space which is not finitely generated may be referred to as an infinite dimensional space.

Theorem 1 :

Let V be a vector space spanned by a finite set of vectors $\beta_1, \beta_2, ... \beta_m$. Then any independent set of vectors in V is finite and contains no more than *m* elements.

Proof :

To prove the theorem it suffices to show that every subset S of V which contains more than *m* vectors is linearly dependent. Let S be such a set. In S there are distinct vectors $\alpha_1, \alpha_2, ... \alpha_n$ where n > m. Since $\beta_1, ... \beta_m$ span V, there exists scalars A_{ij} in F such that $\alpha_j = \sum_{i=1}^{m} A_{ij}\beta_i$. For any n scalars $x_1, x_2, ..., x_n$ we have

$$x_1\alpha_1 + ... + x_n\alpha_n = \sum_{j=1}^{n} x_j\alpha_j$$

$$= \sum_{j=1}^{n} x_j \sum_{j=1}^{m} A_{ij}\beta_i$$

$$= \sum_{j=1}^{n} x_j \sum_{i=1}^{m} \left(A_{ij}x_j\right)\beta_i$$

$$= \sum_{i=1}^{m} \left(\sum_{j=1}^{n} A_{ij}x_j\right)\beta_i$$

Since n > m, the homogeneous system of linear equation $\sum_{j=1}^{n} A_{ij}x_j = 0, 1 \leq i \leq m$ has non trivial solution i.e. $x_1, x_2, ... x_n$ are not all 0. So for $x_1 a_1 + x_2\alpha_2 + ... + x_n\alpha_n = 0, x_1, x_2, ..., x_n$ are not all 0. Hence S in a linearly dependent set.

Corollony 1 : If V is a finite - dimensional vector space, then any two bares of V have the same number of elements.

Proof :

Since V is finite dimensional, it has a finite basis $\{\beta_1, \beta_2, ... \beta_m\}$. By above theorem every basis of V is finite and contains no more than m elements. Thus if $\{\alpha_1, \alpha_2, ..., \alpha_n\}$ is a basis, $n \leq m$. By the same argument $m \leq n$. Hence m = n.

This corollany allows us to define the dimension of a finite dimensional space V by dimV. This leads us to reformulate, Theorem 1 as fallows :

Corollany 2 : Let V be a finite - dimensional vector space and let *n* = dim V. Then (a) any subset of V which contains more than *n* vectors in linearly dependent (b) no subset of V which contains fewer than n vectors can span V.

Lemma. Let S be a linearly independent subset of a vector space V. suppose β is a vector in V which in not in the subspace spanned by S. then the set obtained by adjoining β to S in linearly independent.

Proof :

Suppose $\alpha_1,....,\alpha_m$ are distinct vectors in S and that $e_1x_1+..+e_m\alpha_m+l\beta=0$. Then $le=0$ for otherwise $\beta=\left(-\frac{e_1}{l}\right)\alpha_1+..+\left(-\frac{e_1}{l}\right)\alpha_m$ and β is in the subspace spanned by S. thus $e_1\alpha_1+..+e_m\alpha_m=0$ and since S is a linearly independent set each e_i=0.

Theorem 2 :

If W is a subspace of a finite dimensional vector space V every linearly independent subset of W is finite and is part of a basis for W.

Proof :

Suppose S_0 is a linearly independent subset of W. if S is a linearly independent subset of W containing S_0. Hence S is also a linearly independent subset of V. Since V in finite dimensional, S contains no more than dim V elements.

We extend S_0 to a basis for W, as follows. If S_0 spans W, then S_0 is a basis for W and our job is done. If S_0 does not span W, we use the preceding lemma to find a vector β, in W such that the set $S_1=S_0\cup\{\beta_1\}$ is independent. If S_1 spans W, our work is over. If not, apply the lemma to obtain a vector β_2 in n such that $S_2=S_1\cup\{\beta_2\}$ is independent. If we continue in this way them by at most dim V steps. we reach a set $S_m=S_0\cup\{\beta_1,....\beta_m\}$ which is a basis for W.

Carollony 1 : If W is a proper subspace of finite - dimensional vector space V, then W is finite dimensional and dim W < dim V.

Proof :

Let us consider W contains a vector $\alpha\neq 0$. So there is a basis of W containing α which contains no more than dim V elements. Hence W is finite-dimensional and dim W $\leq$ dim V. Since W is a proper subspace, there is a vector β in V which is not in W.

Adjoining β to any basis of W, we obtain a linearly independent subset of V. Thus dim W < dim V.

Theorem 3 :

If W_1, and W_2 are finite-dimensional subspaces of a vector spaces V, then W_1+W_2 is finite-dimensional and dim W_1 + dimW_2 = dim $\left(W_1 \cap W_2 + \dim(W_1 + W_2)\right)$.

Proof :

$W_1 \cap W_2$ has a finite basis $\{\alpha_1,...,\alpha_n,\beta_1,...,\beta_m\}$ which is part of a basis $\{\alpha_1,...,\alpha_n,\beta_1,...,\beta_m\}$ for W, and part of a basis $\{\alpha_1,...,\alpha_n,\gamma_1,...,\gamma_n\}$. The subspaces W_1+W_2 is spanned by the vectors $\{\alpha_1,...,\alpha_n,\beta_1,...,\beta_m,\gamma_1,....\gamma_n\}$ and these vectors form an independent set. For suppose

$$\sum x_i\alpha_i + \sum y_j\beta_j + \sum z_r\gamma_r = 0.$$

Then

$-\sum z_r\gamma_r = \sum x_i\alpha_i + \sum y_j\beta_j$ which shows that $\sum z_r\gamma_r$ belong to W_1. As $\sum z_r\gamma_r$ also belongs to W_2 it follows that $\sum z_p\gamma_p = \sum e_i\alpha_i$ for certain sectors $c_1,....,c_k$. Because the set $\{\alpha_1,...,\alpha_n,\gamma_1,...,\gamma_n\}$ is independent, each of the scalars $z_r = 0$.

Thus

$\sum x_i\alpha_i + \sum y_j\beta_j = 0$ and since $\{\alpha_1,...,\alpha_n,\beta_1,...,\beta_m\}$ in also an independent set, each $x_i = 0$ and each $T_j=0$. Thus $\{\alpha_1,...,\alpha_n,\beta_1,...,\beta_m,\gamma_1,....\gamma_n\}$ is a basis for W_1+W_2. Finally dim W_1 + dimW_2 $= (k+m)+(k+n)$

$$= k + (m + k + n)$$

$$= \dim(W_1 \cap W_2) + \dim(W_1 + W_2)$$

Example 1 :

The basis set of the vector space of all 2X2 matrices over the field F is $\left\{\begin{bmatrix}1 & 0\\0 & 0\end{bmatrix},\begin{bmatrix}0 & 1\\0 & 0\end{bmatrix},\begin{bmatrix}0 & 0\\1 & 0\end{bmatrix},\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}\right\}$ So the dimension of that vector space is 4.

Exercises : 2.3

1) Show that the vectors $\alpha_1 = (1,0,-1), \alpha_2 = (1,2,1), \alpha_3 = (0,-3,2)$ form a basis for $\mathbb{R}^3$.

2) Tell with reason whether or not the vectors (2,1,0), (1,1,0) and (4,2,0) form a basis of $\mathbb{R}^3$.

3) Show that the vectors $\beta_1 = (1,1,0)$ and $\beta_2 = (1,i,1+i)$ are in the subspace W of c^3 spanned by (1,0,i) and (1+I, 1, -1), and that β_1 and β_2 form a basis of W.

4) Prove that the space of all *mxn* matrices over the field F has dimension *mn* by exhibiting a basis for this space.

5) If $\{\alpha_1, \alpha_2, \alpha_3\}$ is a basis of $V_3(\mathbb{R})$ show that $\{\alpha_1 + \alpha_2, \alpha_2 + \alpha_3, \alpha_3 + \alpha_1\}$ is also a basis of $V_3(\mathbb{R})$.

Answer

Exercises 2.1

1. (0,0,0,0) 2. Yes

Exercises 2.2

1. (i) not a subspace (ii) Subspace
   (iii) not a subspace (iv) not a subspace
   (v) not a subspace.

2. (i) tpul (ii) false (iii) false (iv) tpul

3. (i) dependent (ii) independent
   (iii) dependent (iv) independent

4. no

❖❖❖❖

3

LINEAR TRANSFORMATION

Unit Structure :

3.0 INTRODUCTION

If X and Y are any two arbitrary sets, there is no obvious restriction on the kind of maps between X and Y, except that it is one-one or onto. However if X and Y have some additional structure, we wish to consider those maps which in, some sense 'pruserve' the extra structure on the sets X and Y. A 'linear transformation' pruserves algebraic operations. The sum of two vectors is mapped to the sum of their images and the scalar multiple of a vector is mapped to the same scalar multiple of its image.

3.1 OBJECTIVE

This chapter will help you to understand

- What is linear transformation.
- Zero and image of it.
- Application of linear transformation in matrix.

3.2 DEFINITION AND EXAMPLES

Let U(F) and V(F) be two vector spaces over the some field F. A linear transformation from U into V is a function T from U into V such that $T(a\alpha + l\beta) = aT(\alpha) + lT(\beta)$ for all α, β in U and for all a, b in F.

Example 1 :

The function $\mathrm{T}: V_3(\mathbb{R}) \to V_2(\mathbb{R})$ defined by $\mathrm{T}(a,b,c)=(a,b)$ $\forall a,b,c \in \mathbb{R}$. Let $\alpha \leq (a_1,b_1,c_1), \beta \approx (a_2,b_2,c_2) \in V_3(\mathbb{R})$. If $a,b \in \mathbb{R}$ then

$$\begin{aligned}
\mathrm{T}(a\alpha + b\beta) &= \mathrm{T}\left[a(a_1,b_1,c_1)+b(a_2,b_2,c_2)\right] \\
&= \mathrm{T}(aa_1+ba_2, ab_1+bb_2, ac,+bc_2) \\
&= (aa_1+ba_2, ab_1+bb_2) \\
&= (aa_1, ab_1)+(ba_2, bb_2) \\
&= a(a_1,b_1)+b(a_2,b_2) \\
&= a\mathrm{T}(a_1,b_1,c_1)+b\mathrm{T}(a_2,b_2,c_2) \\
&= a\mathrm{T}(\alpha)+b\mathrm{T}(\beta)
\end{aligned}$$

$\therefore$ T is a linear transformation from $V_3(\mathbb{R})$ to $V_2(\mathbb{R})$.

Example 2 : The most basic example of linear transformation is

$T: F^n \to F^m$ defined by $\mathrm{T}\begin{pmatrix}\lambda_1\\ \lambda_2\\ .\\ a\\ .\\ \lambda_n\end{pmatrix} = A\begin{pmatrix}\lambda_1\\ \lambda_2\\ .\\ .\\ .\\ \lambda_n\end{pmatrix}$ *where A is a fixed mxn* matrix.

Example 3 : Let V(F) be the vector space of all palimonies over dit f(x) = $a_0 + a_1 x$ be a polymonial of $+ \ldots + a_n x^n EV$ n in the inderminate F. Let us define $Df(x) = a_1 + 2a_1 x + .. + na_n x^{n-1}$ if n > 1 and Df(x) = 0 if f(x) is a constant polynomial. Then the corresponding D from V into V is a linear transformation on V.

Example 5 :

Let $V(\mathbb{R})$ be the vector space of all continuous functions from $\mathbb{R}$ into $\mathbb{R}$. If $f \in V$ and we define T by $(\mathrm{T}f)(x) = \int_o^x f(t)\, dt \forall x \in \mathbb{R}$, then T is a linear transformation from V into V.

Some particular transformation :

1) Zero Transformation : Let U(F) and V(F) be two vector spaces. The functions T, from U into V defined by $T(\alpha)=0$ (zero vector of

V) $\forall \alpha \in U$ in a linear transformation from U into V. it is called zero transformation and is denoted if $\widehat{O}$.

2) Identity operator : Let V(F) be a vector space. The function I form V into V defined by $\mathrm{I}(\alpha)=\alpha \forall \alpha \in V$ is a linear transformation from V into V, I is known as identity operator on V.

3) Negative of a linear transformation : Let U(F) and V(F) be two vector spaces. Let T be a linear transformation from U into V. The corresponding - T defined by $(-\mathrm{T})(\alpha)=-[\mathrm{T}(\alpha)] \forall \alpha \in U$ is a linear transformation from U into V. - T is called the negative of the linear transformation of T.

Some properties of linear transformation :

Let T be a linear transformation from a vector space U(F) into a vector space V(F). Then

i) T (O) = O where O on the left hand side is zero vector of U and O on the right hand side

ii) $\mathrm{T}(-\alpha)=-\mathrm{T}(\alpha) \forall \alpha \in U$

iii) $\mathrm{T}(\alpha-\beta)=\mathrm{T}(\alpha)-\mathrm{T}(\beta) \forall \alpha, \beta \in U$

iv) $\mathrm{T}(a_1\alpha_1+a_2\alpha_2+\ldots+a_n\alpha_n)=a_1\mathrm{T}(\alpha_1)+a_2\mathrm{T}(\alpha_2)+\ldots+a_n\mathrm{T}(\alpha_n)$
where $\alpha_1, \alpha_2, \ldots . \alpha_n \in U$ and $a_1, a_2, \ldots a_n \in F$

3.2 IMAGE AND KERNEL OF A LINEAR TRANSFORMATION

Definition : Let U(F) and V(F) be two vector spaces and let T be a linear transformation from U into V. Then the range of T is the set of all vectors in V such that $\beta=T(\alpha)$ for some α in U. This is called the image set of U under T and by $\mathrm{I_m}$ T, i.e. $I_m\mathrm{T}=\{\mathrm{T}(\alpha): \alpha \in U\}$.

Definition : Let U(F) and V(F) be two vector spaces and let T be a linear transformation from U into V. Then the kernel of T written as kept T is the set of all vectors α in U such that $\mathrm{T}(\alpha)=O$ (zero vector of V). Thus ker T = $\{\alpha \in U : T(\alpha)=O \in V\}$, ker T is also called null space of T.

Theorem 1:

If U(F) and V(F) are two vector spaces and T is a linear transformation from U into V then i) Ker T is a subspace of U (ii) I_mT is a subspace of V.

Proof :

i) ker $\mathrm{T} = \{\alpha \in U : \mathrm{T}(\alpha) = O \in V\}$ Since $\mathrm{T}(O) = O \in V$, therefore at least $0 \in$ ker T. Thus ker T is a non-empty subset of U. Let $\alpha_1, \alpha_2 \in$ ker T, Then $T(\alpha_1) = O, T(\alpha_2) = O$.

Let $a, b \in F$. Then $a\alpha_1 + b\alpha_2 \in U$ and $\mathrm{T}(a\alpha_1 + b\alpha_2) = a\mathrm{T}(\alpha_1) + lT(\alpha_2)$ $= aO + bO = O + O = O \in V \therefore a\alpha_1 + b\alpha_2 \in \ker \mathrm{T}$.

Thus Ker T is a subspace

ii) Obviously I_m T is a non-empty subset of V.
Let $\beta_1, \beta_2 \in I_m\mathrm{T}$. The $\exists \alpha_1, \alpha_2$ EU such that $\mathrm{T}(\alpha_1) = \beta_1, \mathrm{T}(\alpha_2) = \beta_2$
Then $a\beta_1 + b.\beta_2 = a\,\mathrm{T}(\alpha_1) + b.\mathrm{T}(\alpha_2)$

$$= \mathrm{T}(x_1 + b.\alpha_2)$$

Now, U is a vector space.
$\therefore$ $\alpha_1, \alpha_2 \in U$ and $a, b \in F$
$\Rightarrow$ $a\,\alpha_1 + b\alpha_2 \in U$
Consequently $\mathrm{T}(a\,\alpha_1 + b.d_2) = a\,\beta_1 + b\,\beta_2 \in \mathrm{Im}\,\mathrm{T}$
Thus, $\mathrm{Im}\,T$ is s subspace of V.

Theorem 2 : Rank nullify theorem.

Let U(F) and V(F) be two vector spaces and T be a linear transformation. Suppose U is finite dimensional. Then, $\dim U = \dim Ker\,T + \dim \mathrm{Im}\,T$

Proof : If $ImT = \{0\}$, then $kerT = V$ and theorem is proved for the trivial case.
Let $\{v_1, v_2, ..., v_r\}$ be a basis of ImT for $r \geq 1$.
Let $v_1, v_2, ..., v_r \in U$ such that $v_i = T(v_i)$
Let $\{v_1, v_2, ..., v_q\}$ be the basis of $kerT$.
We have to show that $\{v_1, v_2, ..., v_r, v_1, v_2, ..., v_q\}$ forms a basis of U.

Let $u \in U$. Then $T(u) \in ImT$. Hence there are real numbers $c_1, c_2, ..., c_r$ such that

$$T(u) = v_1 v_1 + v_2 v_2 + ... + v_r v_r$$
$$= v_1 T(u_1) + v_2 T(u_2) + ... + v_r T(u_r)$$
$$= T_1(v_1 u_1 + v_2 u_2 + ... + v_r u_r)$$

$$\therefore \; T(u - \{v_1 u_1 + v_2 u_2 + ... + v_r u_r\}) = 0$$
$$\therefore \; u - \{v_1 u_1 + v_2 u_2 + ... + v_r u_r\} \in ker\, T$$

This would again mean that there are numbers $a_1, a_2, ..., a_q$ such that

$u - \{v_1 u_1 + v_2 u_2 + ... + v_r u_r\}$

$= a_1 u_1 + a_2 u_2 + ... + a_q u_q$

i.v. $u = v_1 u_1 + v_2 u_2 + ... + v_r u_r + a_1 u_1 + a_2 u_2 + ... + a_q u_q$

So, u is generated by $u_1, u_2, ..., u_r, u_1, u_2, ..., u_q$.

Next to show that these vectors are linearly independent, let $x_1, x_2, ..., x_r, y_1, y_2, ..., y_q$ be the real numbers, such that

$x_1 u_1 + x_2 u_2 + ... + x_r u_r + y_1 u_1 + y_2 u_2 + ... + y_q u_q = 0$

Then $0 = T(0)$

$$= T(x_1 u_1 + ... + x_r u_r + y_1 u_1 + ... + y_q u_q)$$
$$= x_1 T(u_1) + ... + x_r T(u_r) + T(y_1 u_1 + ... + y_q u_q)$$
$$x_1 v_1 + ... + x_r v_r = 0$$

But $v_1, ..., v_r$ being basis of $\mathrm{Im}T$ are linearly independent.

So, $x_1 = x_2, ..., x_r = 0$.

$\therefore \; y_, u_1 + ... + y_q u_q = 0$

By the same argument

$y_1 = y_2 = ... = y_q = 0$

So, $u_1, u_2, ..., u_r, u_1, ..., u_q$ are linearly independent.

Thus, $\dim U = r + q$

$$= \dim \mathrm{Im} T + \dim \ker T$$

Example 1 : Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ is defined by $T(x, y, z) = (x + 2y - z,\; y + z,\; x + y - 2z)$.

Let us show that T is a linear transformation. Let us also find $\ker T, \operatorname{Im} T$, their bases and dimensions.

To check the linearity let $a, b \in \mathbb{R}$ and $(x_1, y_1, z_1), (x_2, y_2, z_2) \in \mathbb{R}$

$$T\left(a(x_1, y_1, z_1) + b(x_2, y_2, z_2)\right)$$
$$= T\left(ax_1 + bx_2, ay_1 + by_2, az_1 + bz_2\right)$$
$$= (ax_1 + bx_2 + 2ay_1 + 2by_2 - az_1 - bz_2, ay_1 + by_2 + az_1 + bz_2, ax_1 + bx_2 + ay_1 + by_2 - 2az_1 - 2bz_2)$$
$$= \left(a(x_1 + 2y_1 + z_1) + b(x_2 + 2y_2 - z_2), a(y_1 + z_1) + b(y_2 + z_2), a(x_1 + y_1 - 2z_1) + b(x_2 + y_2 - 2z_2)\right)$$
$$= a(x_1 + 2y_1 - z_1, y_1 + z_1, x_1 + y_1 - 2z_1) + b(x_2 + 2y_2 - z_2, y_2 + z_2, x_2 + y_2 - 2z_2)$$
$$= aT(x_1, y_1, z_1) + bT(x_2, y_2, z_2).$$

Hence, T is a linear transformation.
Now, $(x, y, z) \in \ker T$ iff $T(x, y, z) = (0, 0, 0)$
i.e. iff $(x + 2y - z,\ y + z,\ x + y - 2z) = (0, 0, 0)$
This gives us

$$x + 2y - z = 0$$
$$y + z = 0$$
$$x + y - 2z = 0$$

By second equation $y = -z$ substituting in third equation

$$x - z - 2z = 0 \ \Rightarrow\ x = 3z$$
$$\therefore \frac{x}{3} = \frac{y}{-1} = \frac{z}{1}$$
$$(x, y, z) = (3, -1, 1)z$$

$$\therefore \ker T = \left\{z(3, -1, 1) : z \in \mathbb{R}\right\}$$

So, $\ker T$ is generated by $(3, -1, 1)$. Hence its basis is $\{(3, -1, 1)\}$ and dimension is 1.

Now, $T(x, y, z) = (x + 2y - z,\ y + z,\ x + y - 2z)$
$$= x(1, 0, 1) + y(2, 1, 1) + z(-1, 1, -1)$$

But $(-1, 1, -2) = -3(1, 0, 1) + 1(2, 1, 1)$, Hence, $T(x, y, z) = x(1,0,1) + y(2,1,1) + z\{-3(1,0,1) + 1(2,1,1)\}$

$$= (x - 3z)(1,0,1) + (y + 2)(2,1,1)$$

$\therefore$ Basis of $\operatorname{Im} T = \{(1,0,1),(2,1,1)\}$ and its dimension is 2.

Exercise : 3.1

1. Let F be field of complex numbers and let T be the function from F^3 into F^3 defined by $T(x_1, x_2, x_3) = (x_1 - x_2 + 2x_3, 2x_1 + x_2 - x_3, x_1 - 2x_2)$. Verify that T is a linear transformation.

2. Show that the following maps are not linear.
   i) $T : \mathbb{R}^3 \to \mathbb{R}^3$; $T(x, y, z) = (x, y, 0)$
   ii) $T : \mathbb{R}^2 \to \mathbb{R}^2$; $T(x, y) = (x^2, y^2)$
   iii) $F : \mathbb{R}^3 \to \mathbb{R}^2$; $F(x, y, z) = (|x|, 0)$
   iv) $S : \mathbb{R}^2 \to \mathbb{R}$; $S(x, y) = |x + y|$

3. In each of the following find $T(1,0)$ and $T(0,1)$ where $T : \mathbb{R}^2 \to \mathbb{R}^2$ is a linear transformation.
   i) $T(3,1) = (1,2)$, $T(-1,0) = (1,1)$
   ii) $T(4,1) = (1,1)$, $T(1,1) = (3,-2)$
   iii) $T(1,1) = (2,1)$, $T(-1,1) = (6,3)$

4. Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be the linear transformation defined by $T(x, y, z) = (x + 2y - z,\ y + z,\ x + y - 2z)$. Find a basis and the dimension of $\operatorname{Im} T$ and $\ker T$.

3.3 ALGEBRA ON LINEAR ALGEBRA

Definition :

Let F be a field. A vector space V over F is called on linear algebra over F if there is defined an additional operation in V called multiplication of vectors and satisfying the following postulates.

1. $\alpha\beta \in V \forall \alpha, \beta \in V$
2. $\alpha(\beta\gamma) = (\alpha\beta)\gamma \forall \alpha, \beta, \gamma \in V$
3. $\alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma$ and $(\alpha + \beta)\gamma = \alpha\gamma + \beta\gamma \forall \alpha, \beta, \gamma \in V$
4. $e(\alpha\beta) = (c\gamma)\beta = \alpha(e\beta) \forall \alpha, \beta \in V$ and $e \in F$

If there is an element 1 in V such that $1\alpha = \alpha = \alpha 1 \forall \alpha \in \mathrm{V}$, then we call V a linear algebra with identity over F. Also 1 is then called the identify of V. The algebra V is commutative if $\alpha\beta = \beta\alpha \forall \alpha, \beta \in V$

Polynomials : Let T be a linear transformation on a vector space V(F). Then TT is also a linear transformation on V, we shall write $T^1 = T$ and $T^2 = TT$. Since the product of linear transformation is an associative operation, therefore if *m* is a positive integer, we shall define T^m = TTT … upto *m* times. Obviously T^m is a linear transformation on V. Also we define $T^0 = I$ (identity transformation).

If *m* and *n* are non-negative integers, it can be easily seen that $T^m T^n = T^{m+n}$ and $\left(T^m\right)^n = T^{mn}$, The set L(V,V) of all linear transformation on V is a vector space over the field F. If $a_0, a_1, ..., a_n e F\ a_1 T + a_2 T^2 + .. + a_n T^n \in L(V,V)$ i.e. P(T) is also a linear transformation on V because it is a linear combination over F of elements of L(V,V). We call P(T) as a polynomial in linear transformation T. The polynomials in a linear transformation behave like ordinary polynomials.

3.4 INVERTIBLE LINEAR TRANSFORMATION

Definition : Let U and V be vector spaces over the field F. Let T be a linear transformation from U into V such that T is one-one onto. Then T is called invertible.

If T is one-one and onto then we define a function from V into U, called the inverse of T and denoted by T^{-1} as follows :

Let β be any vector in V. Since T is onto, therefore $\beta \in V \Rightarrow \exists \alpha \in U$ such that $T(\alpha) = \beta$.

Also α determined in this way is a unique element of U because T is one-one and therefore $\alpha_0, \alpha \in U$ and $\alpha_0 \neq \alpha \Rightarrow \beta = T(\alpha) \neq T(\alpha_0)$ we define $T^{-1}(\beta)$ to be α. Then $T^{-1} : V \rightarrow V$ such that $T^{-1}(\beta) = \alpha \Leftrightarrow T(\alpha) = \beta$. The function T^{-1} is itself one-one and onto.

Properties :

1. T^{-1} is also a linear transformation from V into U.
2. Let T be an invertible linear transformation on a vector space V(F). Then T^{-1}T = I = TT^{-1}.
3. If A, B and C are linear transformations on a vector space V(F) such that AB = CA = I, then A is invertible and A^{-1} = B = C.
4. Let A be an invertible linear transformation on a vector space V(F). The A possesses unique inverse. (The proof of the above properties are left for the students)

Example :

If A is a linear transformation on a vector space V such that $A^2 - A + I = \hat{0}$, then A is invertible.

$A^2 - A + I = \hat{0}$, then A^2 - A = - I. first we shall prove that A is one-one.

Let $\alpha_1, \alpha_2 \in V$. Then $A(\alpha_1) = A(\alpha_2)$

$\Rightarrow A[A(\alpha_1)] = A[A(\alpha_2)]$

$\Rightarrow A^2(\alpha_1) = A^2(\alpha_2)$

$\Rightarrow A^2(\alpha_1) \neq A(\alpha_1) = A^2(\alpha_2) - A\ (\alpha_2)$

$\Rightarrow (A^2 - A)(\alpha_1) = (A^2 - A)(\alpha_2)$

$\Rightarrow (-I)(\alpha_1) = (-I)(\alpha_2)$

$\Rightarrow -[I(\alpha_1)] = -[I(\alpha_2)]$

$\Rightarrow -\alpha_1 = -\alpha_2 \Rightarrow \alpha_1 = \alpha_2$

$\therefore$A is one-one.

Now to prove that A is onto Let $\alpha \in V$. Then $\alpha - A(\alpha) \in V$.

We have $A[\alpha - A(\alpha)] = A(\alpha) - A^2(\alpha)$

$= (A - A^2)(\alpha)$

$= I(\alpha) = \alpha$

Thus $\alpha \in V \Rightarrow \exists \alpha - A(\alpha) \in V$ such that $A[\alpha - A(\alpha)] = \alpha$

$\therefore$A is onto.

Hence A is invertible.

Check your progress :

1. Show that the identity operator on a vector space is always invertible.
2. Describe $T:\mathbb{R}^3 \rightarrow \mathbb{R}^4$ which has its range the subspace spanned by the vectors (1,2,0,-4), (2,0,-1,3)
3. Let T and U be the linear operators on $\mathbb{R}^2$ defined by $T(a,b)=(b,a)$ and $U(a,b)=(a,b)$. Give rules like the one defining T and U for each of the transformations U + T, UT, TU, T^2, U^2
4. Show that the operator T on $\mathbb{R}^3$ defined by T (x,y,z) = (x+z, x-z, y) is invertible.

3.5 REPRESENTATION OF TRANSFORMATION BY MATRICES

Matrix of a linear transformation :

Let U be an *n*-dimensional vector space over the field F and Let V be an *m*-dimensional vector space over F. Let $B=\left\{a_{_1},....,a_n\right\}$ and $B=\left\{\beta_1,....,\beta_n\right\}$ be ordered bases for U and V respective. Suppose T is a linear transformation from U into V is a basis of into V.

Now for $\alpha_j \in U, T\left(\alpha_j\right) \in V$ and .

$$\therefore T\left(\alpha_j\right) = a_{1j}\beta_1 + a_{2j}\beta_2 + ... + a_{mj}\beta_m$$

$$= \sum_{i=1}^{n} a_{ij}\beta_i$$

This gives rise to an *mxn* matrix $[a_{jj}]$ whose jth column represents the numbers that appear in the presentation of $T\left(\alpha_j\right)$ as a combination of elements of B. Thus the first column is $\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} = \left[a_{11}, a_{21}, ..., a_{m1}\right]^T$, the second column is $(a_{12}, \ldots, a_{m2})^T$ and so on. We call [aij] the matrix of T with respect to the ordered basis B, B of U, V respectively. We will denote the matrix so induced by $\left[m(T)\right]_{B'}^{B^1}$.

Example :

Let $T:\mathbb{R}^4 \to P_1(\mathbb{R})$ given by $T\left(x_1,x_2,x_3,x_4\right)=x_1+x_3+\left(x_2+x_4\right)x.$ The basis of $\mathbb{R}^4$ be $B_1=\left\{(1,1,1,1),(1,1,1,0),(1,1,0,0),(1,0,0,0)\right\}$ and that of $P_1(\mathbb{R})$ be $B_2=\{1+x,1-x\}.$

$$T(1,1,1,1)=2+2x=2(1+x)+0(1-x)$$

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

$$T(1,1,0,0)=1+x=1(1+x)+0(1-x)$$

$$T(1,0,0,0)=1=\frac{1}{2}(1+x)+\frac{1}{2}(1-x)$$

Then $\left[m(T)\right]_{B_1}^{B_2}=\begin{pmatrix} 2 & \frac{3}{2} & 1 & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix}$

3.6 MATRIX OF SUM OF LINEAR TRANSFORMATION :

Theorem : Let $T_1:V\to W$ and $T_2:V\to W$ be two linear transformation. Let $B_1=\left\{\upsilon_1,\upsilon_2,...,\upsilon_m\right\}$ and $B_2=\left\{\omega_1,\omega_2,...,\omega_m\right\}$ be the bases of V and W respectively.

Then $\left[m\left(T_1+T_2\right)\right]_{B_2}^{B_1}=\left[m\left(T_1\right)\right]_{B_2}^{B_1}+\left[m\left(T_2\right)\right]_{B_2}^{B_1}$

Proof :

For $\upsilon_i\in V, T_1\left(\upsilon_i\right)\in W$ and $T_2\left(\upsilon_i\right)\in W$. Since $B_2=\left\{\omega_1,\omega_2,...,\omega_n\right\}$ is the basis of W, $T_1\left(\upsilon_i\right)=\sum_{j=1}^{n}a_{ij}w_j$ and $T_2\left(\upsilon_i\right)=\sum_{j=1}^{n}b_{ij}w_j$; $\forall i=1,....,m.$

Now $\left(T_1+T_2\right)\left(\upsilon_j\right)=T_1\left(\upsilon_j\right)+T_2\left(\upsilon_j\right)$

$$=\sum_{n=i}^{n}a_{ij}W_j+\sum_{j=i}^{N}b_{ij}w_j$$

$$=\sum_{j=i}^{n}\left(a_{ij}+b_{ij}\right)w_j$$

$$=\sum_{j=i}^{w}c_{ij}w_j$$

Where

$$c_{ij} = a_{ij} + b_{ij}$$

$$\therefore [c_{ij}] = [a_{ij}] + [c_{ij}]$$

$$\Rightarrow [c_{ij}]^T = [a_{ij}]^T + [b_{ij}]^T$$

$$\therefore [m(T_1 + T_2)]_{B_1}^{B_2} = [m(T_1)]_{B_1}^{B_2} + [m(T_2)]_{B_1}^{B_2}$$

Matrix of scalar multiplication of linear transformation :

Theorem : Let $T : V \to W$ be a linear transformation. Let $B_1 = \{\upsilon_1, ..., \upsilon_m\}$ and $B_2 = \{\omega_1, \omega_2, ..., \omega_m\}$ be the bases of V and W respectively.

Then $[m(kT)]_{B_1}^{B_2} = k[m(T)]_{B_1}^{B_2}, k \in \mathbb{R}$.

For $\upsilon_i \in V, T(\upsilon_i) \in W$ and B_2 is the basis of W.

So, $T(\upsilon_i) = \sum_{j=1}^{n} a_{ij}\omega; \forall 1 \le i \le m$

Now $T(k\upsilon_i) = kT(\upsilon_i) = k = \sum_{j=1}^{n} a_{ij}\omega;$

$$= \sum_{j=1}^{n} (ka_{ij})\omega;$$

$$= \sum_{j=1}^{n} b_{ij}\omega;$$

Where $b_{ij} = Ka_{ij}$

$$\Rightarrow [b_{ij}]^T = k[a_{ij}]^T$$

$$\Rightarrow [m(kT)]_{B_1}^{B_2} = k[m(T)]_{B_1}^{B_2}.$$

Matrix of composite linear transformation :

Theorem : Let $T : V \to W$ and $S : W \to U$ be two linear transformations. Let $B_1 = \{\upsilon_1, ..., \upsilon_m\}, B_2 = \{\omega_1, ..., \omega_n\}$ and $B_3 = \{u_1, ..., u_k\}$ be the bases of V, W and U respectively.

Then $[m(ST)]_{B_1}^{B_2} = [m(S)]_{B_1}^{B_2} [m(T)]_{B_1}^{B_2}$

Proof :

For $\upsilon_i \in V$ and $w_j \in W$,
$T(\upsilon_i) \in W$ and $S(w_j) \in U$

$$\therefore T(\upsilon_i) = \sum_{j=1}^{n} a_{ij} w_j$$

$$S(w_i) = \sum_{r=1}^{k} b_{ir} u_r$$

$$S.T(\upsilon_i) = S(T(\upsilon_i)) = S\left(\sum_{j=1}^{n} a_{ij}; w_j\right)$$

$$= \sum_{j=1}^{n} a_{ij} S(w_j)$$

$$= \sum_{j=1}^{n} a_{ij} \left(\sum_{j=1}^{k} b_{jr} u_p\right)$$

$$= \sum_{j=1}^{n} \sum_{j=1}^{k} (a_{ij} b_{jr}) u_r$$

$$= \sum_{r=1}^{k} \sum_{j=1}^{n} (a_{ij} b_{jr}) u_r$$

$$= \sum_{r=1}^{n} c_{ir} u_r$$

Where

$$c_{ir} = \sum_{j=1}^{n} a_{ij} b_{jr}$$

Which is the $(i, p)^{th}$ element of the matrix.

$$[c_{ir}] = [a_{ij}][b_{jr}]$$

$$[c_{ir}]^T = [b_{jr}]^T [a_{ij}]^T$$

$$\therefore [m(S.T)]_{B_1}^{B_2} = [m(S)]_{B_1}^{B_2} [m(T)]_{B_1}^{B_2}$$

Example :

$T : \mathbb{R}^2 \to \mathbb{R}^2$ and $S : \mathbb{R}^2 \to \mathbb{R}^2$ be two liner raw formations defined by
T (X, Y) = (X + Y, X – Y) and
S(X, Y) = (ZX + Y, X + 2Y)
Let in basis is *{(1, 2) , (0, 1)}.*

Then *T(1, 2) = (3, -1) = a(1, 2) + C(0, 1)* which implies *a = 3, b = -7*

T(0, 1) = (1, -1) = $p(1,2)+q(0,1)$ which implies $p=1, q=-3$

So $\left[m(T)\right]_B^B=\begin{bmatrix} \frac{3}{-7} & \frac{1}{-3} \end{bmatrix}$

Similarly $\left[m(S)\right]_B^B=\begin{bmatrix} \frac{4}{-3} & \frac{1}{0} \end{bmatrix}$

$$\left[m(S.T)\right]_B^B=\left[m(S)\right]_B^B\left[m(T)\right]_B^B$$

$$=\begin{bmatrix} \frac{3}{-7} & \frac{1}{-3} \end{bmatrix}\begin{bmatrix} \frac{4}{-3} & \frac{1}{0} \end{bmatrix}$$

$$=\begin{bmatrix} \frac{s}{-q} & \frac{1}{-3} \end{bmatrix}$$

(SOT) (X, Y) = S(T(X,Y) = S(n+y, n-q)
= (3 x + y , 3 x – y)

By above way are can find $\left[m(SOT)\right]_B^B$ and verify the result.

Exercise 3.2

1. Let $T:\mathbb{R}^3\to\mathbb{R}^3$ and $S:\mathbb{R}^3\to\mathbb{R}^3$ be two lincer transformations difined by

$T(x,y,z)=(x,2y,3z)$ and

$S(x,y,z)=(x+y,y+z,z+x)$.

Let $B=\{(1,0,0),(0,1,0),(0,0,1)\}$.

Verify that

$$\left[m(SOT)\right]_B^B=\left[m(s)\right]_B^B\left[m(T)\right]_B^B$$

2. Let $T_1:\mathbb{R}^2\to\mathbb{R}^2$ and $T_2:\mathbb{R}^2\to\mathbb{R}^2$

Be two linear transformations defined by

$T_1(x,y)=(2x+3y,x-2y)$ and

$T_2(x,y)=(x+y,x-y)$,

$B=\{(1,0),(0,1)\}$ be the basis of $\mathbb{R}^2$. Show that B

$$\left[m(T_1+T_2)\right]_B^B=\left[m(T_1)\right]_B^B+\left[m(T_2)\right]_B^B$$

RANK OF MATRIX and linear of transformation :

Row Space and column space Let is consider a matrix A as follows

$$A = \begin{bmatrix} \frac{4}{3} & \frac{6}{0} & \frac{a}{4} & \frac{z}{5} \end{bmatrix}$$

We can consider A as a matrix of two vectans in $\mathbb{R}^4$ an as a matrix of four rentors in $\mathbb{R}^2$.

We will consider linear span of two rectons i.e., $W_p = L\{(4,6,9,2),(3,0,4,5)\}$.

It is called naw space of matrix A. Similany colum space of A is represented by $W_c = L\{(4,3),(6,0),(9,4),(2,5)\}$.

Definition :

Row space and column space : Let A be a matrix of order mxn. Then the subspace of $\mathbb{R}^n$ generated by the now vectors of A is called the now space and the subspace of $\mathbb{R}^m$ generated by the calumn vectors of A is called the Column space of A.

Example :

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

Here now space is $\{R_1, R_2, R_3\}$ Where $R_1 = (1,0,1,0), R_2 = (0,1,0,1), R_3 = (1,1,1,0)$. The set $\{R_1, R_2, R_3\}$ is line on by independent.

So now space = L $\{\mathbb{R}_1 \mathbb{R}_2 \mathbb{R}_3\}$.
Hence dim (now space) = 3.

Now calumniation we have two vectans $C_1 = \{1,0,1\}, C_2 = \{0,1,1\}, C_3 = \{1,0,1\}, C_4 = \{0,1,0\}$.
Here $C_1 = C_3$ and $C_1, C_2, C_4\}$ are line an by independent,
$\therefore$ column space $= L\{C_1, C_2, C_4\}$
$\therefore$ dim (column space) = 3.

The dim of the now space of a matrix A is called now rank of A and the dim of the column space of A in called calumn rank.

In enery matrix dim (now space) = dim (calumn space) if now rank = calumn rank.

Rank of zero matrix is new.

Rank of identify matrix of order n is n.

Rank of A^t = Rank of A where A^t is transpore of A.

or mxn matrix A, now space is subspace of $\mathbb{R}^n$.

$\therefore$ now rank $\leq$ n.

Similanery column rank $\leq m$

$\therefore$ rank of $A \leq min\ (m, n\}$.

Example :

$$A = \begin{bmatrix} \overline{1} & \overline{2} & \overline{3} \\ \underline{4} & \underline{5} & \underline{6} \\ 7 & 8 & 9 \end{bmatrix}$$

Here $R_1 = (1, 2, 3), R_2 = (4,5,6), R_3 = (7,8,9)$

$R_3 = ZR_2 - R_1$

$\{R_1, R_2\}$ are lance by independent

$\therefore$ now rank = Z

$\therefore$ rank of A = Z

Change of Basis.

Sometimes it is imperative to change the basis in representing a linear transformation T, because relative to this new basis, the representation of T may become very much simplified. We shall thru fare turn own attention to established are important result concerning the matrix represecutations of a liner transformation when the basis in changed.

Theaum :

If two sets of vectors $X = \{ x_1, x_2, .. x_n \}$ and $\hat{x} = \{ \hat{x}_1, \hat{x}_2, ... \hat{x}_n \}$ are the bases of a rectum space V_n, these there exists a nonsingular matrix B = [bij] such that $x_i = bij\ x_1 + b_{2i}\ x_2 + .. + bni\ x_n$,

$i = 1, 2, ... n.$

The nonsingular matrix B is defined as a transformation matrix in V_n.

Proof :

Suppose that X and $\hat{x}$ are two bases in V_n. Then $\hat{x}$ i s (I = 1,2, .. n) can be expressed $x_1, n_2, \ldots, x_n$ *i.e.*

$\hat{x}_i = b_{ij}\ x_1 + b_{zi}\ x_2 + .. + b_{ni}\ x_n,$

$i = 1, 2, ,, , n$

Where b_{ij} s are realer.

Let is define its matrix $B = [b_1\ b_2 \ldots b_n]$

Where

$$b_i = [b_{ij}\ b_{zi} \ldots b_{ni}]^T$$

Is an vector.

We have to show that B is non singular.

For realer $x_1 \ldots x_2$ we write, $x_1\ \hat{x}_1 + \ldots .\ x_n\ x_n$

$= x_1\ (b_{11}\ x_1 + \ldots + b_{n1} x_n) + \ldots + x_n\ (b_{1n}\ x_1 + .. + b_{nn} x_n)$

$= (x_1 b_{11} + .. + x_n b_{1n})^{x_1} + \ldots + (x_1\ b_{n1} + .. + x_n b_{nn})^{\ x_n}$

Now

$x_1 b_1 + \ldots + x_n\ b_n = 0$

Implies

$x_1 b_{i1} + .. + x_n b_{in} = 0, i = 1, 2, \ldots, n$ Substituting this in the above equation we have

$x_1\ \hat{x}_1 + \ldots + x_n\ \hat{x}_n = 0.$

Since, $\hat{x}_1, .., \hat{x}_u$ are lines only independent, it follows that $x_1 = \ldots . = x_n = 0$ and hence $b_1, \ldots b_n$ are linearly independent.

$\therefore$ B is non singular.

Theorem :

Suppose that A is an mxn matrix of the linear transformation $T : V_n \rightarrow W_m$ with respect to the bases $X = \{n_1, \ldots, x_n\}$ and

$y = \{y_1, ..., y_m\}$. If Â in an mxn matrix of T with respect to diffluent bases $\hat{x} = \{\hat{x}_1, ... \hat{x}_n\}$ and $\hat{y} = \hat{y}_1, ... \hat{y}_m\}$ than there exist non singular matrix B and C of order n and m respectively, such that $\widehat{A} = C^{-1} AB.$

Proof :

If $A = [a_{ki}]$ in the matrix of T with aspects to the bases X and Y, we have the lunation, $T(x_i) = a_{1i}\ y_1 + a_{2i}\ y_2 + ... + \ a_{mi}\ y_m$ similarly, for $\widehat{A} = [\hat{a}_{ij}]$, we can write, $T(\hat{x}_j) = \widehat{a_1}, \hat{j}, + \widehat{a_2}, \widehat{y_2} + .. + \widehat{a_m}, \widehat{y_m}$.

By the previous formula thus exist non singular coordinate transformation matrix $B = [b_{ij}]$ and $C = [e_{ki}]$ satisfying.

$$\widehat{j_i} = b_{ij}\ x_1 + b_{zj}\ x_2 + .. + b_{nj}\ x_n$$

$$\widehat{y_i} = c_{ij} y_1 + c_{zi} y_2 + .. + c_{mi}\ y_m$$

Hence $T(\widehat{x_j}) = \sum_{c=1}^{m} \widehat{a_i}\,\widehat{y_i} = \sum_{i=1}^{m} \widehat{a_i} \sum_{k=1}^{m} c_{ki} y_n$

$$\sum_{k=1}^{m} \left(\sum_{c=1}^{m} c_{ki}\, \widehat{x_{ij}} \right) y_k$$

Alternatively sinner T in liner, we have,

$$T(\widehat{x_j}) = \sum_{i=1}^{w} b_{ij}\ T(x_i)\,\mathrm{T}$$

$$= \sum_{i=1}^{n} b_{ij} \sum_{n=1}^{m} a_{ki} y_n$$

$$\sum_{k=1}^{m} \left(\sum_{i=1}^{n} a_{ki} b_{ij} \right) y_n$$

$$i.e.\ c\,\hat{a} = AB$$

$$\Rightarrow \hat{a} = c^{-1} AB$$

Ex. A Linear transformation $T : \mathbb{R}^3 \to \mathbb{R}^2$ in defined by

$$T(x_1, x_2, x_3) = (x_1 + x_2 + x_3, zx_2 + x_3)$$

Of the bares in $\mathbb{R}^3$ are

$$v(\ x_1, x_2, x_3) = \{\ (\ 1,0,0)\ (,\ 0,1,0)\ (,\ 0,0,1)\ \}$$

$\hat{v}(\hat{x}_1, \hat{x}_2, \hat{x}_3) = \{(2,2,1), (0,1,0), (1,0,1)\}$ and these in $\mathbb{R}^2$ are

$w(y_1, y_2) = \{(2,1),(1,1)\}$

$\hat{w}(\hat{y}_1, \hat{y}_2) = \{(1,0),(1,1)\}$

Here we will find the matrix A w.p.t. the bases defined by V and W, and $\hat{A}$ w.p.t. $\hat{v}$ and $\hat{w}$ of the linear transformation T. Also we have to determine non singhan matrixes B and C such that $\hat{A} = C^{-1} AB.$

Here

$T(1,0,0) = (1,0) = 1(2,1) - 1(1,1)$

$T(0,1,0) = (1,2) = -1(2,1) + 3(1,1)$

$T(0,0,1) = (1,1) = 0(2,1) + (1,1)$

$\therefore [m(T)]^w = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 3 & 1 \end{bmatrix}$

i.e. $T(V) = WA.$

Similarly considering the bases $\hat{V}$ of $\mathbb{R}^3$ and $\hat{w}$ of $\mathbb{R}^2$ we find

$$[m(T)]_{\hat{v}}^{\hat{w}} = \begin{bmatrix} 0 & -1 & 1 \\ 5 & 2 & 1 \end{bmatrix} = \hat{A}$$

$s.t. T(\hat{v}) = \hat{w}\hat{A}$

The matrix B and C are determined by the change of relationship as

$\hat{v} = VB$ and $\hat{w} = wc$

$$\therefore \hat{v} = VB \Rightarrow \begin{bmatrix} 2 & 0 & 1 \\ 2 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

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

$\hat{w} = wc.$

$$i \therefore \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

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

We can such the result $c\widehat{A} = AB$

LINEAR FUNCTIONALS : DUAL SPACE

Definition

A linear transformation T from a vector space V one a field $\mathbb{F}$ into this field $\mathbb{F}$ of scalars in called a linear functional on the space V. The set of all function on V in a victor space defined as the dual space of V and in denoted by V.

Example :

Let V be the vector space of all mal valued function internals over the interval $a \leq t \leq b$. Then the transformation $f : v \rightarrow \mathbb{R}$ defined by $x(t) 1 \rightarrow \int_a^t x(t)\, dt$ is a linear functional on V. This mapping f assigns to each internals function *x(e)* a real number on the internal $a \leq t \leq c$.

Example :

Let $\{x_1, x_2, \ldots, x_n\}$ be a basic of the n dimensional vector space v one $\mathbb{F}$. Any vector x in V can be reprinted by $x = x_1\, x_1 + x_2\, x_2 + \ldots + x_n\, x_n$, when $x_i\, s$ one realans in $\mathbb{F}$ we now consider a fined vector Z in V and spurned it by $z = c_1\, n_1 + c_2\, x_2 + \ldots. + c_n\, x_n$ when $c_i\,'s$ are sealers in $\mathbb{F}$.

Denote by $w = [e_1\, e_2 \ldots e_n]^T$ *and* $n = [x_1\, x_2 \ldots x_n]^T$, the coordinate vectors of z and x respective by. Thus the linear transformation $f_2 : v \rightarrow \mathbb{F}$ defined by $f_z(x) = c_1\, x_1 + c_2\, x_2 + \ldots + c_n\, x_n = w^t u$ in a linear functional on V. In fact, $f \rightarrow (x)$ is abstained as an inner product of coadunate vectors of z and x. Restricting Z to be the basis vector $x_i s$, we get Z number of linear functional $f_n(i = 1, 2, \rightarrow, n)$ in V given by $f_{xi}(n) = x_i$ because now the eliminating vector $e_i = [0, -1, \ldots o]^T$ is the co ordinate vector of x_i with respect to the basis vectors $\{x_1, x_2, \ldots, x_n\}$ and should replaes W. This functional may be reconcile as a linear transformation as V that maps each vector in V onto its i-th co ordinates. Sulative to the respected basis. A note worthy property of thus function is, $f_{xi}(x_j) = f_{ij}, i, j = 1, \ldots n\, x_i$ when x_{ij} in the kpronceluers delta defined by $\begin{matrix} x_{ij} = 0, i \neq j \\ = 1, i = j \end{matrix}$ because e_j is the co ordinate vector of x_j with uspect to the basis $\{x_1, x_2, \ldots, x_n\}$ and $u = e_j$ *and* $w = e_i$. Mover if x in a zero vecor in V, then $f_0(x) = 0$ the sealer zero in $\mathbb{F}$.

Now we shall state and prove a very email they are on dual basic.

Theaum :

Let $\{x_1, x_2, \ldots, x_n\}$ be a basis of the n-delusional vector space V one a field $\mathbb{F}$. Then the unique linear functional f_{x1}, f_{x2}, f_{xn} defined by $f_{xi}(x_j) = f_{ij}, i, j = 1, 2, \ldots, n$ farm a basis of the dual space V* defined as the dual basis of $\{x_1, x_2, \ldots, x_n\}$ and any element f in V* can be expressed as $f = f(x_1) f_{x1} + \ldots + f(x_n) f_{xn}$ and for each vector X in V, we have $x = f_{x},(x)\, x_1 + f_{x2}(x)\, x_2 + \ldots + f_{x4}(x)\, x_n$.

Proof :

We have to prove that (a) f_{xi}'s are linear by independent in V*.

(i) Any vector f in v* can be expressed as a liner combination of f_{xi}'s.

Suppose that far realer $x_1, x_2, \ldots x_n$ *in* $\mathbb{F}$ we have $x_1 f_{n1} + x_2 f_{x2} + \ldots + x_n f_{xn} = f_q$

Then

$(x_1 f_{x1} + x_2 f_{x2} + \ldots + x_n f_{xn})\, n_i = f_0(x_i)$ which, by taking into account that $t_{ni}(n_j) = f_{ij}, i, j = 1, 2, \ldots, n$ *and* $f_0(x) = 0$ we get.

$x_i = 0, i = 1, 2, \ldots n.$ Hence, f_{ni}'s are linearly independent in V*.

Suppose now that f is any linear functional in V*. Than we can find n sealers $a_1, a_2, --, a_n$ *in* $\mathbb{F}$ satisfying $f(x\,i) = a\,i,\ i = 1, 2, \ldots, n$ because f n pursuable a known liner functioned. For any vector $x = x_1 n_1 + x_2 x_2 + \ldots + x_n x_n$ in V, because of the inanity of f, we get $f(x) = x_1 f(x_1) + x_2 + (x_2) + \ldots + x_n f(x_n)$

$$= x_1\ a_1 + x_2\ a_2 + ... + x_n\ a_n$$
$$= a_1\ f_{n1}(n) + a_2\ f_{n2}(x) + ... + \alpha_n\ f_{xu}\ (x)$$
$$= (a_1\ f_n, + a_2\ f_{n2} + .. + a_n\ f_{nu})\ (n).$$

Since x is arbitrary and a combination of linear functional is again a linear functional, we have the rulation.

$$f = a_1\ f_{x1} + a_2\ f_{n2} + ... + a_n\ f_{nn}\ .$$

Definition : Dual Basis :

For any basis $\{x_1, x_2, ... x_n\}$ of a vector space V one a field $\mathbb{F}$, thus exists a unique basis $\{f_1, f_2, ..., f_n\}$ of V* such that $f_i(x_j) = f_{ij}, i, j = 1,2,.., n$ of the basis $f_1, f_2, ..., f_n\}$ of V* is said to be dual to the given basis $x_1, x_2, ... x_n\}$ *of V.*

Suppose :

$B = \{(3, 2), (1,1)\}$ be a basis in $V = \mathbb{R}^2$. We will find a dual basis $f_{x1}, f_{x2}\}$ *in* V^* *when* $B = \{x_1, x_2\}$. Let, $x = x_1\ x_1 + x_2\ x_2$ *i.e.*

$$i.e. \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = [x_1\ x_2] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$
$$= \begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Which gives $x_1 = x_1 - x_2\ , x_2 = -2\ x_1 + 3\ x_2$

i. e. $f_{n1}(x) = x_1 - x_2, f_{n2}(x) = -2n_1 + 3x_2$

Example :

$\therefore$ *Let* $B = \{n_1 = 1, n_2 = 1 + t, x_3 = t + t^2\}$ in a basis in $P_2(t)$ one $\mathbb{R}$.

$$x(t) = x_1\ x_1 + x_2\ x_2 + x_3\ x_3$$
$$= x_1\ (1) + x_2\ (1 + t) + x_3\ (t + t^2)$$
$$= (x_1 + x_2) + (x_2\ x_3)\ t + x_3 + 2$$
$$\therefore x(t) = a_o + a_1\ t + a_2\ t^2$$
$$\Rightarrow \alpha_0 = x_1 + x_2, a_1 = x_2 + x_3\ , a_2 = x_3$$
$$\Rightarrow x_1 = a_0 - a_1 + a_2\ , x_2 = a_1 - a_2, x_3 = a_2$$

$\therefore$ me obtain the dual basis as,

$$f_n, (n) = a_0 - a_1 + a_2\ , f_{n2}(x) = a_1\ a_2\ \ f_3\ (n) = a_2.$$

Answer

Exercise : 3.1

3. (i) $\left(\frac{2}{3},1\right)($ $-1,-1)$

(ii) $\left(-\frac{2}{3},1\right),\left(\frac{11}{3},-3\right)$

(iii) $1,0,1)$ $($ $4,2)$

4. $\{(1,0,1),(2,1,1)\}$ in a basis of Inc(T) din(Im T) = 2.

$\{(3,-1,1)\}$ is a basis of kur T and dim (ku T) = 1

❖❖❖❖

4

DETERMINANT

Unit Structure :

4.0 INTRODUCTION

In previous three chapters we have discarded about vcetans line on equations and lincer transformations. Every when we see that we need to check wheather the vectors are linearly independent or not. In this chapter we mill develop a computational technique to find that by using determinants.

4.1 OBJECTIVE

- This chapter will help you to know about determinants and its properties.
- Expiring of determinants by various methods.
- Calculation of rank of a matrix using determinants.
- Existence and uniqueness of a system of equations.

4.2 DETERMINANT AS N-FARM

To discuses determinants, we always consider a square meters

Determinant of a square meters in a value associated with the matrix.

Definition :

Define a determinant function as dut : $M(n,\mathbb{R}) \to \mathbb{R}$ where $M(n,\mathbb{R})$ is a collection of square meters of order n, such that

(i) the value of determinant remain same by adding any multiple of 1st now to i th now i.c.

Det $(R_1, R_2, --, R_i + K\ R_j, R_i + 1, R_n$

$= \det(R_1, R_2, .., R_j, .., R_n)$

$fan\ i \neq j$

(ii) the value of the determinant changes by sign by swapping any two rows i.e. det $(R_1, R_2, .., R_j, .., R_j, .., R_n)$

$= -\det(R_1, R_2, .. R_j, .., R_i, .., R_n)$

(iii) if element of any now in multiplied by k then am value of determinant in is k tines its original value i.c.

Det $(R_1, R_2, .., \kappa R_i, .., R_n)$

$= k \det(R_1, R_2, .., R_i, .., R_n)$

$far\ k \neq 0$

(iv) det (I) = 1, where In an identify matrix.

If is n-linear skin symmetries function on $\mathbb{R}^n\ x\, \mathbb{R}^n\ x .. x\, \mathbb{R}^n \to \mathbb{R}.$

$\therefore A \in M\ (n, \mathbb{R})$

$\Rightarrow \det A = \det(R_1, R_2, .., R_n)$

$or \det A = \det(C_1, C_2, .., C_n)$

Which R_i denote i-th row or C_i denote i-th column of meters A.

Each $R_i\ or\ C_i \in \mathbb{R}^n$

$\therefore A \in \mathbb{R}^n\ x\, \mathbb{R}^n\ x - x\, \mathbb{R}^n$

$\therefore$ Determinant in an n-linear skew symmetric function from $\mathbb{R}^n\ x\, \mathbb{R}^n\ x .. x\, \mathbb{R}^n\ to\ \mathbb{R}$

e.g. the function $\mathbb{R}^2\ x\, \mathbb{R}^2 - \mathbb{R}$ in give by $\left(\begin{pmatrix} a \\ c \end{pmatrix}, \begin{pmatrix} b \\ d \end{pmatrix}\right) \to ad - be$

there $\begin{pmatrix} a \\ c \end{pmatrix} \& \begin{pmatrix} b \\ d \end{pmatrix}$ are ordered pairs from $\mathbb{R}^2$ each.

Also $\begin{bmatrix} a\ b \\ c\ d \end{bmatrix} = ab - leilinear,$

Skeen symmetric and alternating function.

4.3 EXPANSION OF DETERMINANT

Let $A = (a_{ij})$ be an arbitrary nxn matrix as follows

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{1j} & a_{in} \\ a_{i1} & a_{i2} & a_{ij} & a_{in} \\ a_{n1} & a_{n2} & a_{nj} & a_{nn} \end{pmatrix}$$

Let A_{ij} be the (n-1) x (n-2) matrix obtained by deleting the i-th row and j-th column from A.

$$A_{ij} \begin{pmatrix} a_{11} & a_{12} & a_{ij+1} & a_{1j+1} & a_{1n} \\ a_{-11} & a_{i-12} & a_{i-1j-1} & a_{i-1+1} & a_{i-1n} \\ a_{i+1} & a_{i+12} & a_{i+1j-1} & a_{i+1j+1} & a_{i+1n} \\ a_{n1} & a_{n2} & a_{nj-1} & a_{nj+1} & a_{nn} \end{pmatrix}$$

We will give an expressian for the determinant of an nxn matrix in terms of determinants of (n-1) x (n-1) matrix. We define,

$$\det A = (-1)^{i+2} a_{i1} \det(a_{i1}) + --- + (-1)^{i+n} a_{in} \det(A_{in})$$

This sum is called the expansion of the determinant according to the i-th expand det A according to the first now,

$$\det A = a_{11} \det(A_{11}) - a_{12} \det(A_{12})$$

$$+ a_{13} \det(A_{13}) - a_{14} \det(A_{14})$$

Where A in an 4 x 4 matrix.

For example det

$$A = \begin{pmatrix} 3 & 2 & -1 \\ 5 & 0 & 2 \\ 1 & 3 & -4 \end{pmatrix}$$

Then

$$\det A = 3\begin{vmatrix} 0 & 2 \\ 3 & -4 \end{vmatrix} - 2\begin{vmatrix} 5 & 2 \\ 1 & -4 \end{vmatrix} + (-1)\begin{vmatrix} 5 & 0 \\ 1 & 3 \end{vmatrix}$$

= 3(-6) -2 (-20-2) -1 (15)

= -18 + 44 – 15 = 44 – 33 = 11

4.4 SOME PROPERTIES

The determinant satisfies the following properties :

1. As a function of each column vector, the determinant in linear, i.e. if the j-th column c_j is equal to a sum of two, column vectors, say $c_j = c + c$ then

$$D(c_1,..,c_{j-1}\ ,\ c+c^1,c_{j+1},..c_n)$$
$$=D\,(c_1,..,\chi_{j-1},c,..,c_n)$$
$$+D\,(c_1,..,c_{j-1},c^1,..,c_n)$$

2. If two columns are equal i.e. if $c_j = c_k \ where\ j \neq k,$ hun det (A) = 0

3. If one adds a sealer multiple of one column to another three the value of the determinant does not change i.e.
$$D(c_1,..c_k + x\,c_i\ ,...,c_n) = D(c1,..,c_n)$$

We can prove this profile as follows.
$$D(c_1,..,c_{k+x}\ c_i,..c_n)$$
$$=D(c_1,..c_j,..c_k,..,c_n)$$
$$+x\,d\,(c_1,..c_j,..,c_j\ ,..,c_n)$$
$$=D\,(c_1,..c_j\ ,..,c_n\ ,..c_n)$$

As the second determinant on right has same columns and hence its value is zero.

All the properties stated above are valid for both row and column operations.

Using the above properties we can computer the determinants very efficiently. By using the property if a sealer multiple of are now (or column) in added to another now then the value of the determinant does not change, we will by to make as many enemies in the meters square to 0 and then expand.

Example :

$$D = \begin{vmatrix} 2 & 1 & 2 \\ 0 & 3 & -1 \\ 4 & 1 & 1 \end{vmatrix}$$

First we will interchange first two columns to keep first entry as 1.

$$\therefore D = -\begin{vmatrix} 1 & 2 & 2 \\ 3 & 0 & -1 \\ 1 & 4 & 1 \end{vmatrix}$$

Next we will make first two entries of 2nd and 3rd columns as zero. So we subtract trick the first now from 2nd and 3rd columns.

$$D = -\begin{vmatrix} 1 & 0 & 0 \\ 3 & -6 & -7 \\ 1 & 2 & -1 \end{vmatrix}$$

So if nc expand it along the first now nc get only a 2 x 2 determinant as.

$$D = -\begin{vmatrix} -6 & -7 \\ 2 & -1 \end{vmatrix} = -(6+14) = -20$$

Exercises 4.1

1. Computer the following determinates.

$$(i)\begin{vmatrix} 3 & 0 & 1 \\ 1 & 2 & 5 \\ -1 & 4 & 2 \end{vmatrix} (ii)\begin{vmatrix} 2 & 0 & 4 \\ 1 & 3 & 5 \\ 10 & -1 & 0 \end{vmatrix}$$

$$(iii)\begin{vmatrix} 3 & 1 & 2 \\ 4 & 5 & 1 \\ -1 & 2 & -3 \end{vmatrix} (iv)\begin{vmatrix} 2 & 4 & 3 \\ -1 & 3 & 0 \\ 0 & 2 & 1 \end{vmatrix}$$

2. Computer the following determinants.

$$(i)\begin{vmatrix} 1 & 1 & -2 & 4 \\ 0 & 1 & 1 & 3 \\ 2 & -1 & 1 & 0 \\ 3 & 1 & 2 & 5 \end{vmatrix} (ii)\begin{vmatrix} -1 & 1 & 2 & 0 \\ 0 & 3 & 2 & 1 \\ 0 & 4 & 1 & 2 \\ 3 & 1 & 5 & 7 \end{vmatrix}$$

$$(iii)\begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 2 \end{vmatrix} (iv)\begin{vmatrix} 1 & 1 & 1 & 1 \\ 2 & -2 & -1 & 3 \\ 4 & 4 & 1 & 9 \\ 8 & -8 & -1 & 27 \end{vmatrix}$$

4.5 SOME BASIC RESULTS

(1) If k is constant and A is n x n matrix thus $|KA| = k^n |A|$ where KA is obtained by multiplying each element of A by K.

Proof :

Each term of the determinants $|KA|$ has a factor k. This k can be taken out column from each now i.e.
$|KA| = k\ x\ k\ x .. x\ k\ (n - thus)|A|$

$\therefore |KA| = k^n |A|$

(2) If A and B are two n x n matrix $|AB| = |A||B|$ But $|A + B| \neq |A| + |B|$

This can be proved by simple examples.

(3) $|A| = \frac{1}{|A^{-1}|}$

Proof :

$\therefore |AB| = |A||B|, considering$
$B = A^{-1}\ we\ get,$
$|AA^1| = |A||A^{-1}|$
$|I| = |A||A^{-1}|$
$1 = |A||A^{-1}|$
$\Rightarrow |A| = \frac{1}{|A^{-1}|}$

(4) $|A^t| = |A|$ *where* A^t is transfer of matrix A.

Proof :

The determinant can be expanded by any now or column. Let $|A|$ is expanded using i-th now. This i-th now in i-th column in $|A^t|$. So the expansion remain same i.e. the value of the determinant remain same.
$\therefore |A^t = |A||$

(5) If A has a now (or column) of zeros then $|A| = 0$.

Proof :

By expanding along the new now, the value of the determinant becomes zero.

(6) If A and B are square meters then $\left|BAB^{-1} = |A|\right|$

Proof :

$$\left|BAB^{-1}\right| = |B||A|\left|B^{-1} = |B||A|\right|\frac{1}{|B|}$$

$$= |A|$$

(7) The determinants is linear in each now (column) if the other rows (column) are fined.

$$\begin{vmatrix} a_1 + k & a_2 + k & a_3 + k \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}$$

$$= \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} + \begin{vmatrix} k & k & k \\ b_2 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}$$

$$\begin{vmatrix} a_1 & a_2 & a \\ kb_1 & kb_2 & kb_3 \\ c_1 & e_2 & e_3 \end{vmatrix} = k \begin{vmatrix} a_1 & a_2 & a_3 \\ b_2 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}$$

4.6 LAPLAER EXPANSION

Definition :

Mainor – The minor of an element of a square matrix in the determinant obtained by deleting the now and column which interest in that element.

So minor of the element a_{ij} is obtained by deleting i-th now and j-th column, deleted by M_{ij}.

For example, let

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

Then,

Minor of 1 is $M_{11} = \begin{vmatrix} 5 & 6 \\ 8 & 9 \end{vmatrix} = -3$

Minor of 8 is $M_{32} = \begin{vmatrix} 1 & 3 \\ 4 & 6 \end{vmatrix} = -6\,etc.$

Laplaer expansion –

$$|A| = \sum_{J=1}^{n} (-1)^{i+j}\, a_{ij}\, M_{ij}$$

If wc consider i-th previous example.

$$M_{11} = \begin{vmatrix} 5 & 6 \\ 8 & 9 \end{vmatrix} = -3, M_{12} = \begin{vmatrix} 4 & 6 \\ 7 & 9 \end{vmatrix} = -6$$

$$M_{13} = \begin{vmatrix} 4 & 5 \\ 7 & 8 \end{vmatrix} = 7, M_{21} = \begin{vmatrix} 2 & 3 \\ 8 & 9 \end{vmatrix} = -6$$

$$M_{22} = \begin{vmatrix} 1 & 3 \\ 7 & 9 \end{vmatrix} = -12, M_{23} = \begin{vmatrix} 1 & 2 \\ 7 & 8 \end{vmatrix} = -6$$

$$M_{31} = \begin{vmatrix} 2 & 3 \\ 5 & 6 \end{vmatrix} = -3, M_{32} = \begin{vmatrix} 1 & 3 \\ 4 & 6 \end{vmatrix} = -6$$

$$M_{33} = \begin{vmatrix} 1 & 2 \\ 4 & 5 \end{vmatrix} = -3$$

$\therefore$ BY Laplaer Expansion

$$\begin{aligned} |A| &= a_{11}\, M_1 - a_{12}\, M_{12} + a_{13}\, M_{13} \\ &\quad - \alpha_{21}\, M_{21} + a_{22}\, M_{22} - a_{23}\, M_{23} \\ &\quad + a_{31}\, M_1 - a_2\, M_{32} + a_{33}\, M_{33} \\ &= 1\, x\,(-3) - 2\, x\,(-6) + 3\, x\, 7 - 4\, x\,(6) \\ &\quad + 5\, x\,(-12) - 6\, x\,(-6) + 7\,(-3) - 8\,(-6) \\ &\quad + a\, x\,(-3) \\ &= -3 + 12 + 21 + 24 - 60 + 36 - 21 + 48 - 27 \\ &= 30 \end{aligned}$$

4.7 THE RANK OF A MATRIX

Theorem :

Let $c_1, c_2, \ldots, c_n$ be column vectors of disunion n. They are linear by dependent if and only it.
$\det(c_1, c_2, \ldots c_n) = 0$

Proof :

Let $c_1, c_2, \ldots, c_n$ are linear by dependent. So there exists a solution.
$x_1\, c_1 + x_2\, e_2 + \ldots + x_n\, c_n = 0$ with numbers $x_1, \ldots, x_n$ not all 0.
Let $x_j \neq 0$

$$\therefore x_j\, c_j = -\, x_1\, c_1 \ldots -\, x_{j-1}\, c_{j-1} \ldots -\, x_n\, c_n$$

$$i.p.\, c_j = -\frac{x_1}{x_j} c_1 \ldots. - \frac{x_n}{x_j} c_n$$

$$= \sum_{\substack{k=1 \\ k \neq j}}^{n} a_k c_k$$

Thus

$$\text{Det } (A) = \det(c_1, \ldots, c_i, \ldots, e_n)$$

$$= \det\left(c_1, \ldots, \sum_{k=1}^{n} a_k\, c_k, \ldots c_n\right)$$

$$= \sum_{\substack{k=1 \\ k \neq j}}^{n} \det(c_1, \ldots, a_k\, c_k, \ldots, c_n)$$

Where c_k occurs in the j-th place. But c_k also occurs in the k-th place and $k \neq j$. Hence two column of the determinant in equal. So the value of the determinant in 0.

Conversely :

If all the columns of a matrix A in linear by independent, then the matrix A is now equivalent to a triangular matrix B as

$$B = \begin{pmatrix} b_{11} & b_{12} & b_{in} \\ 0 & b_{22} & b_{zn} \\ 0 & 0 & b_{nn} \end{pmatrix}$$

When all the diagonal elements $b_{11}, b_{22}, \dots, b_{nn} \neq 0.$

But by the rule of expiations. Det $(B) = b_{11}\, b_{22} \dots, b_{nn} \neq 0.$

Now, B is obtained by some operation like multiplying a row by non zero scalar.

Which multiplies the determinant by this scalar; or interchanging rows, which multiplies the determinant by -1 , or adding a multiple of are now to another, which does not charge the value of the determinant since $\det(B) \neq 0$ it follows that det $(A) \neq 0.$

Hence the proof.

Corollary :

If $c_1, c_2, - c_n$ are column vectors of $\mathbb{R}^n$ such that $D(c_1, c_2, \dots c_n) \neq 0,$ and if B in a column vector, then thus exist number $x_1, \dots, x_n$ such that $x_1\, c_1 + \dots + x_n\, c_n = B.$

These numbers are uniquely determined by B.

Proof :

$\therefore D(e_1, c_2, \dots, c_n) \neq 0$

$c_1, c_2, \dots c_n$ are linear by independent and hens from a basis of $\mathbb{R}^n$. So, $BE\mathbb{R}^n$ can be written as a linear combination of $c_1, c_2, \dots, c_n$ for some unique numbers $x_1, x_2, \dots, x_n$.

$\therefore x_1\, c_1 + x_2\, c_2 + \dots + x_n\, c_n = B$ for unique $x_1, x_2, \dots, x_n$.

The above corollary can be placed an important feature of system of linear equations like :

If a system of n linear equations in n unknowns has a matrix of coefficients whose determinant in not zero, then the system has a unique solution.

Now recall that the rank of a matrix in dimension of row space on column space i.e. number of independent now of column vectors in the matix.

Let a 3 x 4 matrix is given and we have to find out its rank. Its rank is at most 3. If we can show at last one 3 x 3 determinant from the co-efficient matrix is non-new, we consider the rank of the matrix in 3. If all 3 x 3 determinants are new, we have to check 2 x 2 determinants. If at last one of then is non-zero, we can conclude the rank of the matrix in 2.

For example let

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

$$\begin{vmatrix} 3 & 5 & 1 \\ 2 & -1 & 1 \\ 5 & 4 & 2 \end{vmatrix} = 0 , \begin{vmatrix} 5 & 1 & 4 \\ -1 & 1 & 1 \\ 4 & 2 & 5 \end{vmatrix} = 0$$

One can check that any such 3 x 3 determinant from A in zero.

So, rank $(A) \neq 3$.

Now $\begin{vmatrix} 3 & 5 \\ 2 & -1 \end{vmatrix} = -13 \neq 0.$

No need to check other 2 x 2 determinant. We can conclude that rank (A) = 2.

Exercise 4.2

1. Find the rank of the following matrix.

$$i) \begin{pmatrix} 3 & 1 & 2 & 5 \\ 1 & 2 & -1 & 2 \\ 1 & 1 & 0 & 1 \end{pmatrix} (\quad ii) \begin{pmatrix} 1 & 1 & -1 & 2 \\ 2 & -2 & 0 & 2 \\ 2 & -8 & 3 & -1 \end{pmatrix}$$

$$iii) \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 2 \end{pmatrix} (\quad iv) \begin{pmatrix} 3 & 1 & 1 & -1 \\ -2 & 4 & 3 & 2 \\ -1 & 9 & 7 & 3 \\ 7 & 4 & 2 & 1 \end{pmatrix}$$

2. Find the values of the determinants using Laplaec expansion.

$$i) \begin{pmatrix} 2 & 4 & 6 \\ 7 & 8 & 3 \\ 5 & 9 & 2 \end{pmatrix} (\quad ii) \begin{pmatrix} 2 & 0 & 3 \\ 5 & 1 & 0 \\ 0 & 4 & 6 \end{pmatrix}$$

$$iii)\begin{pmatrix} 1 & -2 & -3 \\ -2 & 3 & 4 \\ 3 & -4 & 5 \end{pmatrix} \quad iv)\begin{pmatrix} 3 & 2 & 0 \\ 1 & 1 & 3 \\ 2 & -1 & 2 \end{pmatrix}$$

3. Check only uniqueness of the solution for the following systems of equations.

i) $2_n - y + 3_z = 9$

$n + 3_y - z = 4$

$3_n + 2_y + z = 10$

ii) $2_n + y - z = 5$

$x - y + 2_z = 3$

$-x + 2_y + z = 1$

iii) $4_n + y + z + w = 1$

$n - y + 2_z - 3_w = 0$

$2_n + y + 3_z + 5_w = 0$

$n + y - z - w = 2$

iv) $x + 2_y - 3_z + 5_w = 0$

$2_n + y - 4_z - w = 1$

$n + x + z + w \quad = 0$

$-n - y - z + w \quad = 4$

[Hint : Just check that determinant of the co-efficient matrix is non few far the uniqueness]

4.8 GRAMER'S RULE

Determinants can be used to solve a system of linear equations.

Thearim :

Let $c_1, c_2, \dots, c_n$ be column vectors such that $D(c_1, e_2, \dots, e_n) \neq 0$. Let B be a column vector and $x_1, x_2, \dots, x_n$ are numbers such that $x_1\, c_1 + x_2\, c_2 + \dots + x_n\, c_n = B,$ then for each J = 1, 2, … , n.

$$nj = \frac{D(c_1, e_2, B, \quad , e_n)}{D(e_1, e_2, \quad , e_n)}$$

Where B column vector replaces the column c_j in numerator of x_j.

Proof :

$$D(c_1, c_2, \dots, B, \dots c_n)$$
$$= D(c_1, c_2, \dots, x_1 c_1 + n_2 c_2 + \dots x_n c_n),$$
$$= D(c_1, c_2, \dots x_1 c_1, \dots, c_n)$$
$$+ D(c_1, c_2, \dots, x_2 c_2, \dots, c_n)$$
$$+ D(c_1, c_2 \dots, x_j c_j, \dots c_n)$$
$$+ d(c_1, c_2 \dots, x_n c_n, \dots c_n)$$
$$= x_1 D(c_1, c_2, \dots, c_1, \dots, c_n)$$
$$+ \dots\dots\dots$$
$$+ x_j D(c_1, c_2, \dots c_j, \dots c_n)$$
$$+ \dots\dots\dots\dots$$
$$+ x_n D(c_1, c_2, \dots., c_n, \dots, c_n)$$

In every term of this sam except the j-th term, two column vectors are equal. Hence every term except the j-th term is equal to 0. so we get.

$$D(c_1, c_2, \dots, B, \dots c_n)$$
$$= x_j D(c_1, c_2, \dots, c_j, \dots c_n)$$
$$\therefore x_j = \frac{D(c_1, c_2, .., B, \dots c_n)}{D(c_1, c_2, \dots, c_j, \dots c_n)}$$

So we can salve a system of equations using above rule. This rule in known as Gramers rule.

Example :

$3_x + 2_y + 4_z = 1$

$2_n - y + z \quad = 0$

$x + 2_y + 3_z \quad = 1$

By Gramer's rule

$$x = \frac{\begin{matrix} 1 & 2 & 4 \\ 0 & -1 & 1 \\ 1 & 2 & 3 \end{matrix}}{\begin{matrix} 3 & 2 & 4 \\ 2 & -1 & 1 \\ 1 & 2 & 3 \end{matrix}}$$

$$y = \frac{\begin{matrix} 3 & 1 & 4 \\ 2 & 0 & 1 \\ 1 & 1 & 3 \end{matrix}}{\begin{matrix} 3 & 2 & 4 \\ 2 & -1 & 4 \\ 1 & 2 & 3 \end{matrix}}$$

$$z = \frac{\begin{matrix} 3 & 2 & 1 \\ 2 & -1 & 0 \\ 1 & 2 & 1 \end{matrix}}{\begin{matrix} 3 & 2 & 4 \\ 2 & -1 & 1 \\ 1 & 2 & 3 \end{matrix}}$$

$$x = -\frac{1}{5}, \; y = 0, \; z = \frac{2}{5}$$

Exercise 4.3

1. Solne the following equation by Gramer's rule.

i) $x + y + z = 6$
$x - y + z = 2$
$x + 2_y - z = 2$

ii) $x + y - 2_z = -10$
$2_n - y + 3_z = -1$
$4_x + 6_y + z = 2$

iii) $-2_x - y - 3_z = 3$
$z_x - 3_y + z = -13$
$2_x - 3_z = -11$

iv) $4_x - y + 3_z = 2$
$x + 5_y - 2_z = 3$
$3_x + 2_y + 4_z = 6$

Answer

Exercise 4.1

1. (i) -42 (ii) -114 (iii) 14 (iv) -9

2. (i) -18 (ii) -45 (iii) 4 (iv) 192

Exercise 4.2

1. (i) 3 (ii) 2 (iii) 4 (iv) 2

2. (i) -204 (ii) 72 (iii) -6 (iv) 23

3. (i) unique (ii) unique (iii) unique (iv) unique.

Exercise 4.3

1. (i) (1,2,3) (ii) (5,3,4)
   (iii) (-4, 2, 1) (iv) (0,1,1)

❖❖❖❖

Chapter 6

Characteristic Polynomial

Chapter Structure

6.1 Introduction

In the following chapter we will attempt to understand the concept of characteristic polynomial. In earlier two units of this course we have seen that studying matrices and studying linear transformations on a finite dimensional vector space is one and the same thing. Although this a case it is important to note that we get different matrices if we change the basis of a underlying vector space. We also saw that although we get different matrices for different bases, corresponding matrices are similar. Being similar is an equivalence relation on the space of $n \times n$ matrices, it is interesting to seek bases of underlying vector space in which for a given linear transformation corresponding matrix is simplest in appearance and because of similarity being equivalence relation we do not loose anything important as far as linear transformation is concerned. Studying so called eigenvalues of a linear transformation addresses the issue of such bases in which given linear transformation has a matrix in simplest form. During the quest of finding such bases mentioned above we come to know various beautiful properties of linear transformation and its relation to linear structure on the vector space.

6.2 Objectives

After going through this chapter you will be able to:
• find eigenvalues of a given linear transformation
• find basis of a vector space in which matrix of a linear transformation has diagonal or at least triangular form
• Properties of eigenvalues that characterize properties of linear transformation

Let V be a finite dimensional vector space over a field of scalars F. Let T be a linear transformation on V.

Definition 16. Eigenvalue (also known as characteristic value) of a linear transformation T is a scalar α in F such that there exist a nonzero vector $v \in V$ with $T(v) = \alpha v$. any such v is known as eigenvector corresponding to eigenvalue α. Collection of all such $v \in V$ for a particular eigenvalue is a subspace of V known as eigen space or characteristic space associated with α.

Theorem 6.2.1. *Let T be a linear transformation on a finite dimensional space V. Then α is characteristic value of T if and only if the operator $T - \alpha I$ is singular.*

Proof. this is the proof of the theorem. □

Remark 6.2.1. A linear transformation is singular if determinant of its matrix is zero.
Thus α is eigenvalue of T if and only if determinant of matrix of $T - \alpha I$ is zero. We see this determinant is a polynomial in α and hence roots of the polynomial $det(T - xI)$ are eigenvalues of T.

Definition 17. **det**$(T - xI)$ is known as characteristic polynomial of T.

Example 7. Find eigenvalues and eigenvectors of the following matrix

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

Solution:

Step 1: Consider $\det(A - \lambda I) = 0$

$$\mathbf{det}A = \mathbf{det}\begin{pmatrix} 1-\lambda & 0 & 1 \\ 2 & 3-\lambda & 1 \\ 1 & 1 & 1-\lambda \end{pmatrix} = 0$$

This gives characteristic polynomial of A and its roots are eigenvalues of A. Eigenvector is a solution vector of homogeneous system of linear equations $(A-\lambda I)x = 0$ where λ is an eigenvalue of A.
Thus characteristic polynomial of A is obtained by finding determinant in above equation and it is found to be the following $p(\lambda)$

$$p(\lambda) = \lambda^3 - 5\lambda^2 + 5\lambda - 1$$

Step 2: Roots of $p(\lambda)$ are $2+\sqrt(3), 1$ and $2-\sqrt{3}$. These are eigenvalues of A.

Step 3: Solve following system of linear equations and we consider first eigenvalue $2+\sqrt{3}$

$$(A-(2+\sqrt{3})I)x = 0$$

Solving we get $X_1 = \left(-\frac{3}{2} + \frac{1}{2}(2+\sqrt{3}), -\frac{1}{2} + \frac{1}{2}(2+\sqrt{3}, 1 \right)$

Step 4: similarly for other two eigenvalues we get following eigenvectors

$$X_3 = \left(\frac{1}{2} + \frac{1}{2}(2+\sqrt{3}), -\frac{1}{2} + \frac{1}{2}(2+\sqrt{3}, 1 \right)$$
$$X_2 = \left(-1, 1, 0 \right)$$

Remark 6.2.2. Roots of characteristic polynomial may repeat and behavior of a linear transformation (or its corresponding matrix) depends crucially on multiplicity of eigenvalues and dimension of corresponding eigen spaces. One simple example of matrix with repeated eigenvalues is the following matrix

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 3 \end{pmatrix}$$

Definition 18. Algebraic multiplicity of an eigenvalue Multiplicity of an eigenvalue as a root of characteristic polynomial is known as algebraic multiplicity of corresponding eigenvalue.

Definition 19. Geometric multiplicity of an eigenvalue Dimension of the eigenspace or characteristic space of an eigenvalue is known as the geometric multiplicity of a that eigenvalue.

Theorem 6.2.2. *Let α be an eigenvalue of a linear transformation T. If $f(x)$ is a polynomial in indeterminate x. Then $f(T)v = f(\alpha)v$*

Proof. First consider the case of $f(x)$ being a monomial. Let $f(x) = x^k$. Let us apply induction on k.
Let $k = 1$. In this case $f(x) = x$. i.e. $f(T) = T$. Here it follows that $f(T)v = f(\alpha)v$.
Let us assume the lemma for $k = r$. Thus we assume that $T^r(v) = (\alpha^r)v$. Consider $k = r + 1$
. $T^{r+1}(v) = T^rTv = (T^r)\alpha v = \alpha(T^r)v$.
Using induction hypothesis we get that
$T^{r+1}(v) = \alpha^{r+1}v$
. Thus the lemma is established for all monomials.
Let $f(x) = a_0 + a_1x + ... + a_kx^k$. Thus $f(T)v = a_0v + a_1Tv + ... + a_kT^kv$. Using the lemma for monomials we get that $f(T)v = a_0 + a_1\alpha v + ... + a_k\alpha^kv = f(\alpha)v$. And the result is established for all polynomials.

□

Remark 6.2.3. This is very important lemma and in future we will use it at various occasions.

Theorem 6.2.3. *Let α_1 and α_2 be two distinct eigenvalues of a linear transformation T on a finite dimensional vector space V. Let v_1 and v_2 be respective eigenvectors. Then v_1 and v_2 are linearly independent.*

Proof. Suppose otherwise that v_1 and v_2 are linearly dependent. Then there exist a nonzero constant c such that $v_2 = cv_1$.Therefore $T(v_2) = cT(v_1)$. Thus we get that $\alpha_2v_2 = \alpha_1cv_1$. Which is same as $v_2 = \frac{\alpha_1c}{\alpha_2}v_1$. This leads to the conclusion that $\frac{\alpha_1c}{\alpha_2} = c$. Since v_1 and v_2 are distinct (as α_1 and α_2 are distinct) we have $c \neq 1$ leading to $\frac{\alpha_1}{\alpha_2} = 1$. This is same as $\alpha_1 = \alpha_2$. This is a contradiction to α_1 and α_2 being distinct. Therefore v_1 and v_2 are linearly independent. □

Now recall that geometric multiplicity of an eigenvalue is a number of linearly independent eigenvectors corresponding to that eigenvalue. In other words geometric multiplicity is nothing but dimension of eigen space (ie.characteristic space) of an eigenvalue. Note however that direct sum of all eigenspaces of a linear operator need not exhaust entire vector space on which linear transformation is defined and now onwards our attempt will be to see what best next can be done in case we fail to recover vector space V from direct sum of eigen spaces corresponding to a particularly given linear transformation. Question we want to address is that in which circumstances does direct sum of eigen spaces exhaust entire vector space. We will see that these linear transformations are precisely the one which are diagonalizable. In the following section we will make these ideas precise.

Theorem 6.2.4. *Let T be a linear transformation on a finite dimensional vector space V. Let $\alpha_1, \alpha_2, ..., \alpha_k$ be the distinct eigenvalues of*

T and let W_i be the eigenspace corresponding to the eigenvalue α_i. B_i be an ordered basis for W_i. Let $W = W_1 + W_2 + ... + W_k$. Then $dimW = dimW_1 + dimW_2 + ... + dimW_k$. Also $B = (B_1, B_2, ...B_n)$ is a ordered basis for W.

proof Vectors in B_i for $1 \leq i \leq k$ are linearly independent eigenvectors of T corresponding to eigenvalue α_i. Also vectors in B_i are linearly independent to those in B_j for $i \neq j$.Because they are eigenvectors corresponding to different eigenvalues. Thus vectors in B are all linearly independent. Note that vectors in B span W. This is because that is how W is defined.

Theorem 6.2.5. *Set of all linear transformations on a finite dimensional vector space forms a vector space over the field F. Here naturally the binary operation is composition of functions. This new vector space is isomorphic to space of $n \times n$ matrices and hence has dimension n^2. Let us denote this space by $L(V, V)$.*

Let T be a linear transformation on finite dimensional vector space V.Thus $T \in L(V, V)$. Consider the first (n^2+1) powers of T in $L(V, V)$.:

$$I, T, T^2, ..., T^{n^2}$$

Since dimension of $L(V, V)$ is n^2 and above we have $n^2 + 1$ elements, these must be linearly dependent.ie. there exist $n^2 + 1$ scalars, not all zero, such that we have:

$$c_0 + c_1T + c_2T^2 + ... + c_{n^2}T^{n^2} = 0 \tag{6.1}$$

Or which is same thing as saying that T satisfies polynomial of degree n^2. We have now definition:
Definition Any polynomial $f(x)$ such that $f(T) = 0$ is known as annihilating polynomial of a linear transformation T.
Polynomial given in 1 is one such polynomial. Thus set of annihilating polynomials is nonempty and we can think of annihilating polynomial of least degree which is monic.

Definition Annihilating polynomial of least degree which is monic is known as minimal polynomial of a linear transformation T.

Example 8. Find minimal polynomial of the following matrix

$$A = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 5 \end{pmatrix}$$

Solution:

Step 1: Find characteristic polynomial of A.
Characteristic polynomial of A is the following:

$$p(\lambda) = (\lambda - 2)^3(\lambda - 5)$$

Step 2: By definition of minimal polynomial, minimal polynomial $m(\lambda)$ must divide characteristic polynomial hence it must be one of the following:

$$(\lambda - 2)^3(\lambda - 5)$$
$$(\lambda - 2)^2(\lambda - 5)$$
$$(\lambda - 2)(\lambda - 5)$$

Step 3: Note that minimal polynomial is a polynomial of least degree which is satisfied by the matrix A. Only polynomial amongst above polynomial is the second one hence minimal polynomial is $(\lambda - 2)^2(\lambda - 5)$

Remark 6.2.4. 1. Set of all annihilating polynomials of a linear transformation T is an ideal in $F[x]$. Since F is a field, this ideal is a principal ideal and monic generator of this ideal is nothing but minimal polynomial of T.

2. Since minimal polynomial is monic it is unique.

Theorem 6.2.6. *Let T be a linear transformation on a n-dimensional vector space V. The characteristic and minimal polynomial of T have the same roots except for multiplicities.*

Proof. Let p be the minimal polynomial for T. Let α be a scalar. We want to show that $p(\alpha) = 0$ if and only if α is an eiegnvalue of T.
First suppose that $p(\alpha) = 0$. Then by remainder theorem of polynomials,

$$p = (x - \alpha)q \tag{6.2}$$

where q is a polynomial. Since $dedq < degp$, the definition of minimal polynomial p tells us that $q(T) \neq 0$. Choose a vector v such that

$q(T)v \neq 0$. Let $q(T)v = w$. Then

$$\begin{aligned} 0 &= p(T)v \\ &= (T - \alpha I)q(T)v \\ &= (T - \alpha I)w \end{aligned}$$

and thus α is an eigenvalue.
Conversely, let α be an eigenvalue of T, say $T(w) = \alpha w$ with $w \neq 0$. Since p is a polynomial we have seen that

$$p(T)w = p(\alpha)w$$

Since $p(T) = 0$ and $w \neq 0$, we have that $p(\alpha) = 0$. Thus eigenvalue α is a root of minimal polynomial p.

□

Remark 6.2.5. 1. Since every root of the minimal polynomial is also a root of characteristic polynomial we see that minimal polynomial divides characteristic polynomial. This is famous Caley Hamilton theorem and it states that linear transformation T satisfies characteristic polynomial in the sense that if $f(x)$ is a characteristic polynomial then $f(T) = 0$.

2. Similar matrices have the same minimal polynomial

Check Your Progress

1. Let $A := \begin{pmatrix} 4 & 4 & 4 \\ -2 & -3 & -6 \\ 1 & 3 & 6 \end{pmatrix}$ Compute (a)the characteristic polynomial (b) the eigenvalues (c) All eigenvectors (d) Identify algebraic and geometric multiplicities of each of the eigenvalue.

2. Let A be the real 3×3 matrix. Find minimal polynomial of A

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

6.3 Digonalizable Linear Transformations

Definition 20. Let V be a vector space of dimension n,and T be a linear transformation on V. Let W be a subspace of V. We say that W is invariant under T if for each vector $v \in W$ the vector $T(v)$ is in W. i.e. if $T(W) \subseteq W$.

Definition 21. Let T be a linear transformation on a finite dimensional vector space V. We say that T is diagonalizable if there exist a basis of V consisting of all eigenvectors of T.

Remark 6.3.1. Matrix of digonalizable T in the basis of V consisting of eigenvectors of T is diagonal matrix with eigenvalues along the diagonal of the matrix.

Example 9. Find the basis in which following matrix is in diagonal form

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

Solution: Since characteristic polynomial of the matrix is $p(\lambda) = \lambda^3 - 7\lambda^2 + 9\lambda - 3$ we get following eigenvalues for A

$$3 + \sqrt{(6)}, 1 \text{ and } 3 - \sqrt{6}.$$

Since all eigenvalues are distinct given matrix is and in the basis formed by three independent eigenvectors matrix becomes diagonal.

$$\begin{aligned} X_1 &= \left(-\tfrac{5}{2} + \tfrac{1}{2}(3 + \sqrt{6}), \tfrac{3}{2} - \tfrac{1}{6}(3 + \sqrt{6}, 1 \right) \\ X_2 &= \left(-1, 1, 0 \right) \\ X_3 &= \left(-\tfrac{5}{2} + \tfrac{1}{2}(3 - \sqrt{6}), \tfrac{3}{2} - \tfrac{1}{6}3 - \sqrt{6}, 1 \right) \end{aligned}$$

Required diagonal matrix is the matrix whose diagonal is formed by three eigenvalues respectively.

Check Your Progress

Let A be a matrix over any field F. Let χ_A be the characteristic polynomial of A and $p(t) = t^4 + 1 \in F[t]$. State with reason whether the following are true or false

1. Let $\chi_A = p$, then$A$ is invertible
2. If $\chi_A = p$, then A is diagonalizable over F
3. If $p(B) = 0$ for some matrix B be 8×8 matrix, then p is the characteristic polynomial of B.

4. There is unique monic polynomial $q \in F[t]$ of degree 4 such that $q(A) = 0$

Theorem 6.3.1. *Let T be a diagonalizable linear transformation on a n dimensional space V. Let $\alpha_1, \alpha_2, ..., \alpha_k$ be the distinct eigenvalues of T. Let $d_1, d_2, ..., d_k$ be the respective multiplicities with which these eigenvalues are repeated. Then characteristic polynomial for T is*

$$f = (x - \alpha_1)^{d_1} + (x - \alpha_2)^{d_2} + ... + (x - \alpha_k)^{d_k}$$

$$d_1 + d_2 + ... + d_k = n$$

Proof. If T is diagonalizable then in the basis consisting of eigenvectors matrix of T is a diagonal matrix with all eigenvalues lying along the diagonal. We know that characteristic polynomial of a diagonal matrix is a product of linear factors of the form-

$$f = (x - \alpha_1)^{d_1} + (x - \alpha_2)^{d_2} + ... + (x - \alpha_k)^{d_k}$$

□

Check Your Progress

1. Let T be the linear operator in R^4 which is represented in the standard ordered basis by the following matrix

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ a & 0 & 0 & 0 \\ 0 & b & 0 & 0 \\ 0 & 0 & c & 0 \end{pmatrix}$$

   Under what conditions on $a, b, andc$ is T diagonalizable.

2. Let N be 2×2 complex matrix such that $N^2 = 0$¿ Prove that either $N = 0$ or N is similar over C to

$$\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

Lemma 6.3.1. *Let W be an invariant subspace for T. The characteristic polynomial for the restriction operator T_w divides the characteristic polynomial for T. The minimal polynomial for T_w divides the minimal polynomial for T.*

Proof. We have

$$A = \begin{bmatrix} B & C \\ 0 & D \end{bmatrix} \tag{6.3}$$

where $A = [T]_B$ and $B = [T_w]'_B$. Because of the block form of the matrix

$$det(A - xI) = det(B - xI)det(xI - D) \tag{6.4}$$

That proves statement of characteristic polynomials. Note that we have used notation I for identity matrices of three different sizes.
Note that kth power of the matrix A has the block form

$$A^k = \begin{bmatrix} B^k & C_k \\ 0 & D^k \end{bmatrix} \tag{6.5}$$

where C_k is some $r \times (n - r)$ matrix. Therefore, any polynomial which annihilates A also annihilates B (and D too). So the minimal polynomial for B divides the minimal polynomial for A □

6.4 Triangulable Linear Transformations

Definition 22. Triangulable Linear Transformation The linear transformation T is called triangulable if there is an ordered basis of V in which T is represented by a triangular matrix.

Lemma 6.4.1. *Let V be a finite dimensional vector space and let T be a linear transformation on V such that minimal polynomial for T is a product of linear factors*

$$p = (x - \alpha_1)^r_1 + ... + (x - \alpha_k)^r_k \tag{6.6}$$

where $\alpha_i \in F$.
Let W be a proper subspace of V which is invariant under T. Then there exist a vector $v \in V$ such that

1. *v is not in W;*

2. *$(T - \alpha I)v$ is in W, for some characteristic value α of the transformation T.*

Proof. Let u be any vector in V which is not in W Then there exist a polynomial g such that $g(T)u \in W$. Then g divides the minimal polynomial p for T. Since u is not in W, the polynomial g is not constant. Therefore,

$$g = (x - \alpha_1)^{l_1} + (x - \alpha_2)^{l_2} + ... + (x - \alpha_k)^{l_k}$$

where at least one of the integers l_i is positive. We choose j such that $l_j > 0$ then $(x - \alpha_j)$ divides g. Hence

$$g = (x - \alpha_j)h \tag{6.7}$$

By the definiion of g, the vector $v = h(T)u$ can not be in W. But,

$$(T - \alpha_j I)v = (T - \alpha_j I)h(T)u = g(T)u \tag{6.8}$$

is in W. □.

We obtain triangular matrix representation of a linear transformation by applying following procedure:

1. Apply above lemma to trivial subspace $W = 0$ to get vector v_1.

2. Once $v_1, v_2, v_3, ...v_{l-1}$ are determined form a subspace W spanned by these vectors and apply above lemma to this W to obtain v_l in the following way-
   Note that the subspace W spanned by $v_1, v_2, ..., v_{l-1}$ is invariant under T. Therefore by above lemma there exist a vector v_l in V which is not in W such that $(T - \alpha_l I)v_l$ is in W for certain eigenvalue α_l of T. This can be done because minimal polynomial of T is factored into linear factors and above lemma is applicable. We will illustrate this procedure with the help of example.

Theorem 6.4.1. *In the basis obtained by above procedure the matrix of T is triangular*

Proof. By above procedure we get ordered basis $\{v_1, v_2, ...v_n\}$ This basis is such that $T(v_j)$ lies in the space spanned by $v_1, v_2, ...v_j$ and we have following form-

$$T(v_j) = a_{1j}v_1 + a_{2j}v_2 + ... + a_{jj}v_j, 1 \leq j \leq n \tag{6.9}$$

In this type of representation we get that the matrix of T is triangular. □.

Check Your Progress

Let

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

Check whether above matrix is similar over a field of real numbers to a triangular matrix? If so find such a triangular matrix.

Theorem 6.4.2. Primary Decomposition Theorem
Let T be a linear transformation on a finite dimensional vector space V over the field F. Let p be a minimal polynomial for T,

$$p = p_1^{r_1}...p_k^{r_k} \tag{6.10}$$

where each p_i are distinct irreducible monic polynomials over F and the r_i are positive integers. Let W_i be the null space of $p_i(T)^{r_i}$, $i = 1, 2, ..., k$. Then

1. *$V = W_1 \oplus ... \oplus W_k$;*

2. *each W_i is invariant under T;*

3. *if T_i is the transformation induced on W_i by T, then the minimal polynomial for T_i is $p_i^{r_i}$.*

.

Proof. Before proceeding to a proof of above theorem we note that real point is in obtaining so called primary decomposition stated in the theorem explicitly for a given linear transformation. Thus we present a proof in the form of algorithm which for given T will produce primary decomposition of a given linear transformation T. Following steps describe the method to obtain primary decomposition theorem

1. For given T, obtain minimal polynomial for T. Let it be in the following form

$$p = p_1^{r_1}...p_k^{r_k} \tag{6.11}$$

2. For each i, let

$$f_i = \frac{p}{p_i^{r_i}} = \Pi_{j \neq i} p_j^{r_j} \tag{6.12}$$

Note that all f_i are distinct and are relatively prime.

3. Find polynomials g_i such that

$$\sum_{i=1}^{k} f_i g_i = 1 \tag{6.13}$$

4. Let $E_i = h_i(T) = f_i(T)g_i(T)$

5. For $i \neq j$ we have

$$E_1 + ... + E_k = I \tag{6.14}$$
$$E_i E_j = 0, i \neq j \tag{6.15}$$

6. These E_i serve the purpose of obtaining invariant subspaces W_i which decompose V into direct sum, and each E_i is a projection operator.

7. It can be verified that minimal polynomial of T_i which is a restriction of T to W_i is $p_i^{r_i}$.

6.5 Nilpotent Linear Transformations

Definition 23. Nilpotent Transformation Let N be a linear transformation on the vector space V. N is said to be nilpotent if there exist some positive integer r such that $N^r = 0$.

Theorem 6.5.1. *Let T be a linear transformation on the finite dimensional vector space V over the field F. Suppose that minimal polynomial for T decomposes over F into linear a product of linear polynomials. Then there is a diagonalizable transformation D on V and nilpotent operator N on V such that*

$$T = D + N; \tag{6.16}$$
$$DN = ND. \tag{6.17}$$

The transformation D and N are uniquely determined and each of them is a polynomial in T.

We will now see the process to find D and N for a given linear transformation T.

1. Calculate minimal polynomial of T and factor into linear polynomials $p_i = x - \alpha_i$.

2. In the notation of above theorem, we calculate E_i and note that range of E_i is null space W_i of $(T - \alpha_i I)^{r_i}$.

3. Let $D = \alpha_1 E_1 + ... + \alpha_k E_k$ and observe that D is a diagonalizable transformation. We call D diagonalizable part of T.

4. Let $N = T - D$. We prove below that N so defined is nilpotent transformation.

Proof that N defined as above is a nilpotent transformation. Note that the range space of E_i is the null space W_i of $(T-\alpha_i)^{r_i}$.

$$I = E_1 + E_2 + ... + E_k \tag{6.18}$$
$$\Rightarrow T = TE_1 + ... + TE_k \tag{6.19}$$
$$D = \alpha_1 E_1 + ... + \alpha_k E_k \tag{6.20}$$

Therefore $N = T - D$ becomes

$$N = (T-\alpha_1 I)E_1 + ... + (T-\alpha_k I)E_k \tag{6.21}$$
$$N^2 = (T-\alpha_1 I)^2 E_1 + ... + (T-\alpha_k I)^2 E_k \tag{6.22}$$
$$N^r = (T-\alpha_1 I)^r E_1 + ... + (T-\alpha_k I)^r E_k \tag{6.23}$$

When $r \geq r_i$ for every i, then $N^r = 0$, because the transformation $(T-\alpha_k I)^i$ will be a null transformation on the range of E_i. Therefore N is nilpotent transformation.

Example 10. Find the basis in which following matrix has triangular form and find that triangular form.

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

Solution: Process to find triangular form of a matrix is as follows-

Step 1: Find at least one eigenvalue and corresponding eigenvector of A. For above matrix characteristic polynomial is $f(\lambda) = \lambda^3$. Hence 0 is an eigenvalue which is repeated thrice. Eigenvectors are $\begin{pmatrix} -1, 0, 1 \end{pmatrix}, \begin{pmatrix} 0, 0, 0 \end{pmatrix}$ and $\begin{pmatrix} 0, 0, 0 \end{pmatrix}$

Step 2: Now Note that $u_1 = \begin{pmatrix} -1, 0, 1 \end{pmatrix} \in \ker A$ and $\ker A \subset \ker A^2 \subset \ker A^3$. If $u_2 \in \ker A^2 - \ker A$ then $Au_2 \in \ker A =< u_1 >$ $Au_2 = \alpha u_1$ for some scalar α.

$$A = \begin{pmatrix} 2 & -2 & 2 \\ 0 & 0 & 0 \\ -2 & 2 & -2 \end{pmatrix}$$

and $\ker A^2 =< \begin{pmatrix} 1, 1, 0 \end{pmatrix}, \begin{pmatrix} 0, 1, 1 \end{pmatrix} >$ Taking $u_2 = \begin{pmatrix} 1, 1, 0 \end{pmatrix}$ and $u_3 = \begin{pmatrix} 1, 0, 0 \end{pmatrix} >$.

Step 3: In the basis u_1, u_2 and u_3, given matrix A has triangular form-

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & -2 \end{pmatrix}$$

6.6 Chapter End Exercise

1. Let A be an invertible matrix. If v is an eigenvector of A, show it is also an eigenvector of both A^2 and A^{-2}. What are the corresponding eigenvalues?

2. Let C be a 2×2 matrix of real numbers. Give a proof or counter example to the assertion that if C has two distinct eigenvalues then so does C^2.

3. Let A be $n \times n$ have an eigenvalue λ with corresponding eigenvector v, then state with reason whether following is true or false

   (a) $-\lambda$ is an eigenvalue of $-A$

   (b) If v is an also an eigenvector of B with eigenvalue μ, then $\lambda\mu$ is an eigenvalue of AB.

   (c) Let $\kappa \in F$. Then $\kappa\lambda$ is an eigenvalues of κA.

4. Let

$$A = \begin{pmatrix} 6 & -3 & -2 \\ 4 & -1 & -2 \\ 10 & -5 & -3 \end{pmatrix}$$

   Is A similar over the firld R to a diagonal matrix? Is A is similar over the field C to a diagonal matrix?

5. Let A and B be an $n \times n$ matrices over the field F. Prove that if $(I - AB$ is invertible, then $(I - BA)$ is invertible and

$$(I - BA)^{-1} = I + B(I - AB)^{-1}A$$

6. Use above result to prove that A and B are $n \times n$ matrices over the field F, then AB and BA have precisely the same characteristic values in F.

7. Let $a, bandc$ be elements of a field F, and let A be the following matrix over F;

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

Prove that the characteristic polynomial of A is $x^3 - ax^2 - bx - c$ and that this is also the minimal polynomial for A.

8. Find a 3×3 matrix for which the minimal polynomial is x^2.

9. Is it true that every matrix A such that $A^2 = A$ is similar to a diagonal matrix. If true, prove your assertion otherwise give a counter example.

10. Let T be a linear operator on V. if every subspace of V is invariant under T, then prove that T is a scalar multiple of the identity operator.

11. Let T be a linear operator on a finte dimensional vector space over an algebraically closed field F. Let f be a polynomial over F. Prove that α is a characteristic value of $f(T)$ if and only if $\alpha = f(\kappa)$, where κ is a characteristic value of T.

Chapter 5

Inner Product Spaces

Chapter Structure

5.1 Introduction

Vector space structure on a set is purely an algebraic structure. We simply mention the way, by means of this structure, how to add and subtract two vectors. In general we also talk about geometrical properties of vectors. Then question arise which concepts in general describe geometric properties of vectors. Once such concepts are there with us then we can discuss terms like orthogonality in case of vector spaces where apparently elements do not look like Euclidean vectors. Inner product is that concept. Inner product in case of Euclidean vectors is simply a dot product of two vectors. We have seen that at elementary level dot products to a quite larger extent describe geometry of Euclidean spaces. In other words all propositions of geometry, in one or other way, are consequence of the fact that dot product is defined on Euclidean spaces. Thus we define a real valued function known as inner product on a vector space and try to see what impressions this function makes on vector space structure and linear transformations defined on these vector spaces.

5.2 Objectives

After going through this chapter you will be able to:
• decide whether given real valued function is an inner product on a vector space
• decide whether given pair of vectors is orthogonal
• decide how vivid vector space structure becomes and different look linear transformations get because of defining inner product on vector spaces.

5.3 Inner Product

Definition 1. Let F be a field of real numbers or field of complex numbers and V is a vector space over F. An inner product on V is a function which assigns to each ordered pair of vectors $u, v \in V$ a scalar $< u, v >\in F$ in such a way that for all $u, v, w \in V$ and all scalars α

1. $\langle u+v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$
2. $\langle \alpha u, v\rangle = \alpha\langle u, v\rangle$
3. $\langle u, v\rangle = \overline{\langle v, u\rangle}$
4. $\langle u, u\rangle \neq 0$ if $u \neq 0$

Observation 5.3.1. Without complex conjugation in the definition, we would have the contradiction:

$$\langle u, u\rangle > 0 \text{ and } \langle \imath u, \imath u\rangle = -1\langle u, u\rangle > 0 \text{ for } u \neq 0$$

Example 1. On $\mathbb{R}^n$ there is an inner product which is known as **Standard Inner Product**. It is defined as the dot product of two coordinate vectors.

Example 2. For $u = (x, y), v = (x_1, y_1)$ in $\mathbb{R}^2$, let

$$\langle u, v\rangle = xx_1 - yx_1 - xy_1 + 4yy_1$$

Then $\langle u, v\rangle$ defines an inner product on the vectors of $\mathbb{R}^2$.

Example 3. Let V be $F^{n\times n}$, the space of all $n \times n$ matrices over F. Here F is either field of real numbers or field of complex numbers. Then the following defined is an inner product on V.

$$\langle A, B\rangle = \sum_{j,k} A_{jk}\overline{B_{jk}}$$

- Verify that above inner product can be expressed in the following way-
$$\langle A, B\rangle = \mathbf{tr}(AB^*) = \mathbf{tr}(B^*A)$$

- Let V and W be vector spaces over a field F. Where F is either a field of real numbers or field of complex numbers. Let T be a non-singular linear transformation from V into W. If $\langle .,.\rangle$ is an inner product on V. Then prove that $\langle Tu, Tv\rangle$ is an inner product on W.

- Let v be the vector space of all continuous complex valued functions on the unit interval, $0 \leq t \leq 1$. Let
$$\langle f, g\rangle = \int_0^1 f(t)\overline{g(t)}dt$$
Prove that $\langle .,.\rangle$ so defined is an inner product.

- Let V be a vector space on the field of complex numbers. Let $\langle .,.\rangle$ be an inner product defined on V. Then prove that the following holds.
$$\langle u, v\rangle = \mathbf{Re}\langle u, v\rangle + \imath\mathbf{Re}\langle u, \imath v\rangle$$

Definition 2. An inner product space is a real or complex vector space, together with a specified inner product on that space. A finite dimensional real inner product space is called Euclidean space and a finite dimensional complex inner product space is called a unitary space.

Definition 3. The quadratic form determined by the inner product is the function that assigns to each vector u the scalar $||u||^2$ defined as
$$||u||^2 = \langle u, u\rangle$$

Note that $||u||$ satisfies the following identity
$$||u \pm v||^2 = ||u||^2 \pm 2\mathbf{Re}\langle u, v\rangle + ||v||^2 \forall u, v \in V$$

- For a real inner product prove the following:
$$\langle u, v\rangle = \frac{1}{4}||u+v||^2 - \frac{1}{4}||u-v||^2$$

- For a complex inner product prove the following
$$\langle u, v\rangle = \frac{1}{4}||u+v||^2 - \frac{1}{4}||u-v||^2 + \frac{\imath}{4}||u+iv||^2 - \frac{\imath}{4}||u-iv||^2$$
Above identities are known as **polarization identities**

Theorem 5.3.1. *Let V be an inner product space, then for any vectors u, v in V and scalar α*

1. $||\alpha u|| = |\alpha|||u||$;
2. $||u|| \geq 0$ *for* $u \neq 0$;
3. $|\langle u, v\rangle| \leq ||u||||v||$;
4. $||u + v|| \leq ||u|| + ||v||$.

Proof. 1. First follows easily in the following way

$$\begin{aligned} ||\alpha u||^2 &= \langle \alpha u, \alpha u\rangle \\ &= \alpha\overline{\alpha}\langle u, u\rangle \\ &= \alpha\overline{\alpha}||u||^2 \end{aligned}$$

Therefore it follows that

$$||\alpha u|| = |\alpha|||u||$$

2. This follows immediately from definition of inner product.

3. This inequality is true if $u = 0$. Suppose $u \neq 0$ Let

$$w = v - \frac{\langle v, u\rangle}{||u||^2}u$$

then $\langle w, u\rangle = 0$ and $0 \leq ||w||^2$

$$\begin{aligned} ||w||^2 &= (\langle w - \frac{\langle w, u\rangle}{||u||^2}u, w - \frac{\langle w, u\rangle}{||u||^2}u\rangle) \\ &= \langle w, w\rangle - \frac{\langle w, u\rangle\langle u, w\rangle}{||u||^2} \\ &= ||w||^2 - \frac{|\langle u, v\rangle|^2}{||u||^2} \end{aligned}$$

Hence third inequality.

4. Using third inequality in the following way we get the fourth inequality–

$$\begin{aligned} ||u + v||^2 &= ||u||^2 + \langle u, v\rangle + \langle v, u\rangle + ||v||^2 \\ &= ||u||^2 + 2Re\langle u, v\rangle + ||v||^2 \\ &\leq ||u||^2 + 2||u||||v|| + ||v||^2 \\ &= (||u|| + ||v||)^2 \end{aligned}$$

Therefore

$$||u + v|| \leq ||u|| + ||v||$$

The third inequality above is called **Cauchy-Schwarz inequality**. Equality occurs in the third if and only if u and v are linearly dependent. □

5.4 Orthogonalization

Definition 4. Let u and v be vectors in an inner product space V. Then u is orthogonal to v if $\langle u, v\rangle = 0$. If S is a set of vectors in V then S is said to be orthogonal set if for each pair of distinct vectors $u, v \in S$ we have $\langle u, v\rangle = 0$. Such an orthogonal set is said to be orthonormal if for every $u \in S$ we have that $||u|| = 1$.

Example 4. Find unit vector orthogonal to $v_1 = \begin{pmatrix} 1, 1, 2 \end{pmatrix}, v_2 = \begin{pmatrix} 0, 1, 3 \end{pmatrix}$

Solution: If w is a vector which is orthogonal to v_1 and v_2 then it satisfies that $\langle w, v_1\rangle = 0$ and $\langle w, v_2\rangle = 0$ This leads to homogeneous system of linear equation-

$$\begin{aligned} x + y + 2z &= 0 \\ y + 3z &= 0 \end{aligned}$$

where $w = \begin{pmatrix} x, y, z \end{pmatrix}$ Upon solving this system of equation we get $x = 1, y = -3, z = 1$. Normalizing this vector we get unit vector orthogonal to v_1 and v_2 which is thus $v_1 = \begin{pmatrix} 1/\sqrt{11}, -3/\sqrt{11}, 1/\sqrt{11} \end{pmatrix}$.

Theorem 5.4.1. *An orthogonal set of non-zero vectors is linearly independent.*

Proof. Let S be an finite or infinite orthogonal set of nonzero vectors in a given inner product space. Suppose $u_1, u_2, ..., u_n$ are distinct vectors in S and that

$$v = \alpha_1 u_1 + \alpha_2 u_2 + ... + \alpha_n u_n$$

then

$$\begin{aligned} \langle v, u_k\rangle &= \langle \sum_j \alpha_j u_j, u_k\rangle \\ &= \sum_j \alpha_j \langle u_j, u_k\rangle \\ &= \alpha_k \langle u_k, u_k\rangle \end{aligned}$$

Since $\langle u_k, u_k\langle \neq 0$ it follows that

$$\alpha_k = \frac{\langle v, u_k\rangle}{||u_k||^2}$$

Thus when $v = 0$ we get that each $\alpha_k = 0$; therefore S is an independent set. □

Corollary 5.4.1. *If a vector v is a linear combination of an orthogonal sequence of non zero vectors $u_1, u_2, ..., u_n$, then v is the particular linear combination*

$$v = \sum_{k=1}^{m} \frac{\langle v, u_k \rangle}{||u_k||^2} u_k$$

Theorem 5.4.2. *Let V be an inner product space and let $u_1, u_2, ..., u_n$ be any independent vectors in V. Then one may construct orthogonal vectors $v_1, v_2, ...v_n$ in V such that for each $k = 1, 2, ...n$ the set*

$$v_1, v_2, ..., v_n$$

is a basis for the subspace generated by $u_1, u_2, ..., u_n$.

Proof. We will achieve claim of made in the theorem by explicitly determining $v_1, v_2, ...v_n$ when $u_1, u_2, ..., u_n$ are given. This process is known as **Gram-Schmidt orthogonalization process**
Let

$$v_1 = u_1$$

Suppose $v_1, v_2, ..., v_m$ vectors of sought n vectors are constructed such that these vectors span the subspace spanned by $u_1, u_2, ..., u_m$. Then v_{m+1} is defined as follows

$$v_{m+1} = u_{m+1} - \sum_{k=1}^{m} \frac{\langle u_{m+1}, v_k \rangle}{||v_k||^2} v_k$$

Then $v_{m+1} \neq 0$ otherwise u_{k+1} is a linear combination of $v_1, v_2, ..., v_m$ and hence linear combination of $u_1, u_2, ..., u_m$. Furthermore, if $1 \leq j \leq m$ then

$$\begin{aligned}
\langle v_{m+1}, v_j \rangle &= \langle u_{m+1}, v_j \rangle - \sum_{k=1}^{m} \frac{\langle u_{m+1}, v_k \rangle}{||v_k||^2} \langle v_k, v_j \rangle \\
&= \langle u_{m+1}, v_j \rangle - \langle u_{m+1}, v_j \rangle \\
&= 0
\end{aligned}$$

Therefore $v_1, v_2, ..., v_{m+1}$ is an orthogonal set consisting of $m = 1$ nonzero vectors in the subspace spanned by $u_1, u_2, ..., u_{m+1}$. This set therefore is a basis for this subspace. This completes the construction and proof of the theorem. □

Corollary 5.4.2. *Every finite dimensional inner product space has an orthonormal basis.*

Proof. Let V be a finite dimensional inner product space and $u_1, u_2, ..., u_n$ be a basis for V. We apply the Gram-Schmidt process to construct orthogonal basis $v_1, v_2, ..., v_n$. Then we obtain orthonormal basis simply by replacing v_k by $\frac{v_k}{||v_k||}$. □

Example 5. Consider the following basis of Euclidean space $\mathbb{R}^3$.

$$v_1 = (1,1,1)\ v_2(0,1,1) \text{ and } v_3 = (0,0,1)$$

Transform this basis to orthonormal basis using Gram-Schmidt orthogonalization process.

Solution:

$$u_1 = \frac{v_1}{||v_1||} = \frac{(1,1,1)}{\sqrt{3}}$$

Now we find w_2 as follows-

$$w_2 = v_2 - \langle v_2, u_1\rangle u_1 = (0,1,1,) - \tfrac{2}{3}(\tfrac{(1,1,1)}{\sqrt{3}}) = (-\tfrac{2}{3}, \tfrac{1}{3}, \tfrac{1}{3})$$

Normalizing w_2 we get

$$u_2 = \frac{w_2}{||w_2||} = (-\tfrac{2}{\sqrt{6}}, \tfrac{1}{\sqrt{6}}, \tfrac{1}{\sqrt{6}})$$

Now we write

$$w_3 = v_3 - \langle v_3, u_1\rangle u_1 - \langle v_3, u_2\rangle u_2 = (0, -\tfrac{1}{2}, \tfrac{1}{2})$$

Normalizing we get $u_3 = (0, -\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$
u_1, u_2 and u_3 together form the required orthogonal basis.

Example 6. Find real orthogonal matrix P such that P^tAP is diagonal for the following matrix A.

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

Solution: First find characteristic polynomial for A which is $(\lambda - 1)^2(\lambda - 4)$. Thus eigenvalues of A are 1 (with multiplicity two) and 4 (with multiplicity one) Solve $(A - \lambda I)X = 0$ for $\lambda = 1$ and we get following homogeneous system of equations:

$$\begin{aligned} -x - y - z &= 0 \\ -x - y - z &= 0 \\ -x - y - z &= 0 \end{aligned}$$

That is $x + y + z = 0$ This system has two independent solutions. One such solution is $v_1 = (1, -1, 0)$.We seek a second solution $v_2 = (a, b, c)$ which is orthogonal to v_1 that is such that

$$a + b + c = 0 \text{ and } a - b = 0$$

One of the solution to these equations is $v_2 = (1, 1, -2)$ Next we normalize v_1 and v_2 to obtain the unit orthogonal solutions

$$u_1 = (1/\sqrt{2}, -1/\sqrt{2}, 0), u_2 = (1/\sqrt{6}, 1/\sqrt{6}, -2/\sqrt{6})$$

Similarly solution of $(A - \lambda I)X = 0$ for $\lambda = 4$ is $v_3 = (1, 1, 1)$ and normalize it to obtain $u_3 = (1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3})$ Matrix P whose columns are vectors u_1, u_2 and u_3 form an orthogonal matrix such that $P^t AP$ is diagonal matrix and corresponding diagonal matrix is the following-

$$P = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{6} & 1/\sqrt{3} \\ -1/\sqrt{2} & 1/\sqrt{6} & 1/\sqrt{3} \\ 0 & -2/\sqrt{6} & 1/\sqrt{3} \end{bmatrix}$$

$$P^t AP = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{bmatrix}$$

Definition 5. If W is a finite dimensional subspace of a inner product space V and $u_1, u_2, ..., u_n$ is any orthonormal basis for W, then the vector u defined as follows is known as **orthogonal projection of** $v \in V$**.**

$$u = \sum_k \frac{\langle v, u_k \rangle}{||u_k||^2} u_k$$

The mapping that assigns to each vector in V its orthogonal projection is called **orthogonal projection of** V **on** W

Definition 6. Let V be an inner product space and S be any set of vectors in V. The orthogonal complement of S is the set $S^\perp$ of all vectors in V which are orthogonal to every vector in S.

Theorem 5.4.3. *Let W be a finite dimensional subspace of an inner product space V and let E be orthogonal projection of V on W. Then E is idempotent linear transformation of V onto W , $W^\perp$ is the null space of E and*

$$V = W \bigoplus W^\perp$$

In this case, $I - E$is the orthogonal projection of V on $W^\perp$. It is an idempotent linear transformation of V onto $W^\perp$ with null space W.

Bessel's Inequality

Let $v_1, v_2, ...v_n$ be an orthonormal set of nonzero vectors in an inner product space V. If u is any vector in V, then-

$$\sum_k \frac{|\langle u, v_k \rangle|^2}{||v_k||^2} \leq ||u||^2$$

Theorem 5.4.4. *Let V be a finite dimensional vector space and f be a linear functional on V. Then there exists a unique vector v in V such that $f(u) = \langle u, v \rangle$ for all u in V.*

Note that v lies in the orthogonal complement of the null space of f.

Theorem 5.4.5. *For any linear transformation/operator T on a finite dimensional inner product space V, there exists a unique linear operator T^* on V such that*

$$\langle Tu, v\rangle = \langle u, T^*v\rangle$$

for all u and v in V.

Proof. Let v be any vector in V. Then $u \mapsto \langle Tu, v\rangle$ is a linear functional on V. As we have stated above there exist unique vector v' in V such that $\langle Tu, v\rangle = \langle u, v'\rangle$ for every u in V. Let T^* denote the mapping $v \mapsto v'$

$$v' = T^*v$$

For any u, v, w in V consider the following for any scalar α

$$\begin{aligned}
\langle u, T^*(\alpha v + w)\rangle &= \langle Tu, \alpha v + w\rangle \\
&= \langle Tu, \alpha v\rangle + \langle Tu, w\rangle \\
&= \overline{\alpha}\langle Tu, v\rangle + \langle Tu, w\rangle \\
&= \overline{\alpha}\langle u, T^*v\rangle + \langle u, T^*w\rangle \\
&= \langle u, \alpha T^*v\rangle + \langle u, T^*w\rangle \\
&= \langle u, cT^*v + T^*w\rangle
\end{aligned}$$

Thus

$$\langle u, T^*(\alpha v + w)\rangle = \langle u, cT^*v + T^*w\rangle$$

hence T^* is linear. Now we prove uniqueness. Note that for any v in V, the vector T^*v is uniquely determined as the vector v' such that $\langle Tu, v\rangle = \langle u, v'\rangle$ for every u. □

Theorem 5.4.6. *Let V be a finite dimensional inner product space and let $B = u_1, u_2, .., u_n$ be an ordered orthonormal basis for V. Let T be a linear operator on V and let A be the matrix of T in the basis B. Then $A_{kj} = \langle Tu_j, u_k\rangle$.*

Proof. Since B is an orthonormal basis, we have

$$u = \sum_{k=1}^{n} \langle u, u_k\rangle u_k$$

The matrix A is defined by

$$Tu_j = \sum_{k=1}^{n} A_{kj} u_k$$

and since

$$Tu_j = \sum_{k=1}^{n} \langle Tu_j, u_k\rangle u_k$$

we have that $A_{kj} = \langle Tu_j, u_k\rangle$. □

Corollary 5.4.3. *Let V be a finite dimensional inner product space, and let T be a linear operator on V. In any orthonormal basis for V, the matrix of T^* is the conjugate transpose of the matrix of T.*

Proof. Let $B = u_1, u_2, .., u_n$ be an orthonormal basis for V, let $A = [T]_B$ and $B = [T^*]_B$. From above theorem we have

$$\begin{aligned} A_{kj} &= \langle Tu_j, u_k \rangle \\ B_{kj} &= \langle T^* u_j, u_k \rangle \end{aligned}$$

By the definition of T^* we have

$$\begin{aligned} B_{kj} &= \langle T^* u_j, u_k \rangle \\ &= \overline{\langle u_k, T^* u_j \rangle} \\ &= \overline{Tu_k, u_j} \\ &= \overline{A_{jk}} \end{aligned}$$

□

5.5 Adjoint of a Linear Transformation

Definition 7. Let T be a linear operator on an inner product space V. Then we say that T has an adjoint on V if there exists a linear operator T^* on V such that $\langle Tu, v \rangle = \langle u, T^* v \rangle$ for all u and v in V.

Note that for a linear operator on finite dimensional inner product space there always exist an adjoint but there exist infinite diemnsional inner product spaces and linear operators on them for which there is no corresponding adjoint operator.

Theorem 5.5.1. *Let V be a finite dimensional inner product space. If T and U are linear operators on V and α is scalar,*

1. $$(T + U)^* = T^* + U^*$$

2. $$(\alpha T)^* = \overline{\alpha} T^*$$

3. $$(TU)^* = U^* T^*$$

4. $$(T^*)^* = T$$

Proof. 1. Let u and v be in V. Then

$$\begin{aligned}\langle (T+U)u, v\rangle &= \langle Tu + Uu, v\rangle \\ &= \langle Tu, v\rangle + \langle Uu, v\rangle \\ &= \langle u, T^*v\rangle + \langle u, U^*v\rangle \\ &= \langle u, T^*v + U^*v\rangle \\ &= \langle u, (T^* + U^*)v\rangle\end{aligned}$$

From the uniqueness of adjoints we have that $(T+U)^* = T^* + U^*$

2. Consider

$$\begin{aligned}\langle \alpha Tu, v\rangle &= \langle Tu, \overline{\alpha}v\rangle \\ &= \langle u, T^*\overline{\alpha}v\rangle \\ &= \langle u, \overline{\alpha}T^*v\rangle\end{aligned}$$

From the uniqueness of adjoints we get $(\alpha T)^* = \overline{\alpha}T^*$.

3. Note the following

$$\langle TUu, v\rangle = \langle Uu, T^*v\rangle = \langle u, U^*T^*v\rangle$$

Uniqueness of adjoint of a linear operator proves the third identity.

4. Note the following

$$\langle T^*u, v\rangle = \overline{\langle v, T^*u\rangle} = \overline{\langle Tv, u\rangle} = \langle u, TV\rangle$$

and the fourth identity follows.

□Note that if T is a linear operator on finite dimensional complex inner product space, then

$$T = U_1 + iU_2$$

where $U_1 = U_1^*$ and $U_2 = U_2^*$.This expression for T is unique and

$$\begin{aligned}U_1 &= \frac{1}{2}(T + T^*) \\ U_2 &= \frac{1}{2i}(T - T^*)\end{aligned}$$

Definition 8. A linear operator T such that $T = T^*$ is called **self adjoint** or **Hermitian**.

5.5.1 Unitary Operators

Definition 9. Let V and W be inner product spaces over the same field and let T be a linear transformation from V into W. We say that T preserves inner product if $\langle Tu, Tv\rangle = \langle u, v\rangle$ for all u, v in V.An **isomorphism** of V onto W is a vector space isomorphism which preserves inner products.

Definition 10. A **unitary operator** on an inner product space is an isomorphism onto itself.

Theorem 5.5.2. *Let V and W be inner product spaces over the same field and let T be a linear transformation from V into W. Then T preserves inner product if and only if $||Tu|| = ||u||$ for every u in V.*

Proof. If T preserves inner product then it follows that $||Tu|| = ||u||$. For converse part, we prove the result for real inner product spaces. For complex inner product spaces result follows on similar lines except that we have to consider polarization identity for complex inner product spaces. So, let our inner product spaces are real and let $||Tu|| = ||u||$. Consider polarization identity-

$$\langle u, v\rangle = \frac{1}{4}||u+v||^2 - \frac{1}{4}||u-v||^2$$

$$\begin{aligned} \langle Tu, Tv\rangle &= \frac{1}{4}||Tu+Tv||^2 - \frac{1}{4}||Tu-Tv||^2 \\ &= \frac{1}{4}||T(u+v)||^2 - \frac{1}{4}||T(u-v)||^2 \\ &= \frac{1}{4}||u+v||^2 - \frac{1}{4}||u-v||^2 \\ &= \langle u, v\rangle \end{aligned}$$

□

Theorem 5.5.3. *Let U be a linear operator on an inner product space V. Then U is unitary if and only if the adjoint U^* of U exists and $UU^* = U^*U = I$.*

Proof. Suppose U is unitary. Then U is invertible and

$$\langle Uu, v\rangle = \langle Uu, UU^{-1}v\rangle = \langle u, U^{-1}v\rangle$$

for all u and v from V. From definition of adjoint of an operator then it follows that U^{-1} satisfies properties of adjoint and hence U^{-1} is the adjoint of U. It is trivial now to see $UU^* = U^*U = I$ Conversely let adjoint exists and it is U^{-1}. We need to show that U preserves inner product.

$$\begin{aligned}\langle Uu, Uv\rangle &= \langle u, U^*Uv\rangle \\ &= \langle u, Iv\rangle \\ &= \langle u, v\rangle\end{aligned}$$

for all u and v. □

Definition 11. A complex $n \times n$ matrix A is called unitary, if $A^*A = I$.

Definition 12. A real or complex $n \times n$ matrix A is said to be orthogonal, if $A^tA = I$.

Definition 13. Let A and B be complex $n \times n$ matrices. We say that B is unitarily equivalent to A if there is an $n \times n$ unitary matrix P such that $B = P^{-1}AP$. We say that B is orthogonally equivalent to A if there is an $n \times n$ orthogonal matrix P such that $B = P^{-1}AP$.

5.5.2 Normal Operators

Definition 14. Let V be a finite dimensional inner product space and T a linear operator on V. We say that T is normal if it commutes with its adjoint.i.e., $TT^* = T^*T$.

Observation 5.5.1. Any self-adjoint operator is normal. Any unitary operator is normal. Any scalar multiple of normal operator is normal. Note however that sums and products of normal operators are not, in general, normal.

Theorem 5.5.4. *Let V be an inner product space and T be a self adjoint linear operator on V. Then each characteristic value of T is real, and characteristics vectors associated of T associated with distinct characteristic values are orthogonal.*

Proof. Let α be a characteristic value of T. Thus $Tu = \alpha u$ for some nonzero vector u.Then

$$\begin{aligned}\alpha\langle u, u\rangle &= \langle \alpha u, u\rangle \\ &= \langle Tu, u\rangle \\ &= \langle u, Tu\rangle \\ &= \langle u, \alpha u\rangle \\ &= \overline{\alpha}\langle u, u\rangle\end{aligned}$$

Since $u \neq 0$, we must have $\alpha = \overline{\alpha}$.i.e., α is real. Suppose we also have $Tv = \beta v$ with $v \neq 0$. Then

$$\begin{aligned} \alpha\langle u, v\rangle &= \langle Tu, v\rangle \\ &= \langle u, Tv\rangle \\ &= \langle u, \beta v\rangle \\ &= \overline{\beta}\langle u, v\rangle \\ &= \beta\langle u, v\rangle \end{aligned}$$

If $\alpha \neq \beta$, then it follows that $\langle u, v\rangle = 0$. Proving orthogonality of u and v. □

Theorem 5.5.5. *On a finite dimensional inner product space of positive dimension, every self adjoint operator has a (nonzero) characteristic vector.*

Proof. Let V be a finite dimnsional inner product space of dimension n, where $n > 0$. Let T be a self adjoint operator on V. Choose an orthonormal basis B for V and let $A = [T]_B$ Since $T = T^*$, we have $A = A^*$. Let W be the space of $n \times 1$ matrices over C, with inner product $\langle X, Y\rangle = Y^*X$. Then $U(X) = AX$ defines a self adjoint operator U on W. The characteristic polynomial, $det(xI - A)$, is a polynomial of degree n over the field of complex numbers. Every polynomial over C has a root. Thus there exists a complex number α such that $det(\alpha I - A) = 0$.This means that $A - \alpha I$ is singular, or that there exist a non zero X such that $AX = \alpha X$. Since multiplication by A is self adjoint it follows that α is real. If V is real then one may choose X with real entries. For then A and $A - \alpha I$ have real entries, and since $A - \alpha I$ is singular, the system $(A - \alpha I)X = 0$ has a nonzero real solution X. In this way we have that $Tu = \alpha u$. □.

Theorem 5.5.6. *Let V be a finite dimensional inner product space and let T be any linear operator on V. Suppose W is a subspace of V which is invariant under T. Then the orthogonal complement of W is invariant under T^*.*

Proof. : W is invariant under T means if u is in W then Tu is in W.Let v be in $W^{\perp}$. Let $u \in W$. Then $Tu \in W$. Now

$$\begin{aligned} 0 &= \langle Tu, v\rangle \\ &= \langle u, T^*v\rangle \end{aligned}$$

This shows that $u \perp T^*v$. This proves that T^*v is in $W^{\perp}$. Therefore if v be in $W^{\perp}$ then T^*v is in $W^{\perp}$. Hence the proof. □

Theorem 5.5.7. *If T is a normal operator on finite dimensional inner product space V then the operator U defined for any scalar α by $U = T - \alpha I$ is normal.*

Proof. Note that $(T - \alpha I)^* = T^* - \overline{\alpha} I$.

$$\begin{aligned} UU^* &= (T - \alpha I)(T - \alpha I)^* \\ &= (T - \alpha I)(T^* - \overline{\alpha} I) \\ &= (T^* - \overline{\alpha} I)(T - \alpha I) \\ &= U^* U \end{aligned}$$

Thus U as defined above is normal. □

Theorem 5.5.8. *Let V be a finite dimensional inner product space and T a normal operator on V. Suppose u is a vector in V. Then u is a characteristic vector of T with characteristic value α if and only if u is a characteristic vector of T^* with characteristic value $\overline{\alpha}$*

Proof. Suppose U is any normal operator on V. Then

$$||Uu||^2 = \langle Uu, Uu \rangle = \langle u, U^* Uu \rangle = \langle u, UU^* u \rangle = \langle U^* u, U^* u \rangle = ||U^* u||^2$$

Which implies that $||Uu|| = ||U^* u||$. If α is any scalar then we saw in above theorem that the operator $U = T - \alpha I$ is normal. Thus

$$||(T - \alpha I)u|| = ||(T^* - \overline{\alpha} I)u||$$

and $(T - \alpha I)u = 0$ if and only if$(T^* - \overline{\alpha} I)u = 0$. □

Definition 15. A complex $n \times n$ matrix is called normal if $AA^* = A^* A$

Theorem 5.5.9. *write theorem statement here.*

Proof. this is the proof of the theorem. □If there is some corollary to this theorem then you may write like this:

Corollary 5.5.1. *corollary to above theorem.*

If you want give some problems for practice then write this:

Check Your Progress Prove or give counter example for the following assertions where v, w, z are vectors in a real inner product space H.

1. If $\langle v, w \rangle = 0$ and $\langle v, z \rangle = 0$ then $\langle w, z \rangle = 0$
2. If $\langle v, z \rangle = \langle w, z \rangle$ for all $z \in H$, then $v = w$
3. If A is an $n \times n$ symmetric matrix then A is invertible.

5.6 Chapter End Exercise

1. Prove that an angle inscribed in a semicircle is a right angle.

2. Let v, w be vectors in the plane R^2 with lengths $3, 5$ respectively. What is the maxima and minima of the length of $v + w$?

3. Let A be the following matrix. Show that the bilinear map $R^3 \times R^3 \to R$ defined by $\langle x, y \rangle = x^T A y$ is a scalar product.

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

4. Let $S \subset R^4$ be the vectors that satisfy $X = (x_1, x_2, x_3, x_4)$ that satisfy $x_1 + x_2 - x_3 + x_4 = 0$. What is dimension of S. Find orthogonal complement of S.

5. Let $L : R^3 \to R^3$ be a linear map with the property that $Lv \perp v$ for every $v \in R^3$ Prove that L can not be invertible.
   Is a similar assertion is true for linear map $L : R^2 \to R^2$?

6. In a complex vector space with hermitian inner product on it, if a matrix A satisfies $< x, Ax >= 0$ for all vectors x, show that $A = 0$.

7. Let A be a square matrix of real numbers whose columns are (non zero) orthogonal vectors. Then show that $A^T A$ is a diagonal matrix.

Chapter 7

Bilinear Forms

7.1 Introduction

Linear transformation is a linear function of one variable. Then question arises of defining a linear function of two variables, concept of bilinear form arose out of this need. But again what does it mean by linearity in two variables? Natural answer to this question is that what defines a bilinear form. Theory of bilinear forms (and multilinear forms) has developed by generalizing concepts from one variable in a most natural way. Here natural means we take for generalization obvious choices and establish that they are unambiguous. We have taken most of the text from book by Hoffman and Kunz and care has been taken that reader will have to go to original text, the least number of times.

7.2 Objectives

After going through this chapter you will be able to:
• Check whether given expression is bilinear form and classify whether it is degenerate, non-degenerate, symmetric, skewsymmetric bilniear form
• Find matrix of a bilinear form in the given basis and switching from one basis to the other
• Diagonalization of a bilinear form and find its signature

7.3 Bilinear Form and its types

Definition 24. Bilinear Form Bilinear form on a vector space V is a function of two variables on V, with values in the field F satisfying the bilinear axioms which are-

$$\begin{aligned} f(v_1 + v_2, w) &= f(v_1, w) + f(v_2, w) \\ f(\alpha v, w) &= \alpha f(v, w) \\ f(v, w_1 + w_2) &= f(v, w_1) + f(v, w_2) \\ f(v, \alpha w) &= \alpha f(v, w) \end{aligned}$$

for all $v, w, v_1, w_1, v_2, w_2 \in V$ and $\alpha \in F$

Bilinear form will be denoted by $\langle v, w \rangle$.

Definition 25. A bilinear form is said to be symmetric if

$$\langle v, w \rangle = \langle w, v \rangle$$

and skew symmetric if

$$\langle v, w \rangle = -\langle w, v \rangle$$

Definition 26. Two vectors u, v are called orthogonal with respect to symmetric form if $\langle u, v \rangle = 0$

Definition 27. A basis B of V is called orthonormal basis with respect to the form if,

$$\langle v_i, v_j \rangle = 0 \text{ for all } i \neq j \text{ and } \langle v_i, v_i \rangle = 1 \text{ for all } i.$$

Remark 7.3.1. If the form is either symmetric or skew symmetric, then the linearity in the first variable follows from linearity in the second variable.

Example Let A be an $n \times n$ matrix in F and define

$$\langle v, w \rangle = X^t A Y$$

where X and Y are co ordinates of v and w respectively in some basis of V.
Then we see that this defines a bilinear form on V. This coincides with usual inner product on V if $A = I$.

Definition 28. A matrix A is called symmetric if $A^t = A$.

Theorem 7.3.1. *Bilinear form given in above example is symmetric if and only if matrix A is symmetric.*

Proof. Assume that A is symmetric. Since Y^tAX is a 1×1 matrix, it is equal to its transpose:$Y^tAX = (Y^tAX)^t = X^tA^tY = X^tAY$ and hence $\langle Y, X\rangle = \langle X, Y\rangle$ and it follows that form is symmetric. Conversely let the form is symmetric. Set $X = e_i$ and $Y = e_j$ where e_i and e_j are elements of fixed basis. We find that $\langle e_i, e_j\rangle = e_i^tAe_j = a_{ij}$ while $\langle e_j, e_i\rangle = e_j^tAe_i = a_{ji}$ and as the form is symmetric we get that $a_{ij} = a_{ji}$ and the matrix A is symmetric. □

Computation of the value of bilinear form Let $v, w \in V$ and let X and Y be their coordinates in the basis B so that $v = BX$ and $w = BY$ Then

$$\langle v, w\rangle = \langle \sum_i v_i x_i, \sum_j v_j y_j\rangle$$

This expands using bilinearity to $\sum_{i,j} x_i y_j \langle v_i, v_j\rangle = \sum i, j x_i a_{ij} y_j = X^tAY$

$$\langle v, w\rangle = X^tAY$$

Thus if we identify V with F^n using basis B then bilinear form $<,>$ corresponds to X^tAY

Corollary 7.3.1. *Let A be a matrix of a bilinear form with respect to a basis. The matrices A' which represents the same form with respect to different bases are the matrices $A' = QAQ^t$ where Q is arbitrary matrix in $GL_n(F)$.*

Proof. The change of basis is effected by $B = B'P$ for some matrix P. Then $X' = PX, Y' = PY$. If A' is the matrix of the form with respect to a new basisB', then by definition of A', $\langle v, w\rangle = X'^tA'Y' = X^tP^tA'PY$ but we also have $\langle v, w\rangle = X^tAY$. Therefore

$$P^tA'P = A$$

□

Theorem 7.3.2. *The following properties of a real $n \times n$ are equivalent*

1. *A represents dot product, with respect to some basis of R^n*
2. *There is an invertible matrix $P \in GL_n(R)$ such that $A = P^tP$*
3. *A is symmetric and positive definite.*

Proof. **1 implies 2:** The matrix of the dot product with respect to the standard basis is the identity matrix:$X.Y = X^tIY$. If we change basis, the matrix of the form changes to

$$A = (P^{(-1)t})I(P^{-1}) = (P^{(-1)t})(P^{-1})$$

where P is the matrix of change of basis. Thus A is of the form P^tP and assertion in (2) follows.

2 implies 3: P^tP is always a symmetric and positive definite hence this implication in (2) to (3) follows.

3 implies 1: If A is symmetric and positive definite then for $\langle X, Y\rangle = X^tAY$ is also symmetric and positive definite.

□

Definition 29. A bilinear form which takes on both positive as well as negative values is called indefinite form

For example the Lorentz form defined bellow is an indefinite bilinear form.

$$X^tAY = x_1y_1 + x_2y_2 + x_3y_3 - c^2x_4y_4$$

The coefficient c representing speed of light can be normalized to 1, and then the matrix of the form with respect to given basis is given by

$$\begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & -1 \end{pmatrix}$$

Theorem 7.3.3. *Suppose the symmetric form $\langle , \rangle >$ is not identically zero, then there exist a vector $v \in V$ which is not self orthogonal: $\langle v, v\rangle \neq 0$.*

Proof. Since the form is not identically zero, we have two vectors $u, v \in V$ such that $\langle u, v\rangle \neq 0$. If $\langle v, v\rangle \neq 0$ or $\langle u, u\rangle \neq 0$ then we have the theorem proved. Otherwise suppose $\langle v, v\rangle = 0$ and $\langle u, u\rangle = 0$. Define $w = u + v$ and expand $\langle w, w\rangle$ using bilinearity. We get,

$$\begin{aligned} \langle w, w\rangle &= \langle u + v, u + v\rangle \\ &= \langle u, u\rangle + \langle u, v\rangle + \langle v, u\rangle + \langle v, v\rangle \\ &= 0 + \langle u, v\rangle + \langle v, u\rangle + 0 \\ &= 2\langle u, v\rangle \end{aligned}$$

since $\langle u, v\rangle \neq 0$ it follows that $\langle w, w\rangle \neq 0$. □

Definition 30. Let W be a subspace of V then following defined set is a subspace of V known as orthogonal complement of W.

$$W^\perp = v \in V | \langle v, W\rangle = 0$$

Theorem 7.3.4. *Let $w \in V$ be a vector such that $\langle w, w \rangle \neq 0$. Let $W = \alpha w$ be the span of w. Then V is the direct sum of W and its orthogonal complement:*

$$V = W \bigoplus W^{\perp}$$

Proof. We prove this theorem in two main steps.

1. $W \bigcap W^{\perp} = 0$

2. W and $W^{\perp}$ span V

First assertion follows as w is not self orthogonal and therefore $\langle \alpha w, w \rangle = 0 \Leftrightarrow \alpha = 0$. For second step we need to show every vector $v \in V$ can be expressed as $v = \alpha w + v'$ for some unique α and $v' \in W^{\perp}$. If we take $\alpha = \frac{\langle v,w \rangle}{\langle w,w \rangle}$ and set $v' = v - \alpha w$ then the claim follows. □

Definition 31. A vector $v \in V$ is called null vector for the given form if $< v, w >= 0$ for all $w \in V$.

Definition 32. The null space of the form is the set of all null vectors of V

$$N = v | \langle v, V \rangle = 0 = V^{\perp}$$

Definition 33. A symmetric form is said to be nondegenerate if the null space is 0.

Definition 34. An orthogonal basis $\mathbf{B} = (v_1, v_2, ..., v_n)$ for V, with respect to a symmetric form $\langle , \rangle$ is a basis of V such that $v_i \perp v_j$ for all $i \neq j$

Remark 7.3.2. The matrix A of a form is defined by $a_{ij} = \langle v_i, v_j \rangle$, the basis B is orthogonal if and only if A is diagonal matrix. If the symmetric form $\langle , \rangle$ is nondegenerate and basis $B = (v_1, v_2, ..., v_n)$ is orthogonal, then $\langle v_i, v_i \rangle \neq 0$ for all i, the diagonal entries of A are nonzero.

Theorem 7.3.5. *Let $\langle , \rangle$ be a symmetric form on a real vector space V.*

Vector space form *There is an orthogonal basis for V. More precisely, there exist a basis $B = (v_1, v_2, ..., v_n)$ such that $\langle v_i, v_j \rangle = 0$ for $i \neq j$ and such that for each i, $\langle v_i, v_i \rangle$ is either* $1, -1,$ *or* 0.

Matrix form *Let A be a real symmetric $n \times n$ matrix. There is a matrix $Q \in GL_n(R)$ such that QAQ^t is a diagonal matrix each of whose diagonal entries is 1, -1, or 0.*

Remark 7.3.3. Matrix form of the above theorem follows from its vector space form by noting that symmetric matrix A is a matrix of symmetric form on a vector space.

Proof. We apply induction on dimension n of vector space V. Assume the result to be true for all vector spaces of dimension less than or equal to $n-1$. Let V be a vector space of dimension n. Let the form be not identically zero. Then we know that there is a vector $v = v_1$ which is not self orthogonal i.e. $\langle v_1, v_1\rangle \neq 0$. Let W be the span of v_1. Then by earlier theorem we have that $V = W \oplus W^{\perp}$. and so a basis for V can be obtained by combining v_1 with any basis $(v_2, ..., v_n)$ of $W^{\perp}$. using induction hypothesis, since dimension of $W^{\perp}$ i $n-1$, we can take $(v_2, ..., v_n)$ to be orthogonal. Then $(v_1, v_2, ..., v_n)$ is orthogonal basis of V. For, $\langle v_1, v_i\rangle = 0$ if $i > 1$ because $v_i \in W^{\perp}$, and $\langle v_i, v_j\rangle = 0$ if $i, j > 1$ and $i \neq j$, because $(v_2, ..., v_n)$ is an orthogonal basis. We normalize the basis so constructed by solving $c^{-2} = \pm\langle v_i, v_i\rangle$ and replacing v_i by cv_i. Then $\langle v_i, v_i\rangle$ is changed to ± 1. □

Remark 7.3.4. We can permute an orthogonal basis obtained in above theorem so that indices with $\langle v_i, v_i\rangle = 1$ are the first ones, and indices with $\langle v_i, v_i\rangle = -1$ will appear afterwards and those with $\langle v_i, v_i\rangle = 0$ will appear in the last. Then the matrix A of the form will be

$$\begin{pmatrix} I_p & & \\ & -I_m & \\ & & 0_z \end{pmatrix}$$

Theorem 7.3.6. Sylvester's Law *The numbers p, m, z appearing in above matrix are uniquely determined by the form. In other words they do not depend on the choice of orthogonal basis B such that $\langle v_i, v_i\rangle = \pm 1$* ***or*** *$0$.*

Theorem 7.3.7. *Let T be a normal operator on a finite dimensional complex inner product space V or a self adjoint operator on a finite dimensional real inner product space V. Let $\alpha_1, \alpha_2, ...\alpha_k$ be the distinct characteristic values of T. Let W_j be the characteristic space associated with α_j and E_j the orthogonal projection of V on W_j. Then W_j is orthogonal to W_i when $i \neq j$, V is the direct sum of $W_1, ..., W_K$, and*

$$T = \alpha_1 E_1 + ... + \alpha_k E_k$$

Proof. Let u be a vector in W_j and v be a vector in W_i, and suppose that $i \neq j$. Then $\alpha_j\langle u, v\rangle = \langle Tu, v\rangle = \langle u, T^*v\rangle = \langle u, \overline{\alpha_i}v\rangle$. Hence $(\alpha_j - \alpha_i)\langle u, v\rangle = 0$ and since $\alpha_j - \alpha_i \neq 0$ it follows that $\langle u, v\rangle = 0$. Thus W_j is orthogonal to W_i when $i \neq j$. From the fact that V has an orthonormal basis consisting of characteristic vectors it follows that

$V = W_1 + ... + W_k$. If $u_j \in W_j (j = 1, ...k)$ and $u_1 + ... + u_k = 0$, then

$$\begin{aligned} 0 &= \langle u_i, \sum_j u_j \rangle \\ &= \sum_j \langle u_i, u_j \rangle \\ &= ||u_i||^2 \end{aligned}$$

for every i, so that V is a direct sum of $W_1, ...W_k$. Therefore $E_1 + ...E_k = I$ and

$$\begin{aligned} T &= TE_1 + ... + TE_k \\ &= \alpha_1 E_1 + ... + \alpha_k E_K \end{aligned}$$

□

Such a decomposition of T is known as the **spectral resolution** of T.
Because $E_1, ...E_k$ are canonically associated with T and $I = E_1 + ... + E_k$ the family of projections $E_1, ...E_k$ is called the **resolution of the identity** defined by T.

7.3.1 Solved Problems

If you want give some problems for practice then write this:

Example 11. Let $\langle ., . \rangle$ be a bilinear form on $\mathbb{R}^2$ defined by

$$\langle (x_1, x_2), (y_1, y_2) \rangle = 2x_1y_1 - 3x_1y_2 + x_2y_2$$

1. Find the matrix A of this bilinear form in the basis $\{u_1 = (1, 0)$ and $u_2 = (1, 1)\}$
2. Find the matrix B of given bilinear form in the basis $\{v_1 = (2, 1)$ and $v_2 = (1, -1)\}$
3. Find the transition matrix P from the basis $\{u_i\}$ to $\{v_i\}$ and verify that $B = P^tAP$

Solution:

1. Set $A = (a_{ij})$ where $a_{ij} = \langle u_i, u_j \rangle$

$$a_{11} = \langle u_1, u_1 \rangle = \langle (1, 0), (0, 1) \rangle = 2 - 0 + 0 = 2$$

Rest of the entries in the matrix are calculated using following formulae

$$\begin{aligned} a_{12} &= \langle u_1, u_2 \rangle \\ a_{21} &= \langle u_2, u_1 \rangle \\ a_{22} &= \langle u_2, u_2 \rangle \end{aligned}$$

Thus the matrix A is as follows

$$A = \begin{pmatrix} 2 & -1 \\ 2 & 0 \end{pmatrix}$$

2. Similarly matrix B is

$$B = \begin{pmatrix} 3 & 9 \\ 0 & 6 \end{pmatrix}$$

3. Now we write $V - 1$ and v_2 in terms of u_1 and u_2.

$$\begin{aligned} (2, 1) &= u_1 + u_2 \\ (1, -1) &= 2u_1 - u_2 \end{aligned}$$

Thus $P = \begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix}$ and so $P^t = \begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix}$ Thus $P^tAP = P = \begin{pmatrix} 3 & 9 \\ 0 & 6 \end{pmatrix} = B$

Example 12. For the following real symmetric matrix A, find a non-singular matrix P such that P^tAP is diagonal and also find its signature .

$$A = \begin{pmatrix} 1 & -3 & 2 \\ -3 & 7 & -5 \\ 2 & -5 & 8 \end{pmatrix}$$

Solution:

First form the block matrix (A, I)

$$(A, I) = \left(\begin{array}{ccc|ccc} 1 & -3 & 2 & 1 & 0 & 0 \\ -3 & 7 & -5 & 0 & 1 & 0 \\ 2 & -5 & 8 & 0 & 0 & 1 \end{array}\right)$$

Apply the row operations $R_2 \to 3R_1 + R_2$ and $R_3 \to 2R_1 + R_3$ to (A, I) and then corresponding column operations $C_2 \to 3C_1 + C_2$ and $C_3 \to -2C_1 + C_3$ to A to obtain

[1] All solved problems are taken from Scaum's Outline Series-Theory and Problems of Linear Algebra by Lipschutz

$$\left(\begin{array}{ccc|ccc} 1 & -3 & 2 & 1 & 0 & 0 \\ 0 & -2 & 1 & 3 & 1 & 0 \\ 0 & 1 & 4 & -2 & 0 & 1 \end{array}\right) \text{ and then } \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & -2 & 0 & 3 & 1 & 0 \\ 0 & 1 & 4 & -2 & 0 & 1 \end{array}\right)$$

Next apply the row operation $R_3 \to R_2 + 2R_3$ and then corresponding column operation $C_3 \to C_2 + 2C_3$ to obtain

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & -2 & 1 & 3 & 1 & 0 \\ 0 & 0 & 9 & -1 & 1 & 2 \end{array}\right) \text{ and then } \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & -2 & 0 & 3 & 1 & 0 \\ 0 & 0 & 18 & -1 & 1 & 2 \end{array}\right)$$

Now A has been diagonalized. Set $P = \begin{pmatrix} 1 & -3 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{pmatrix}$; then

$P^tAP = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 18 \end{pmatrix}$

The signature S of A is $S = 2 - 1 = 1$.

Check Your Progress

1. Determine which of the following bilinear forms are symmetric/skewsymmetric/nondegenerate/ degenerate:

$$A = \begin{pmatrix} -1 & 0 & 1 \\ 0 & 0 & -1 \\ -1 & 1 & 2 \end{pmatrix}$$

$$A = \begin{pmatrix} -1 & 2 & -1 \\ 2 & 0 & -1 \\ -1 & -1 & 2 \end{pmatrix}$$

$$A = \begin{pmatrix} -1 & 0 & 1 \\ 0 & 0 & -1 \\ -1 & 1 & 2 \end{pmatrix}$$

$$A = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 1 & -1 & 0 & 2 \\ 0 & 0 & 0 & 1 \\ 1 & 2 & 1 & 5 \end{pmatrix}$$

7.4 Chapter End Exercise

1. Determine canonical form of the following real nondegenerate symmetric bilinear form

$$A = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 1 & 2 & 3 & 4 \\ 0 & 3 & 4 & 1 \\ 1 & 4 & 1 & 5 \end{pmatrix}$$

2. Let $\langle ., . \rangle$ be a bilinear form on $\mathbb{R}^2$ defined by

$$\langle (x_1, x_2), (y_1, y_2) \rangle = 3x_1y_1 - 2x_1y_2 + 4x_2y_1 - x_2y_2$$

   (a) Find the matrix A of this bilinear form in the basis $\{u_1 = (1, 1)$ and $u_2 = (1, 2)\}$

   (b) Find the matrix B of given bilinear form in the basis $\{v_1 = (1, -1)$ and $v_2 = (3, 1)\}$

   (c) Find the transition matrix P from the basis $\{u_i\}$ to $\{v_i\}$ and verify that $B = P^t AP$

3. Let V be a finite dimensional vector space over a field F and $\langle ., . \rangle$ is a symmetric bilinear form on V. For each subspace W of V, let $W^\perp$ be the set of all vectors $u \in V$ such that $\langle u, v \rangle = 0$for every $v \in W$. Show that

   (a) $W^\perp$ is a subspace of V.

   (b) $V = 0^\perp$

   (c) $V^\perp = 0$ if and only if $\langle \rangle$ is non degenerate

   (d) The restriction of $\langle ., . \rangle$ to W is nondegenerate if and only if $W \cap W^\perp = \{0\}$