



**Uttar Pradesh Rajarshi Tondon
open University**

Bachelor of Science

UGMM - 107

Linear Algebra

Linear Algebra

Block - 1 Vector Spaces	3
Unit - 1 Vector Spaces	5
Unit - 2 D' Alembert Principle	30
Unit - 3 Basis, Dimension and Quotient Spaces	56
Unit - 4 Dual Vector Space	91
Block - 2 Matrix	105
Unit - 5 Matrices	107
Unit - 6 Rank of a Matrix	129
Unit - 7 Determinants	150
Unit - 8 Eigen vector and eigen space	164
Block - 3 Characteristic polynomial Inner product Space Bilinear, quadratic forms	181
Unit - 9 Characteristic polynomial of a matrix	183
Unit - 10 Inner Products space	197
Unit - 11 Bilinear, Quadratic & Hermitian Forms	220

UGMM-107: Linear Algebra

- Unit 1:-** Binary operations, Field, Examples: $\mathbb{Q}, \mathbb{R}, \mathbb{C}, \mathbb{Z}_p$ (p, a prime) of fields. Definition of a vector space, Some basic properties of a vector space, Vector subspace, Subspace spanned by a subsets with examples. Linearly dependence and independence of a subset. Finite dimensional vector spaces,
- Unit 2:-** Basis of a vector space, Dimension of a vector space, finite dimensional vector space, linear sum and direct sum of subspaces, Quotient spaces.
- Unit 3:-** Linear transformations:- Definition examples and some properties, Linear transformations and Isomorphism of vector spaces, Null space and range space. Rank and Nullity of a linear transformations, Fundamental theorem of vector space homomorphism, Rank-Nullity theorem, Non singular, Invertible transformations.
- Unit 4:-** $\text{Hom}(V, W)$ as a vector space, dual space V^* of a vector space V , Dual basis of a vector space with examples, Transpose of a linear transformations. Annihilator of a subset of a vector space, Rank of transpose of a linear transformations.
- Unit 5:-** Matrix representation of a linear transformations, equality, Algebra of matrices, Multiplication of matrices. Vector space of all $m \times n$ matrices over a field. Ring of all n -square matrices. Invertible matrices (or Non-singular matrix), Transpose of matrix, Equivalent matrices, Similar Matrices, and Orthogonal matrices.
- Unit 6:-** Rank of a matrix, Row rank and column rank of a matrix, Elementary matrices, elementary row and column operations of a matrix, elementary matrices, Normal form of a matrix, Echelon form of matrix, Inverse of a non singular matrix. Determinantal rank of a matrix. Non homogeneous and homogeneous linear equations.
- Unit 7:-** Determinant of a square matrix, Cofactor of an element of a determinant, Properties of a determinant, minor of an element of a determinant, Evaluation of a determinant. Laplace expansion of a determinant. Product of two determinants(all statements without proof)
- Unit 8:-** Adjoint of a n - square matrix. Inverse of a matrix of a non-singular matrix by using adjoint of the matrix. Characteristic roots or Eigen values of a linear transformation and Eigen vector and Eigen space.
- Unit 9:-** Characteristic polynomial of a matrix, Diagonalization of a matrix, Caley - Hamilton theorem. Inverse of a matrix of a non singular matrix by Caley – Hamilton theorem. Characteristic polynomial of a linear transformation, Minimal polynomial, Hermitian Matrix, characteristic roots of a complex Hermitian matrix.
- Unit 10:-** Inner product space: Definition and examples of inner product, length of a vector, Cauchy Schwarz inequality, distance between two vectors, angle between two vectors, Orthogonal and orthonormal sets.
- Unit 11:-** Bilinear, quadratic and Hermitian forms, bilinear form on a vector space V . Quadratic forms, matrix of quadratic form, normal form or canonical form.



**Uttar Pradesh Rajarshi Tandon
Open University**

Bachelor of Science

UGMM - 107

Linear Algebra

BLOCK

1

Unit - 1	Vector Spaces	5
Unit - 2	D' Alembert Principle	30
Unit - 3	Basis, Dimension and Quotient Spaces	56
Unit - 4	Dual Vector Space	91

Course Design Committee

Prof. Ashutosh Gupta, School of Computer and Information Science, UPRTOU, Prayagraj	Chairman
Prof. Sudhir Srivastav Dept. of Mathematics, DDU Gorakhpur University, Gorakhpur	Member
Prof. P. K. Singh Dept. of Mathematics, Allahabad University, Prayagraj	Member
Prof. Mona Khare Dept. of Mathematics, Allahabad University, Prayagraj	Member
Dr. A. K. Pandey Associate Professor, ECC Prayagraj	Member
Dr. Vikas Singh Academic Consultant, UPRTOU, Prayagraj	Member
Dr. S. S. Tripathi Academic Consultant, UPRTOU, Prayagraj	Member

Course Preparation Committee

Dr. Kamran Alam Khan Assistant Professor (Dept. of Mathematics), Ram Lubhai Sahani Govt. Mahila Degree College, Pilibhit	Author (Unit – 1 - 3)
Dr. Ashok Kumar Pandey Associate Professor, E.C.C, Prayagraj	Author (Unit 4 - 9)
Dr. S. S. Tripathi Academic Consultant School of Science, UPRTOU, Prayagraj	Editor (Unit 1 - 9)
Prof. Ashutosh Gupta Director, School of Computer and Information Science, UPRTOU, Prayagraj	

© UPRTOU , Prayagraj - 2021**© UGMM - 107 Under Algebra****ISBN :****Reprint – 2021 – 100 Copy**

All Rights are reserved. No Part of this work may reproduced in any form, by mimeograph or any other means, without permission in writing from the Uttar Pradesh Rajarshi Tandon Open University.

Printed and Published by Pro. (Dr.) Arun Kumar Gupta Registrar, Uttar Pradesh Rajarshi Tandon Open University, 2021

Printed By. – M/s K.C.Printing & Allied Works, Panchwati, Mathura -281003.

Unit-1: VECTOR SPACES

Structure

- 1.1 Introduction**
- 1.2 Objectives**
- 1.3 Binary operations**
- 1.4 Fields**
- 1.5 Vector spaces**
- 1.6 Some basic properties of a vector space**
- 1.7 Illustrative examples**
- 1.8 Vector subspaces**
- 1.9 Linear Span and subspace spanned by a subset**
- 1.10 Linearly dependence and independence**
- 1.11 Illustrative examples**
- 1.12 Summary**
- 1.13 Terminal questions**
- 1.14 Further readings**

1.1 Introduction

In pure Mathematics we study abstract structures, forms and their properties. The algebraic structures such as groups, rings, fields and vector spaces are studied in abstract algebra. Linear algebra is that part of abstract algebra which deals with the systems of linear equations, matrices, determinants, vector spaces and linear transformations. Linear algebra finds applications almost everywhere- geometry, analysis, number theory, game theory, Markov chains, graph theory operations research, cryptography, computer science (e.g. data science, machine learning, computer graphics, image processing etc), physics (e.g. quantum mechanics, electrical networks etc), statistics, engineering, genetics, economics, sociology etc.

Let us have a look at the historical development of linear algebra. Leibnitz invented the notion of determinant in 1693 while working on linear systems of equations. Cramer presented the solution (known as Cramer's rule) for linear systems as ratios of determinants in 1750. The great mathematician Carl Friedrich Gauss developed his elimination method in 1811 without using the notation of matrices explicitly. He also worked with linear transformations by implicitly introducing matrices. James Joseph Sylvester contributed greatly to the theory of determinants and he coined the term "matrix" in 1848 which was the latin word for 'womb'. Arthur Cayley developed the matrix algebra in two papers (in 1850 and 1858) and defined matrix multiplication, inverse of a square matrix and proved the important Cayley-Hamilton theorem. The idea of extension of the notion of three-dimensional space to higher dimensional spaces was developed independently by Cayley, Hamilton and Grassmann. Hamilton developed an algebra for his system of quaternions. The mathematician Hermann Grassmann is credited for developing most of the core concepts of linear algebra such as linear dependence, spanning set, basis, dimension etc. He did not formalize the notion of a vector space; however this idea was implicit in his work *Die Ausdehnungslehre* (Theory of Extension, 1844). The modern definition of vector space that we use today was given by Peano in 1888.

In this unit, we shall study the basic concepts of linear algebra such as vector spaces, subspaces, the subspaces spanned by a subset and linear dependence and independence.

1.2 Objectives

After reading this unit, you should be able to:

- Understand the definition of a binary operation and a field

- Define a vector space and subspace
- Discuss various properties of a vector space
- Define linear span and subspace spanned by a subset
- Describe linearly dependent and independent sets

1.3 Binary operation

Let us consider the following observations-

- (1) The result of addition of two natural numbers is also a natural number.

For example, $5, 9 \in \mathbb{N}$ and $5 + 9 = 14 \in \mathbb{N}$

Therefore we have $a + b \in \mathbb{N}$ for all $a, b \in \mathbb{N}$

- (2) The product of any two real numbers is also a real number, i.e.

$$a \cdot b \in \mathbb{R} \text{ for all } a, b \in \mathbb{R}$$

- (3) Let M be a set of all 2×2 matrices having their elements as integers. Then the addition of any two such matrices produces the same kind of matrix.

For example, if $A, B \in M$ such that $A = \begin{bmatrix} 3 & -2 \\ 8 & 3 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & 5 \\ -4 & 2 \end{bmatrix}$, then

$$\text{that } A + B = \begin{bmatrix} 3 & -2 \\ 8 & 3 \end{bmatrix} + \begin{bmatrix} 0 & 5 \\ -4 & 2 \end{bmatrix} = \begin{bmatrix} 3 & 3 \\ 4 & 5 \end{bmatrix} \in M$$

Thus we see that sometimes for every pair of elements a, b of a given set S , the application of operations like addition and multiplication yields a unique element of the same set S . This operation is then called a binary operation. Let us give its formal definition-

Definition: Let S be a non-empty set. A **binary operation** $*$ on S is a function $*$: $S \times S \rightarrow S$ such that to each ordered pair $(a, b) \in S$ there exists a unique element $c \in S$ such that $*(a, b) = c$. We write $a * b$ for $*(a, b)$.

Thus a binary operation is a rule by which we can associate any two elements a, b of S with a unique element $c = a * b$ of S . A binary operation on a set S is also called a composition in S .

There are examples of binary operations other than usual addition and multiplication. For example

- (1) Suppose $\mathbb{Q}^+$ is the set of all positive rational numbers. For $a, b \in \mathbb{Q}^+$, define $*$ such that $a * b = ab/2$. Then you can verify yourself that $*$ is a binary operation on $\mathbb{Q}^+$.

- (2) Let $\wp(S)$ be the power set of any set S . Then the operations $\cup$ and $\cap$ are binary operations on $\wp(S)$, since

$$A \cup B \in \wp(S), \text{ and } A \cap B \in \wp(S) \quad \forall A, B \in \wp(S)$$

Definition: Let S be a non-empty set and $*$ be a binary operation on S . A subset H of S is said to be **closed** under the binary operation $*$, if $a * b \in H$ for all $a, b \in H$.

Definition: A binary operation $*$ on a non-empty set G is **associative** if

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

and is **commutative** if $a * b = b * a$ for all $a, b \in G$.

For example, the usual addition $+$ is associative as well as commutative on $\mathbb{Z}$, i.e. $a + (b + c) = (a + b) + c$ and $a + b = b + a$ for all $a, b, c \in \mathbb{Z}$. However subtraction is neither associative nor commutative on $\mathbb{Z}$.

i.e. $a - (b - c) \neq (a - b) - c$ and $a - b \neq b - a$ for $a, b, c \in \mathbb{Z}$.

A set with one or more operations (unary, binary or other) obeying a particular collection of axioms is termed as '**algebraic structure**'. For example, groupoids, semigroups, monoids, groups, rings, fields and vector spaces are all algebraic structures.

Our main objective in this unit is to explore the vector spaces and related concepts which are central to linear algebra. For that purpose we need some basic knowledge of the algebraic structures-groups, rings and fields. So let us begin with the definition of a group

Definition: A **group** is an ordered pair $(G, *)$ where G is a nonempty set and $*$ is a binary operation on G satisfying the following axioms:

1. $a * b \in G$ for all $a, b \in G$ (**Closure Property**)
2. $a * (b * c) = (a * b) * c$ for all $a, b, c \in G$ (**Associative Law**)
3. There exists an element $e \in G$ called an identity element such that
$$a * e = a = e * a \text{ for all } a \in G \quad (\text{Existence of Identity})$$
4. For each $a \in G$ there exists an element $b \in G$ called an inverse of ' a ' such that

$$a * b = e = b * a \text{ and we write } b = a^{-1} \quad (\text{Existence of Inverse})$$

If the set G is finite then the group $(G, *)$ is called a **finite group**, otherwise we call it an **infinite group**. The number of elements in a finite group is called the **order** of the group. An infinite group is said to be of infinite order.

The group whose only element is the identity e will be denoted by $\{e\}$. It is called the **trivial group**.

Definition: A group $(G, *)$ is said to be **abelian** (or commutative) if the binary operation $*$ is commutative, i.e. $a * b = b * a$ for all $a, b \in G$.

For example, the set of all integers $\mathbb{Z}$ forms an abelian group with respect to addition:

- (1) $a + b \in \mathbb{Z}$ for all $a, b \in \mathbb{Z}$.
- (2) $a + (b + c) = (a + b) + c$ for all $a, b, c \in \mathbb{Z}$.
- (3) We also notice that the set $\mathbb{Z}$ has a special element ' 0 ' such that

$$a + 0 = a = 0 + a \quad \forall a \in \mathbb{Z}$$

Hence 0 is the identity element.

(4) To each $a \in \mathbb{Z}$, there is an element $-a \in \mathbb{Z}$ such that

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

(5) $a + b = b + a$ for all $a, b \in \mathbb{Z}$

Therefore $(\mathbb{Z}, +)$ is an abelian group. You can check for yourself that $(\mathbb{Q}, +)$, $(\mathbb{R}, +)$ and $(\mathbb{C}, +)$ are all infinite abelian groups with $e = 0$ and $a^{-1} = -a$.

Now we are in position to define an important algebraic structure, called 'Field'.

1.4 Fields

Definition A field $(F, +, \cdot)$ is an algebraic structure consisting of a non-empty set F with at least two elements and two binary operations '+' and ' $\cdot$ ', called addition and multiplication respectively, such that the following axioms are satisfied:

(1) $a + b \in F$ for all $a, b \in F$

(2) $a + (b + c) = (a + b) + c$ for all $a, b, c \in F$

(3) $a + b = b + a$ for all $a, b \in F$

(4) There is an element $0 \in F$ called *zero* such that

$$a + 0 = a = 0 + a \text{ for all } a \in F.$$

(5) For each $a \in F$, there exists $-a \in F$ such that

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

(6) $a \cdot b \in F$ for all $a, b \in F$

(7) $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ for all $a, b, c \in F$

(8) $a \cdot b = b \cdot a$ for all $a, b \in F$

(9) There exists $1 \in F$ called *unity* such that

$$1 \cdot a = a = a \cdot 1 \text{ for all } a \in F$$

(10) Every nonzero element of F is invertible, i.e. for each $a(\neq 0) \in F$, there exists $a^{-1} \in F$ such that $a \cdot a^{-1} = 1 = a^{-1} \cdot a$

(11) For all $a, b, c \in F$, $a \cdot (b + c) = a \cdot b + a \cdot c$

$$\text{and } (a + b) \cdot c = a \cdot c + b \cdot c$$

Therefore, we can say that $(F, +, \cdot)$ is a field, if

(i) $(F, +)$ is an abelian group.

(ii) All the non-zero elements of F form an abelian group with respect to multiplication.

(iii) The distributive laws hold:

$$a \cdot (b + c) = a \cdot b + a \cdot c \text{ and } (a + b) \cdot c = a \cdot c + b \cdot c \text{ for all } a, b, c \in F$$

Instead of saying that $(F, +, \cdot)$ is a field, we simply say that F is a field with the understanding that there is no confusion regarding the binary compositions '+' and '·'. We denote $a + (-b)$ by $a - b$ and $a \cdot b$ by ab .

The fields of rational numbers, real numbers and complex numbers:

Each of $\mathbb{Q}$ (set of all rational numbers), $\mathbb{R}$ (set of all real numbers) and $\mathbb{C}$ (set of all complex numbers) forms a field under the usual operations of addition and multiplication, i.e. $(\mathbb{Q}, +, \cdot)$, $(\mathbb{R}, +, \cdot)$ and $(\mathbb{C}, +, \cdot)$ are fields. Here we shall discuss the field of complex numbers only. You can easily verify the field axioms for $\mathbb{Q}$ and $\mathbb{R}$.

We have

$$\mathbb{C} = \{z: z = a + ib, a, b \in \mathbb{R}\}$$

Let us define addition and multiplication of complex numbers as follows:

Let $z_1 = a_1 + ib_1$ and $z_2 = a_2 + ib_2$. Then

$$z_1 + z_2 = (a_1 + a_2) + i(b_1 + b_2)$$

$$\text{and } z_1 z_2 = (a_1 a_2 - b_1 b_2) + i(a_1 b_2 + a_2 b_1)$$

Then we make the following observations:

- (1) $z_1 + z_2 \in \mathbb{C}$ for all $z_1, z_2 \in \mathbb{C}$
- (2) $z_1 + (z_2 + z_3) = (z_1 + z_2) + z_3$ for all $z_1, z_2, z_3 \in \mathbb{C}$
- (3) $z_1 + z_2 = z_2 + z_1$ for all $z_1, z_2 \in \mathbb{C}$
- (4) There is an element $0 = 0 + i0 \in \mathbb{C}$ such that

$$z + 0 = z = 0 + z \text{ for all } z \in \mathbb{C}$$
- (5) For each $z = a + ib \in \mathbb{C}$, there exists $-z = (-a) + i(-b) \in \mathbb{C}$ such that

$$z + (-z) = 0 = (-z) + z$$
- (6) $z_1 z_2 \in \mathbb{C}$ for all $z_1, z_2 \in \mathbb{C}$
- (7) $z_1 (z_2 z_3) = (z_1 z_2) z_3$ for all $z_1, z_2, z_3 \in \mathbb{C}$
- (8) $z_1 z_2 = z_2 z_1$ for all $z_1, z_2 \in \mathbb{C}$
- (9) we have $1 = 1 + i0 \in \mathbb{C}$ such that

$$z1 = z = 1z \text{ for all } z \in \mathbb{C}$$
- (10) Let $z = a + ib \neq 0$. Then both a and b cannot be zero simultaneously. If $w = c + id$ is the multiplicative inverse of z , then

$$zw = 1$$

$$\Rightarrow (a + ib)(c + id) = 1$$

$$\Rightarrow (ac - bd) + i(ad + bd) = 1 + 0i$$

$$\Rightarrow ac - bd = 1, ad + bd = 0$$

Solving these equations, we get

$$c = \frac{a}{a^2 + b^2}, d = \frac{-b}{a^2 + b^2}$$

Since $a^2 + b^2 \neq 0$, hence one of c and d is a non-zero real number. Thus

$w = \frac{a-ib}{a^2+b^2} \in \mathbb{C}$ is the multiplicative inverse of $z = a + ib$.

- (11) For all $z_1, z_2, z_3 \in \mathbb{C}$, we have

$$z_1(z_2 + z_3) = z_1 z_2 + z_1 z_3$$

$$\text{and } (z_1 + z_2)z_3 = z_1z_3 + z_2z_3$$

Therefore $(\mathbb{C}, +, \cdot)$ is a field.

The field of integers modulo p (p , a prime number):

Let a and b be any two integers and n be a fixed integer, we define a composition on the set of integers $\mathbb{Z}$ called **addition modulo n** ($+_n$) as follows-

$$a+_nb = r, \quad 0 \leq r < n$$

Where r is the least non-negative remainder when $a + b$ is divided by n .

For example, $3+_54 = 2$, since 2 is the remainder (non-negative) when $3 + 4$ i.e. 7 is divided by 5. You can check for yourself that

- (i) $a+_nb \equiv a + b \pmod{n}$
- (ii) if $a \equiv b \pmod{n}$, then $a+_nc = b+_nc$.

Similarly we can define multiplication modulo n ($\times_n$). Let a and b be any two integers and n be a fixed integer. The composition on the set of integers $\mathbb{Z}$ called **multiplication modulo n** ($\times_n$) is defined as follows-

For any two integers a and b and a fixed integer n ,

$$a \times_n b = r, \quad 0 \leq r < n$$

Where r is the least non-negative remainder when ab is divided by n .

For example, $7 \times_3 4 = 1$, since 1 is the remainder (non-negative) when 7×4 i.e. 28 is divided by 3. Obviously we have $a \times_n b \equiv ab \pmod{n}$. Also if $a \equiv b \pmod{n}$, then $a \times_n c = b \times_n c$.

Let $\mathbb{Z}_p = \{0, 1, \dots, p-1\}$, where p is a prime number. Then we can show that $(\mathbb{Z}_p, +_p, \times_p)$ is a field called **the field of integers modulo p** .

(1) Let $a, b \in \mathbb{Z}_p$ then $a+_pb \equiv a + b \pmod{p} \in \mathbb{Z}_p$.

(2) Let $a, b, c \in \mathbb{Z}_p$. Then

$$\begin{aligned} a+_p(b+_pc) &= a+_p(b+c) \text{ as } b+_pc \equiv b+c \pmod{p} \\ &= [a+(b+c)] \pmod{p} \\ &= [(a+b)+c] \pmod{p} \\ &= (a+b)+_pc \\ &= (a+_pb)+_pc \end{aligned}$$

(3) There exists $0 \in \mathbb{Z}_p$ such that $a+_p0 = a = 0+_pa \quad \forall a \in \mathbb{Z}_p$

(4) Let $a \in \mathbb{Z}_p$. If $a = 0$, then it is the additive inverse of itself. Let $a \neq 0$. Then $p - a \in \mathbb{Z}_p$ is the additive inverse of $a \in \mathbb{Z}_p$ as

$$(p-a)+_pa = 0 = a+_p(p-a)$$

(5) Let $a, b \in \mathbb{Z}_p$. Then

$$a+_pb = a + b \pmod{p} = b+_pa$$

(6) Let $a, b \in \mathbb{Z}_p$. If any one of a and b are zero or both are zero, then $a \times_p b = 0 \in \mathbb{Z}_p$. Suppose $a \neq 0, b \neq 0$. Then p is neither a divisor of a nor a divisor of b . Therefore $p \nmid ab$. Hence the remainder cannot be zero when ab is divided by p i.e. if $a \times_p b = r$, then $1 \leq r \leq p-1$. Thus $a \times_p b \in \mathbb{Z}_p$

(7) For all $a, b, c \in \mathbb{Z}_p$, we have,

$$\begin{aligned} a \times_p (b \times_p c) &= a \times_p (bc) \text{, since } b \times_p c \equiv bc \pmod{p} \\ &= a(bc) \pmod{p} \\ &= (ab)c \pmod{p} \\ &= (ab) \times_p c \\ &= (a \times_p b) \times_p c \end{aligned}$$

(8) We have $1 \in \mathbb{Z}_p$ such that $1 \times_p a = a = a \times_p 1$ for all $a \in \mathbb{Z}_p$.

(9) If $0 \neq a \in \mathbb{Z}_p$, then by closure property $a \times_p i \in \mathbb{Z}_p$ for all $i \in \mathbb{Z}_p$. Also $i \neq j \Rightarrow a \times_p i \neq a \times_p j$ for $i, j \in \mathbb{Z}_p$, for if $a \times_p i = a \times_p j$, then $ai - aj$ is divisible by p , i.e. $p | a(i - j)$. Which is not possible, since p cannot divide $(i - j)$ as either $i - j < p$ for $i > j$, or $j - i < p$ for $j > i$. Also $p \nmid a$. Hence all the products $a \times_p i \in \mathbb{Z}_p$, $i \in \mathbb{Z}_p$ are distinct elements of $\mathbb{Z}_p$. Therefore there exists $i = k \in \mathbb{Z}_p$ such that $a \times_p k = 1$, i.e. k is the multiplicative inverse of a .

(10) For all $a, b \in \mathbb{Z}_p$ we have,

$$\begin{aligned} a \times_p b &= ab \pmod{p} \\ &= ba \pmod{p} \\ &= b \times_p a \end{aligned}$$

(11) Let $a, b, c \in \mathbb{Z}_p$. Then

$$\begin{aligned} a \times_p (b +_p c) &= a \times_p (b + c) \text{ as } b +_p c \equiv b + c \pmod{p} \\ &= a(b + c) \pmod{p} \\ &= ab + ac \pmod{p} \\ &= ab +_p ac \\ &= a \times_p b +_p a \times_p c \end{aligned}$$

Similarly, you can show that $(a +_p b) \times_p c = a \times_p c +_p b \times_p c$

Therefore $(\mathbb{Z}_p, +_p, \times_p)$ is a field. This is an example of a finite field.

1.5 Vector Spaces

In your previous classes, you were introduced to vectors in three dimensions and you are also familiar with the following properties:

(1) Let V be the set of all vectors in three dimensional Euclidean Space. Then the vectors in V can be added and V forms an abelian group with respect to this addition $+$ (say):

(i) If $\vec{a}$ and $\vec{b}$ are any two vectors in three dimensional Euclidean space

$\mathbb{R}^3$, then $\vec{a} + \vec{b}$ is also a vector in $\mathbb{R}^3$, i.e. $\vec{a} + \vec{b} \in V$ for all $\vec{a}, \vec{b} \in V$

(ii) $\vec{a} + (\vec{b} + \vec{c}) = (\vec{a} + \vec{b}) + \vec{c}$ for all $\vec{a}, \vec{b}, \vec{c} \in V$

(iii) There exists null vector $\vec{0}$ such that $\vec{a} + \vec{0} = \vec{a} = \vec{0} + \vec{a}$ for all $\vec{a} \in V$

(iv) For each $\vec{a} \in V$, there exists $-\vec{a} \in V$ such that

$$\vec{a} + (-\vec{a}) = \vec{0} = (-\vec{a}) + \vec{a}$$

$$(v) \vec{a} + \vec{b} = \vec{b} + \vec{a} \text{ for all } \vec{a}, \vec{b} \in V$$

(2) The multiplication of a vector by any real number produces another vector, i.e. if $\lambda \in \mathbb{R}$, then $\lambda\vec{a}$ is also a vector in $\mathbb{R}^3$. For example, $2\vec{a}$, $(5/9)\vec{a}$, ... etc are all vectors. Hence $\lambda \in \mathbb{R}, \vec{a} \in V \Rightarrow \lambda\vec{a} \in V$

(3) For all $\vec{a}, \vec{b} \in V$ and $\mu \in \mathbb{R}$, we have

$$(i) \lambda(\vec{a} + \vec{b}) = \lambda\vec{a} + \lambda\vec{b}$$

$$(ii) (\lambda + \mu)\vec{a} = \lambda\vec{a} + \mu\vec{a}$$

$$(iii) (\lambda\mu)\vec{a} = \lambda(\mu\vec{a})$$

$$(iv) 1\vec{a} = \vec{a}$$

These properties are not restricted to vectors in three dimensional Euclidean space only. There are many mathematical objects which follow these properties. For example, the complex numbers, the polynomials over a field F , convergent real sequences and many others follow similar properties. So we can generalize this concept to obtain an abstract structure called a 'vector space'. We have the following definition:

Definition Let V be a non-empty set, whose elements are referred to as vectors and $(F, +, \cdot)$ be a field, whose elements are referred to as scalars. Then V is said to be a *vector space over the field F* if the following properties are satisfied :

1. There exists a binary operation $\oplus: V \times V \rightarrow V$, defined by $(\alpha, \beta) \mapsto \alpha \oplus \beta$, called vector addition such that $(V, \oplus)$ is an abelian group, that is,

$$(i) \alpha \oplus \beta \in V \text{ for all } \alpha, \beta \in V;$$

$$(ii) \alpha \oplus (\beta \oplus \gamma) = (\alpha \oplus \beta) \oplus \gamma \text{ for all } \alpha, \beta, \gamma \in V;$$

$$(iii) \alpha \oplus \beta = \beta \oplus \alpha \text{ for all } \alpha, \beta \in V;$$

$$(iv) \text{ There is a unique vector } \mathbf{0} \in V, \text{ called the zero vector, such that}$$

$$\alpha \oplus \mathbf{0} = \alpha = \mathbf{0} \oplus \alpha \text{ for all } \alpha \in V;$$

$$(v) \text{ For each vector } \alpha \in V \text{ there is a unique vector } -\alpha \in V \text{ such that}$$

$$\alpha \oplus (-\alpha) = \mathbf{0} = (-\alpha) \oplus \alpha$$

2. There exists an operation $\odot: F \times V \rightarrow V$, defined by $(a, \alpha) \mapsto a \odot \alpha$ called scalar multiplication, that is, $a \in F, \alpha \in V \Rightarrow a \odot \alpha \in V$

3. For all vectors $\alpha, \beta \in V$ and for all scalars $a, b \in F$, we have

$$(i) 1 \odot \alpha = \alpha$$

$$(ii) (ab) \odot \alpha = a \odot (b \odot \alpha)$$

$$(iii) a \odot (\alpha \oplus \beta) = (a \odot \alpha) \oplus (a \odot \beta)$$

$$(iv) (a + b) \odot \alpha = (a \odot \alpha) \oplus (b \odot \alpha)$$

The word 'vector' should not be confused with the concept of 'vector as a directed segment' which you have studied in geometry and mechanics. Here we use it in a much broader sense.

Here it must be noted that the addition in ' $a + b$ ' is the addition of scalars, i.e. the addition in the field F . Similarly the multiplication in ab represents the multiplication in F .

Note: For the sake of convenience, in future, we shall use the notation ' $\alpha + \beta$ ' for ' $\alpha \oplus \beta$ ' and the notation ' $a\alpha$ ' for ' $\alpha \odot a$ '. With these simplified notations, the definition of a vector space may be given as follows-

A vector space consists of a non-empty set V and a field F such that the following properties are satisfied:

(1) There exists a binary operation $+$ on V , called vector addition such that $(V, +)$ is an abelian group, i.e.

$$(i) \alpha + \beta \in V \text{ for all } \alpha, \beta \in V;$$

$$(ii) \alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma \text{ for all } \alpha, \beta, \gamma \in V;$$

$$(iii) \alpha + \beta = \beta + \alpha \text{ for all } \alpha, \beta \in V;$$

(iv) There is a unique vector $0 \in V$, called the zero vector, such that

$$\alpha + 0 = \alpha = 0 + \alpha \text{ for all } \alpha \in V;$$

(v) For each vector $\alpha \in V$ there is a unique vector $-\alpha \in V$ such that

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

2. There exists an operation called scalar multiplication which associates each $a \in F$ and $\alpha \in V$ to a unique element $a\alpha$ in V , i.e.

$$a \in F, \alpha \in V \Rightarrow a\alpha \in V$$

3. For all vectors $\alpha, \beta \in V$ and for all scalars $a, b \in F$, we have

$$(i) 1\alpha = \alpha$$

$$(ii) (ab)\alpha = a(b\alpha)$$

$$(iii) a(\alpha + \beta) = a\alpha + a\beta$$

$$(iv) (a + b)\alpha = a\alpha + b\alpha$$

1.6 Some basic properties of a vector space

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

$$(1) 0\alpha = 0 \text{ for all } \alpha \in V$$

$$(2) a0 = 0 \text{ for all } a \in F$$

$$(3) (-1)\alpha = -\alpha \text{ for all } \alpha \in V$$

$$(4) (-a)\alpha = a(-\alpha) = -(a\alpha) \text{ for all } a \in F, \alpha \in V$$

$$(5) \text{ If } a\alpha = 0, \text{ then either } a = 0 \text{ or } \alpha = 0$$

Proof: (1) $0 \in F, \alpha \in V \Rightarrow 0\alpha \in V$

$$\text{Hence } 0 + 0\alpha = 0\alpha$$

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

$$\Rightarrow 0 + 0\alpha = 0\alpha + 0\alpha$$

$$\Rightarrow 0 = 0\alpha \text{ , by right cancellation law in the group } (V, +)$$

(2) we have $a0 \in V$, hence

$$0 + a0 = a0$$

$$\Rightarrow 0 + a0 = a(0 + 0)$$

$$\Rightarrow 0 + a0 = a0 + a0$$

$$\Rightarrow 0 = a0 \text{ , by right cancellation law in the group } (V, +)$$

(3) From (1), we have

$$0\alpha = 0 \text{ for all } \alpha \in V$$

$$\begin{aligned}
&\Rightarrow [1 + (-1)]\alpha = 0 \text{ for all } \alpha \in V \\
&\Rightarrow 1\alpha + (-1)\alpha = 0 \text{ for all } \alpha \in V \\
&\Rightarrow \alpha + (-1)\alpha = 0 \text{ for all } \alpha \in V \\
&\Rightarrow (-1)\alpha = -\alpha \text{ for all } \alpha \in V
\end{aligned}$$

(4) we have

$$\begin{aligned}
&0\alpha = 0 \text{ for all } \alpha \in V \\
&\Rightarrow [a + (-a)]\alpha = 0 \text{ for all } a \in F, \alpha \in V \\
&\Rightarrow a\alpha + (-a)\alpha = 0 \text{ for all } a \in F, \alpha \in V \\
&\Rightarrow (-a)\alpha = -(a\alpha) \text{ for all } a \in F, \alpha \in V
\end{aligned}$$

Also $a0 = 0 \Rightarrow a[\alpha + (-\alpha)] = 0$ for all $a \in F, \alpha \in V$

$$\Rightarrow a\alpha + a(-\alpha) = 0 \text{ for all } a \in F, \alpha \in V$$

$$\Rightarrow a(-\alpha) = -(a\alpha) \text{ for all } a \in F, \alpha \in V$$

Therefore we have $(-a)\alpha = a(-\alpha) = -(a\alpha)$ for all $a \in F, \alpha \in V$

(5) We have $a\alpha = 0$. Suppose $a \neq 0$. Then there exists $a^{-1} \in F$ such that

$$\begin{aligned}
&aa^{-1} = 1 = a^{-1}a \\
&a\alpha = 0 \Rightarrow a^{-1}(a\alpha) = a^{-1}0 \\
&\Rightarrow (a^{-1}a)\alpha = 0 \\
&\Rightarrow 1\alpha = 0 \\
&\Rightarrow \alpha = 0
\end{aligned}$$

Hence if $a\alpha = 0$, then either $a = 0$ or $\alpha = 0$.

1.7 Illustrative examples

Example 1.7.1 (The n -tuple space or Affine n -space over F) Let F be a field, and let

$$V = F^n = \{(x_1, x_2, \dots, x_n) : x_1, x_2, \dots, x_n \in F\}$$

Where n is a fixed integer ≥ 1 . Two elements $(x_1, x_2, \dots, x_n)$ and $(y_1, y_2, \dots, y_n)$ of V are said to be equal if and only if $x_i = y_i$ for each $i = 1, 2, \dots, n$. Then V is a vector space over F under vector addition and scalar multiplication given by

$$\begin{aligned}
(x_1, x_2, \dots, x_n) + (y_1, y_2, \dots, y_n) &= (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n) \\
a(x_1, x_2, \dots, x_n) &= (ax_1, ax_2, \dots, ax_n)
\end{aligned}$$

Let $\alpha = (x_1, x_2, \dots, x_n)$, $\beta = (y_1, y_2, \dots, y_n)$, and $\gamma = (z_1, z_2, \dots, z_n)$. Then

(1) $\alpha + \beta = (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n) \in V$ as $x_i + y_i \in F \forall x_i, y_i \in F$.

(2) $(\alpha + \beta) + \gamma = (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n) + (z_1, z_2, \dots, z_n)$
 $= ((x_1 + y_1) + z_1, (x_2 + y_2) + z_2, \dots, (x_n + y_n) + z_n)$

Since addition is associative in F , hence $(x_i + y_i) + z_i = x_i + (y_i + z_i)$ for all $i = 1, 2, \dots, n$. Therefore

$$\begin{aligned}
(\alpha + \beta) + \gamma &= (x_1 + (y_1 + z_1), x_2 + (y_2 + z_2), \dots, x_n + (y_n + z_n)) \\
&= (x_1, x_2, \dots, x_n) + (y_1 + z_1, y_2 + z_2, \dots, y_n + z_n) \\
&= \alpha + (\beta + \gamma)
\end{aligned}$$

(3) Since $x_i + y_i = y_i + x_i$ for all $i = 1, 2, \dots, n$, hence

$$\begin{aligned}\alpha + \beta &= (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n) \\ &= (y_1 + x_1, y_2 + x_2, \dots, y_n + x_n) \\ &= \beta + \alpha\end{aligned}$$

(4) The zero vector here is the n -tuple $\mathbf{0} = (0, 0, \dots, 0)$ as

$$\begin{aligned}\alpha + \mathbf{0} &= (x_1, x_2, \dots, x_n) + (0, 0, \dots, 0) \\ &= (x_1 + 0, x_2 + 0, \dots, x_n + 0) \\ &= (x_1, x_2, \dots, x_n) \\ &= \alpha\end{aligned}$$

Similarly, $\mathbf{0} + \alpha = \alpha$

(5) The additive inverse of $\alpha = (x_1, x_2, \dots, x_n)$ is the vector $-\alpha = (-x_1, -x_2, \dots, -x_n)$ as

$$\begin{aligned}(-\alpha) + \alpha &= (-x_1, -x_2, \dots, -x_n) + (x_1, x_2, \dots, x_n) \\ &= (-x_1 + x_1, -x_2 + x_2, \dots, -x_n + x_n) \\ &= (0, 0, \dots, 0) = \mathbf{0}\end{aligned}$$

Similarly, $\alpha + (-\alpha) = \mathbf{0}$.

Thus $(V, +)$ is an abelian group.

(6) We have $a\alpha = a(x_1, x_2, \dots, x_n) = (ax_1, ax_2, \dots, ax_n) \in V$, as $ax_1, ax_2, \dots, ax_n$ are all elements of F .

(7) For all vectors $\alpha, \beta \in V$ and for all scalars $a, b \in F$, we have

$$(i) \ 1\alpha = 1(x_1, x_2, \dots, x_n) = (1x_1, 1x_2, \dots, 1x_n) = (x_1, x_2, \dots, x_n) = \alpha$$

$$\begin{aligned}(ii) \ (ab)\alpha &= ((ab)x_1, (ab)x_2, \dots, (ab)x_n) \\ &= (a(bx_1), a(bx_2), \dots, a(bx_n)) \\ &= a(bx_1, bx_2, \dots, bx_n) = a(b\alpha)\end{aligned}$$

$$\begin{aligned}(iii) \ a(\alpha + \beta) &= a(x_1 + y_1, x_2 + y_2, \dots, x_n + y_n) \\ &= (a(x_1 + y_1), a(x_2 + y_2), \dots, a(x_n + y_n)) \\ &= (ax_1 + ay_1, ax_2 + ay_2, \dots, ax_n + ay_n) \\ &= (ax_1, ax_2, \dots, ax_n) + (ay_1, ay_2, \dots, ay_n) \\ &= a\alpha + a\beta\end{aligned}$$

$$\begin{aligned}(iv) \ (a + b)\alpha &= (a + b)(x_1, x_2, \dots, x_n) \\ &= ((a + b)x_1, (a + b)x_2, \dots, (a + b)x_n) \\ &= (ax_1 + bx_1, ax_2 + bx_2, \dots, ax_n + bx_n) \\ &= (ax_1, ax_2, \dots, ax_n) + (bx_1, bx_2, \dots, bx_n) \\ &= a\alpha + b\alpha\end{aligned}$$

Thus $V(= F^n)$ is a vector space over the field F . When $F = \mathbb{R}$, the vector space $\mathbb{R}^n$ is called **Euclidean n -space**.

Example 1.7.2 (The space of functions) Let F be any field and let X be any non-empty set. Let V be the set of all the functions from X into F . For $f, g \in V$, define the sum $f + g$ by

$$(f + g)(x) = f(x) + g(x) \quad \forall x \in X$$

and for $c \in F$, define cf by

$$(cf)(x) = cf(x) \quad \forall x \in X$$

Then you can check that V satisfies all the properties of a vector space over F . We shall denote this vector space by $\mathcal{F}_X(F)$.

Example 1.7.3 (The space of matrices) Let F be any field and $M_{m \times n}(F)$ be the set of all $m \times n$ matrices with entries from F . Let $c \in F$ and $A, B \in M_{m \times n}(F)$ such that $A = [a_{ij}]$ and $B = [b_{ij}]$. Define

$$A + B = [a_{ij}] + [b_{ij}] = [a_{ij} + b_{ij}]$$

$$\text{and} \quad cA = c[a_{ij}] = [ca_{ij}]$$

Then $M_{m \times n}(F)$ is a vector space over the field F .

Example 1.7.4 (The space of polynomial functions) Let F be any field and let V be the set of all polynomial functions in one variable with coefficients belonging to F . Let $f, g \in V$ such that

$$f(x) = a_0 + a_1x + \dots + a_mx^m$$

$$\text{and} \quad g(x) = b_0 + b_1x + \dots + b_nx^n$$

where $a_i, b_j \in F$ for all i, j . Define the sum $f + g$ by

$$(f + g)(x) = f(x) + g(x) = \sum_k (a_k + b_k)x^k$$

Where $a_k = 0$ if $k > m$ and $b_k = 0$ if $k > n$. Also defines cf by

$$(cf)(x) = cf(x) = c \sum_{i=0}^m a_i x^i = \sum_{i=0}^m ca_i x^i$$

Then V is a vector space over the field F .

Example 1.7.5 (The space of real sequences) Let V be the set of all real sequences, i.e. $V = \{\langle x_n \rangle : x_n \in \mathbb{R}\}$. Then V is a vector space over $\mathbb{R}$ under the following operations

$$\langle x_n \rangle + \langle y_n \rangle = \langle x_n + y_n \rangle$$

$$a\langle x_n \rangle = \langle ax_n \rangle$$

1.8 Vector subspaces

Definition: Let V be a vector space over the field F . A **vector subspace** W of V is any non-empty subset $W \subset V$ which is itself a vector space over F with the operations of vector addition and scalar multiplication on V .

Example 1.8.1 (a) If V is any vector space, then V is a subspace of V itself.

(b) The subset $\{0\}$ of a vector space V consisting of the zero vector of V is a subspace of V called the **zero subspace** of V .

How will you check whether a given subset W of a vector space V is a subspace of V ? The answer is the following criterion-

Theorem 1 A non-empty subset W of a vector space V is a subspace of V if and only if for each pair of vectors α, β in W and each pair of scalars a, b in F , the vector $a\alpha + b\beta$ is again in W , i.e.

$$a, b \in F \text{ and } \alpha, \beta \in W \Rightarrow a\alpha + b\beta \in W$$

Proof: *If part* First suppose that W is a non-empty subset of a vector space V such that $a\alpha + b\beta \in W$ for all $a, b \in F$ and for all $\alpha, \beta \in W$.

Since $W \neq \emptyset$, there is a vector $\alpha \in W$, taking $a = -1$ and $b = 1$, we have

$$(-1)\alpha + \alpha \in W, \text{ i.e. } 0 \in W.$$

Now let $\alpha \in W$ and $a \in F$. Taking $\beta = 0$ in the given condition, we have

$$a\alpha + 0 \in W \text{ or } a\alpha \in W$$

In particular, if we take $a = -1$, we have $(-1)\alpha \in W$, i.e. $-\alpha \in W$.

Also if $\alpha, \beta \in W$, taking $a = 1, b = 1$ in the given condition, we have

$$1\alpha + 1\beta \in W \text{ or } \alpha + \beta \in W$$

Since $W \subset V$, the commutativity and associativity of vector addition, and the properties of scalar multiplication will hold in W . Thus the remaining properties of vector space also hold in W , i.e. W is a subspace of V .

Only if part Suppose W is a subspace of V . Let $a, b \in F$ and $\alpha, \beta \in W$. Then

$$a \in F, \alpha \in W \Rightarrow a\alpha \in W$$

$$\text{and } b \in F, \beta \in W \Rightarrow b\beta \in W$$

Now $a\alpha \in W, b\beta \in W \Rightarrow a\alpha + b\beta \in W$.

Example 1.8.2 Euclidean 3-space $\mathbb{R}^3$ is a vector space over the field $\mathbb{R}$. Then $W = \{(x, y, 0) : x, y \in \mathbb{R}\}$ is a vector subspace of $\mathbb{R}^3$.

Let $\alpha, \beta \in W$ such that $\alpha = (x_1, y_1, 0)$ and $\beta = (x_2, y_2, 0)$. Let $a, b \in \mathbb{R}$.

$$\begin{aligned} \text{Now } a\alpha + b\beta &= a(x_1, y_1, 0) + b(x_2, y_2, 0) \\ &= (ax_1, ay_1, 0) + (bx_2, by_2, 0) \\ &= (ax_1 + bx_2, ay_1 + by_2, 0) \end{aligned}$$

Since $ax_1 + bx_2 \in \mathbb{R}, ay_1 + by_2 \in \mathbb{R}$, hence $(ax_1 + bx_2, ay_1 + by_2, 0) \in W$.

Thus $a, b \in F$ and $\alpha, \beta \in W \Rightarrow a\alpha + b\beta \in W$. Therefore W is a subspace of $\mathbb{R}^3$.

Theorem 2 The intersection of two subspaces of a vector space V is also a subspace of V .

Proof: Let W_1 and W_2 be any two subspaces of a vector space V . Then $0 \in W_1$ and $0 \in W_2$. Therefore $0 \in W_1 \cap W_2$. Hence $W_1 \cap W_2$ is a non-empty subset of V . Let $a, b \in F$ and $\alpha, \beta \in W_1 \cap W_2$. Then $\alpha, \beta \in W_1$ and $\alpha, \beta \in W_2$. Since W_1 and W_2 are subspaces of V , hence

$$a, b \in F \text{ and } \alpha, \beta \in W_1 \Rightarrow a\alpha + b\beta \in W_1$$

$$\text{and } a, b \in F \text{ and } \alpha, \beta \in W_2 \Rightarrow a\alpha + b\beta \in W_2$$

Now $a\alpha + b\beta \in W_1, a\alpha + b\beta \in W_2 \Rightarrow a\alpha + b\beta \in W_1 \cap W_2$.

Therefore $a, b \in F$ and $\alpha, \beta \in W_1 \cap W_2 \Rightarrow a\alpha + b\beta \in W_1 \cap W_2$

However the union of two subspaces may not be a subspace of vector space. For example, consider two subspaces $W_1 = \{(x, 0, 0) : x \in \mathbb{R}\}$ and $W_2 = \{(0, y, 0) : y \in \mathbb{R}\}$ of $\mathbb{R}^3$. Suppose $\alpha = (1, 0, 0) \in W_1$ and $\beta = (0, 3, 0) \in W_2$. Then $\alpha, \beta \in W_1 \cup W_2$. Now $\alpha + \beta = (1, 0, 0) + (0, 3, 0) = (1, 3, 0) \notin W_1 \cup W_2$. Hence $W_1 \cup W_2$ is not a subspace of $\mathbb{R}^3$. The next theorem tells you the situation in which $W_1 \cup W_2$ is a subspace.

Theorem 3 The union of two subspaces of a vector space V is a subspace if and only if one is contained in the other.

Proof: Let W_1 and W_2 be any two subspaces of a vector space V . First suppose that either $W_1 \subset W_2$ or $W_2 \subset W_1$. Then $W_1 \cup W_2 = W_2$ or W_1 . In either case, $W_1 \cup W_2$ is a subspace of V .

Conversely, suppose that $W_1 \cup W_2$ is a subspace of V . To prove that either $W_1 \subset W_2$ or $W_2 \subset W_1$, suppose on the contrary that $W_1 \not\subset W_2$ and $W_2 \not\subset W_1$.

$$W_1 \not\subset W_2 \Rightarrow \exists \alpha \in W_1 \text{ such that } \alpha \notin W_2$$

$$W_2 \not\subset W_1 \Rightarrow \exists \beta \in W_2 \text{ such that } \beta \notin W_1$$

$$\text{Now } \alpha \in W_1, \beta \in W_2 \Rightarrow \alpha, \beta \in W_1 \cup W_2$$

$$\Rightarrow \alpha + \beta \in W_1 \cup W_2, \text{ since } W_1 \cup W_2 \text{ is a subspace}$$

$$\Rightarrow \alpha + \beta \in W_1 \text{ or } \alpha + \beta \in W_2$$

Since $\alpha \in W_1$ and W_1 is a subspace, hence if $\alpha + \beta \in W_1$, then $(\alpha + \beta) - \alpha \in W_1$, i.e. $\beta \in W_1$. Hence we get a contradiction as $\beta \notin W_1$.

Also $\beta \in W_2$ and W_2 is a subspace, hence if $\alpha + \beta \in W_2$, then $(\alpha + \beta) - \beta \in W_2$, i.e. $\alpha \in W_2$. Again we get a contradiction as $\alpha \notin W_2$.

Hence either $W_1 \subset W_2$ or $W_2 \subset W_1$.

1.9 Linear Span and Subspace spanned by a subset

Definition: Let V be a vector space over a field F . A vector α in V is said to be a **finite linear combination** of vectors $\alpha_1, \alpha_2, \dots, \alpha_n$ in V if there exist scalars $c_1, c_2, \dots, c_n \in F$ such that

$$\alpha = c_1\alpha_1 + c_2\alpha_2 + \dots + c_n\alpha_n = \sum_{i=1}^n c_i\alpha_i$$

Definition Let V be a vector space over a field F and let S be any non-empty subset of V . The set $L(S)$ of all finite linear combinations of elements of S is called the **linear span** of S , i.e.

$$L(S) = \left\{ \sum_{i=1}^n a_i\alpha_i : a_1, a_2, \dots, a_n \in F, \alpha_1, \alpha_2, \dots, \alpha_n \in S \right\}$$

Obviously the elements of $L(S)$ are all vectors. Now we shall show that $L(S)$ is a subspace of V .

Theorem 4 The linear span $L(S)$ is a subspace of the vector space V .

Proof: Let $a, b \in F$ and $\alpha, \beta \in L(S)$ such that $\alpha = \sum_{i=1}^n a_i \alpha_i$ and $\beta = \sum_{j=1}^m b_j \beta_j$. Where $a_1, \dots, a_n; b_1, \dots, b_m \in F, \alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_m \in S$. We have

$$\begin{aligned} a\alpha + b\beta &= a \sum_{i=1}^n a_i \alpha_i + b \sum_{j=1}^m b_j \beta_j \\ &= \sum_{i=1}^n a(a_i \alpha_i) + \sum_{j=1}^m b(b_j \beta_j) \\ &= \sum_{i=1}^n (aa_i) \alpha_i + \sum_{j=1}^m (bb_j) \beta_j \\ &= \text{a finite linear combination of } \alpha\text{'s and } \beta\text{'s} \\ &\Rightarrow a\alpha + b\beta \in L(S) \end{aligned}$$

Thus $L(S)$ is a subspace of V .

Subspace spanned by a subset

Suppose V is a vector space over a field F and $S \subset V$. A vector subspace W of V is said to be spanned or generated by S if W is the smallest subspace of V containing S . In such a case, we write $W = \langle S \rangle$.

Theorem 5 Let V be a vector space over a field F and $S \subset V$. Then $L(S) = \langle S \rangle$.

Proof: Let $\alpha \in S$. Then we can write $\alpha = 1\alpha$. Therefore $\alpha \in L(S)$. Hence $S \subset L(S)$. Now suppose that W is any subspace of V containing S . Since W is closed under vector addition and scalar multiplication, each linear combination of elements of S must belong to W , i.e. $L(S) \subset W$. Hence every subspace of V containing S also contains $L(S)$. Thus $L(S)$ is the smallest subspace of V containing S , i.e. $L(S) = \langle S \rangle$.

1.10 Linearly dependence and independence

Definition Let V be a vector space over a field F . A subset $\{\alpha_1, \dots, \alpha_n\}$ of V is said to be **linearly dependent** if there exist scalars $a_1, \dots, a_n$, not all of which are 0, such that

$$a_1 \alpha_1 + a_2 \alpha_2 + \dots + a_n \alpha_n = 0$$

The set $\{\alpha_1, \dots, \alpha_n\}$ of vectors in V is said to be **linearly independent** if it is not linearly dependent, that is,

$$\sum_{i=1}^n a_i \alpha_i = 0 \Rightarrow a_i = 0 \text{ for all } 1 \leq i \leq n$$

The terminology used here is quite common and makes sense. We usually use the phrase “ α depends (linearly) on β ” to describe the relation “ $\alpha = c\beta$ ”. So if the vector α is linearly dependent on vectors $\alpha_1, \dots, \alpha_n$, there must exist scalars $a_1, \dots, a_n$ such that

$$\begin{aligned}\alpha &= a_1 \alpha_1 + a_2 \alpha_2 + \dots + a_n \alpha_n \\ \text{or } (-1)\alpha + a_1 \alpha_1 + a_2 \alpha_2 + \dots + a_n \alpha_n &= 0\end{aligned}$$

In light of above definition, we can say that the set $\alpha_1, \dots, \alpha_n, \alpha$ is linearly dependent.

For example, consider the vectors $(3, -5, -8), (9, -2, 1), (2, 1, 3)$ in $\mathbb{R}^3$. We see that

$$(3, -5, -8) = 1(9, -2, 1) + (-3)(2, 1, 3)$$

Hence we can say that $(3, -5, -8)$ is linearly dependent on $(9, -2, 1)$ and $(2, 1, 3)$. Also since

$$\begin{aligned}(9, -2, 1) &= 1(3, -5, -8) + 3(2, 1, 3) \\ \text{and } (2, 1, 3) &= \left(\frac{1}{3}\right)(9, -2, 1) + \left(-\frac{1}{3}\right)(3, -5, -8)\end{aligned}$$

Therefore the vector $(9, -2, 1)$ is linearly dependent on $(3, -5, -8)$ and $(2, 1, 3)$. Also the vector $(2, 1, 3)$ is linearly dependent on $(9, -2, 1)$ and $(3, -5, -8)$. We can also write

$$1(3, -5, -8) + (-1)(9, -2, 1) + 3(2, 1, 3) = (0, 0, 0) = 0$$

Thus the vectors $(3, -5, -8), (9, -2, 1), (2, 1, 3)$ are linearly dependent.

What about the set $\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ of vectors in $\mathbb{R}^3$? Is it possible to express anyone of them as a linear combination of remaining two vectors?

Suppose there exist $a_1, a_2, a_3 \in \mathbb{R}$ such that

$$a_1(1, 0, 0) + a_2(0, 1, 0) + a_3(0, 0, 1) = (0, 0, 0) = 0$$

Then this gives

$$\begin{aligned}(a_1, 0, 0) + (0, a_2, 0) + (0, 0, a_3) &= (0, 0, 0) \\ \text{or } (a_1, a_2, a_3) &= (0, 0, 0) \\ \Rightarrow a_1 = 0, a_2 = 0, a_3 = 0\end{aligned}$$

Thus none of the vectors can be written as a linear combination of other two vectors, i.e. the vectors $(1, 0, 0), (0, 1, 0)$ and $(0, 0, 1)$ are linearly independent.

Here you notice one more thing that any set of vectors is either linearly dependent or linearly independent. It cannot be both linearly dependent and linearly independent.

Definition An infinite subset of a vector space V is said to be **linearly independent** if every finite subset of S is linearly independent.

Let us discuss some elementary results concerning linearly dependent and independent vectors.

Theorem 6 Any subset of a vector space containing a zero vector is linearly dependent.

Proof: Let V be a vector space over a field F . Let $\{\alpha_1, \dots, \alpha_n\}$ be a subset of V such that $\alpha_i = 0$. Then

$$0\alpha_1 + 0\alpha_2 + \dots + 1\alpha_i + \dots + 0\alpha_n = 0$$

Since the scalar coefficient of α_i is not zero, hence the set $\{\alpha_1, \dots, \alpha_n\}$ is linearly dependent.

Theorem 7 (i) A subset of a vector space containing a single non-zero vector is linearly independent.

(ii) Every superset of a linearly dependent set of vectors is linearly dependent.

(iii) Every subset of a linearly independent set is linearly independent.

Proof: Let V be a vector space over a field F .

(i) Let $S = \{\alpha\} \subset V$ such that $\alpha \neq 0$. Then for $a \in F$, we have

$$a\alpha = 0 \Rightarrow a = 0$$

Hence the set $S = \{\alpha\}$ is linearly independent.

(ii) Let $S = \{\alpha_1, \dots, \alpha_n\}$ be a linearly dependent subset of V . Then there exist scalars $a_1, \dots, a_n \in F$ not all zero such that

$$a_1\alpha_1 + a_2\alpha_2 + \dots + a_n\alpha_n = 0 \quad \dots\dots\dots (1)$$

Suppose $T \supset S$ such that $T = \{\alpha_1, \dots, \alpha_n, \gamma_1, \dots, \gamma_k\}$. We can rewrite (1) as

$$a_1\alpha_1 + a_2\alpha_2 + \dots + a_n\alpha_n + 0\gamma_1 + \dots + 0\gamma_k = 0$$

Since all a_i 's are not zero, hence the set T is linearly dependent.

(iii) Let T be a linearly independent subset of V . Suppose $S \subset T$. We have to show that S is also linearly independent.

On the contrary, assume that S is linearly dependent. Then by (ii) the superset T of S is also linearly dependent. This contradicts the assumption that T is linearly independent. Hence S must be linearly independent.

Theorem 8 Let V be a vector space over a field F . Then the set $\{\alpha_1, \dots, \alpha_n\}$ of vectors is linearly dependent if and only if some vector from the set can be written as a linear combination of remaining vectors.

Proof: First suppose that the set $\{\alpha_1, \dots, \alpha_n\}$ is linearly dependent. Then there exists scalars $a_1, \dots, a_n$, not all are 0, such that

$$a_1\alpha_1 + a_2\alpha_2 + \dots + a_n\alpha_n = 0$$

Hence there exists $a_i \neq 0$ for some $1 \leq i \leq n$. We can write

$$a_1\alpha_1 + a_2\alpha_2 + \dots + a_i\alpha_i + \dots + a_n\alpha_n = 0$$

$$\text{or } a_i\alpha_i = -\left(\sum_{j \neq i} a_j\alpha_j\right)$$

$$\text{or } \alpha_i = -a_i^{-1}\left(\sum_{j \neq i} a_j\alpha_j\right)$$

$$\text{or} \quad \alpha_i = \sum_{j \neq i} (-a_i^{-1} a_j) \alpha_j$$

Therefore α_i is a linear combination of the remaining elements of the set $\{\alpha_1, \dots, \alpha_n\}$.

Conversely, suppose that there exists some vector α_i among $\alpha_1, \dots, \alpha_n$ which can be expressed as a linear combination of rest of the vectors, i.e.

$$\alpha_i = \sum_{j \neq i} c_j \alpha_j$$

Then we have

$$c_1 \alpha_1 + c_2 \alpha_2 + \dots + c_{i-1} \alpha_{i-1} + (-1) \alpha_i + c_{i+1} \alpha_{i+1} + \dots + c_n \alpha_n = 0$$

$$\text{or} \quad \sum_{j=1}^n c_j \alpha_j = 0, \quad \text{where } c_i = -1$$

Therefore the set $\{\alpha_1, \dots, \alpha_n\}$ is linearly dependent.

1.11 Illustrative examples

Example 1.11.1 From example 1.7.2 we know that the set of all the functions from X into F is a vector space over the field F and is denoted by $\mathcal{F}_X(F)$. Let $\mathcal{C}([a, b])$ be the set of all real valued continuous functions on $[a, b]$. Then $\mathcal{C}([a, b]) \subset \mathcal{F}_{[a, b]}(\mathbb{R})$. Also you can verify that $\mathcal{C}([a, b])$ is itself a vector space over $\mathbb{R}$ under the following operations defined in $\mathcal{F}_{[a, b]}(\mathbb{R})$. Thus $\mathcal{C}([a, b])$ is a vector subspace of $\mathcal{F}_{[a, b]}(\mathbb{R})$.

Example 1.11.2 We know that the set $M_{m \times n}(\mathbb{R})$ of all $m \times n$ matrices with entries from $\mathbb{R}$ is the vector space over $\mathbb{R}$ with respect to the operation of addition of matrices and scalar multiplication of matrices defined in example 1.7.3. Let $S_{n \times n}(\mathbb{R})$ be the set of all $n \times n$ symmetric real matrices. Then $S_{n \times n}(\mathbb{R})$ is a subspace of the vector space $M_{n \times n}(\mathbb{R})$ of all square $n \times n$ matrices (verify it).

Example 1.11.3 Let $V = \{(x_n): x_n \in \mathbb{R}\}$ and W be the set of all convergent real sequences. We know that V is a vector space over $\mathbb{R}$ under the operation given in example 1.7.5 and $W \subset V$. You can verify that W is itself a vector space over $\mathbb{R}$ under the operations defined in V . Therefore W is a subspace of V .

Example 1.11.4 Let V be a vector space over a field F and $\gamma \in V$. Let $W = \{c\gamma: c \in F\}$

Let $\alpha = c\gamma, \beta = d\gamma \in W$ for some $c, d \in F$. Then for $a, b \in F$ we have

$$a\alpha + b\beta = a(c\gamma) + b(d\gamma)$$

$$\begin{aligned}
 &= (ac)\gamma + (bd)\gamma \\
 &= (ac + bd)\gamma \in W, \text{ since } ac + bd \in F.
 \end{aligned}$$

Hence W is a subspace of V .

Now we consider an interesting situation. Suppose if we take $V = \mathbb{R}^2$, $F = \mathbb{R}$ and $\gamma = (p, q)$ then

$$W = \{t\gamma : t \in \mathbb{R}\} = \{(tp, tq) : t \in \mathbb{R}\}$$

Geometrically, for some $t \in \mathbb{R}$, (tp, tq) represents a point on the line joining the origin and the point (p, q) in $\mathbb{R}^2$. Thus the subspace W represents the line through the origin and the point (p, q) in $\mathbb{R}^2$. You can also check that W is the smallest subspace containing the point (p, q) . We can write

$$W = \mathbb{R}\gamma = L(\{\gamma\}) = \langle \{\gamma\} \rangle$$

Example 1.11.5 Let $W \subset \mathbb{R}^3$ such that

$$W = \{(a, b, c) : a, b, c \text{ are rational numbers}\}$$

Then W is not a subspace of $\mathbb{R}^3$. Why?

Example 1.11.6 Let $W = \{(a, b, c) : a, b, c \in \mathbb{R} \text{ such that } a + b + c = 0\}$. Then obviously $(0, 0, 0) \in W$. Hence W is a non-empty set. Let $\alpha, \beta \in W$ such that $\alpha = (a_1, b_1, c_1)$ and $\beta = (a_2, b_2, c_2)$. Then

$$a_1 + b_1 + c_1 = 0, a_2 + b_2 + c_2 = 0$$

For $a, b \in \mathbb{R}$, we have $a\alpha + b\beta = a(a_1, b_1, c_1) + b(a_2, b_2, c_2)$

$$\begin{aligned}
 &= a(a_1, b_1, c_1) + b(a_2, b_2, c_2) \\
 &= (aa_1, ab_1, ac_1) + (ba_2, bb_2, bc_2) \\
 &= (aa_1 + ba_2, ab_1 + bb_2, ac_1 + bc_2) \\
 &= (x, y, z), \text{ say}
 \end{aligned}$$

$$\begin{aligned}
 \text{Now } x + y + z &= (aa_1 + ba_2) + (ab_1 + bb_2) + (ac_1 + bc_2) \\
 &= a(a_1 + b_1 + c_1) + b(a_2 + b_2 + c_2) \\
 &= a \cdot 0 + b \cdot 0 = 0
 \end{aligned}$$

Hence $a\alpha + b\beta \in W$. Therefore W is a subspace of $\mathbb{R}^3$.

Example 1.11.7 Let V be a vector space over a field F and $S, T \subseteq V$. Then

- (i) $S \subseteq T \Rightarrow L(S) \subseteq L(T)$
- (ii) $L(S \cup T) = L(S) + L(T)$

Solution (i) Let $\alpha_1, \dots, \alpha_n \in S$. Then $\alpha = a_1\alpha_1 + a_2\alpha_2 + \dots + a_n\alpha_n \in L(S)$ for $a_1, \dots, a_n \in F$. Since $S \subseteq T$, hence $\alpha_1, \dots, \alpha_n \in T$. Therefore $\alpha \in L(T)$. i.e. $L(S) \subseteq L(T)$.

(ii) Let $\delta \in L(S) + L(T)$. Then there exist $\alpha \in L(S)$ and $\beta \in L(T)$ such that

$$\delta = \alpha + \beta$$

Hence for some $\alpha_1, \dots, \alpha_n \in S$ and $\beta_1, \dots, \beta_m \in T$, we have

$$\alpha = a_1\alpha_1 + a_2\alpha_2 + \dots + a_n\alpha_n$$

$$\text{and } \beta = b_1\beta_1 + b_2\beta_2 + \dots + b_m\beta_m$$

Where $a_1, \dots, a_n, b_1, \dots, b_m \in F$.

Now $\alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_m \in S \cup T$ and hence

$$a_1\alpha_1 + a_2\alpha_2 + \dots + a_n\alpha_n + b_1\beta_1 + \dots + b_m\beta_m \in L(S \cup T)$$

$$\Rightarrow \alpha + \beta \in L(S \cup T)$$

$$\Rightarrow \delta \in L(S \cup T)$$

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

Now suppose $\gamma \in L(S \cup T)$. Then

$$\gamma = \text{linear combination of finite number of elements of } S \cup T$$

$$\Rightarrow \gamma = \text{linear combination of some finite number of elements of } S \\ + \text{linear combination of some finite number of elements of } T$$

$$\Rightarrow \gamma \in L(S) + L(T)$$

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

Therefore $L(S \cup T) = L(S) + L(T)$.

Example 1.11.8 Show that the subset $\{(1,1,-1), (2,-3,5), (-2,1,4)\}$ of $\mathbb{R}^3$ is linearly independent.

Solution: Let $a, b, c \in \mathbb{R}$ such that

$$a(1,1,-1) + b(2,-3,5) + c(-2,1,4) = (0,0,0)$$

$$\Rightarrow (a, a, -a) + (2b, -3b, 5b) + (-2c, c, 4c) = (0,0,0)$$

$$\Rightarrow (a + 2b - 2c, a - 3b + c, -a + 5b + 4c) = (0,0,0)$$

$$\Rightarrow a + 2b - 2c = 0, a - 3b + c = 0, -a + 5b + 4c = 0$$

Solving these equations, we have $a = 0, b = 0, c = 0$.

Hence the given subset $\{(1,1,-1), (2,-3,5), (-2,1,4)\}$ of $\mathbb{R}^3$ is linearly independent.

Example 1.11.8 Show that the subset $\{(-1,2,1), (3,0,-1), (-5,4,3)\}$ of $\mathbb{R}^3$ is linearly dependent.

Solution: Let $a, b, c \in \mathbb{R}$ such that

$$a(-1,2,1) + b(3,0,-1) + c(-5,4,3) = (0,0,0)$$

$$\Rightarrow (-a, 2a, a) + (3b, 0, -b) + (-5c, 4c, 3c) = (0,0,0)$$

$$\Rightarrow (-a + 3b - 5c, 2a + 0b + 4c, a - b + 3c) = (0,0,0)$$

$$\Rightarrow -a + 3b - 5c = 0, 2a + 0b + 4c = 0, a - b + 3c = 0$$

The second equation gives $a = -2c$. The first and third equations become

$$2c + 3b - 5c = 0 \text{ and } -2c - b + 3c = 0$$

Each of these equations is equivalent to $b - c = 0$. This equation is satisfied by many non-zero real values. Hence the system of equations has non-zero solution. For example $a = -2, b = 1, c = 1$ is one such solution. Therefore the given the subset $\{(-1,2,1), (3,0,-1), (-5,4,3)\}$ of $\mathbb{R}^3$ is linearly dependent.

Example 1.11.9 Consider the vector space $\mathcal{F}_{\mathbb{R}}(\mathbb{R})$ of all the functions from $\mathbb{R}$ into $\mathbb{R}$. Let $f, g, h \in \mathcal{F}_{\mathbb{R}}(\mathbb{R})$ such that $f(x) = e^{2x}, g(x) = x^2, h(x) = x$ for all $x \in \mathbb{R}$. Show that the set $\{f, g, h\}$ is linearly independent.

Solution: Let $a, b, c \in \mathbb{R}$ such that

$$\begin{aligned}af + bg + ch &= \hat{0} \\ \Rightarrow af(x) + bg(x) + ch(x) &= \hat{0}(x) \text{ for all } x \in \mathbb{R} \\ \Rightarrow ae^{2x} + bx^2 + cx &= 0 \text{ for all } x \in \mathbb{R}\end{aligned}$$

Taking $x = 0, 1, 2$ we have

$$ae^0 + b0 + c0 = 0, ae^2 + b + c = 0, ae^4 + 4b + 2c = 0$$

Solving these equations, we get $a = 0, b = 0, c = 0$.

This shows that the given set $\{f, g, h\}$ is linearly independent.

1.12 Summary

We conclude with summarizing what we have covered in this unit.

- (1) We have defined binary operation, group and field with examples.
- (2) We have discussed some special fields, i.e. $\mathbb{Q}, \mathbb{R}, \mathbb{C}, \mathbb{Z}_p$ (p, a prime).
- (3) We have defined vector space with examples and discussed its properties.
- (4) Let V be a vector space over the field F . A **vector subspace** W of V is any non-empty subset $W \subset V$ which is itself a vector space over F with the operations of vector addition and scalar multiplication on V .
- (5) A non-empty subset W of a vector space V is a subspace of V if and only if for each pair of vectors α, β in W and each pair of scalars a, b in F , the vector $a\alpha + b\beta$ is again in W , i.e. $a, b \in F$ and $\alpha, \beta \in W \Rightarrow a\alpha + b\beta \in W$.
- (6) The intersection of two subspaces of a vector space V is also a subspace of V .
- (7) The union of two subspaces of a vector space V is a subspace if and only if one is contained in the other.
- (8) Let V be a vector space over a field F and let S be any non-empty subset of V . The set $L(S)$ of all finite linear combinations of elements of S is called the **linear span** of S , i.e.

$$L(S) = \left\{ \sum_{i=1}^n a_i \alpha_i : a_1, a_2, \dots, a_n \in F, \alpha_1, \alpha_2, \dots, \alpha_n \in S \right\}$$

- (9) The linear span $L(S)$ is a subspace of the vector space V .

(10) A vector subspace W of V is said to be spanned or generated by S if W is the smallest subspace of V containing S . In such a case, we write $W = \langle S \rangle$.

(11) Let V be a vector space over a field F and $S \subset V$. Then $L(S) = \langle S \rangle$.

(12) Let V be a vector space over a field F . A subset $\{\alpha_1, \dots, \alpha_n\}$ of V is said to be **linearly dependent** if there exist scalars $a_1, \dots, a_n$, not all of which are 0, such that

$$a_1\alpha_1 + a_2\alpha_2 + \dots + a_n\alpha_n = 0$$

The set $\{\alpha_1, \dots, \alpha_n\}$ of vectors in V is said to be **linearly independent** if it is not linearly dependent, that is,

$$\sum_{i=1}^n a_i\alpha_i = 0 \Rightarrow a_i = 0 \text{ for all } 1 \leq i \leq n$$

(13) An infinite subset of a vector space V is said to be **linearly independent** if every finite subset of S is linearly independent.

(14) Any subset of a vector space containing a zero vector is linearly dependent.

(15) A subset of a vector space containing a single non-zero vector is linearly independent.

(16) Every superset of a linearly dependent set of vectors is linearly dependent.

(17) Every subset of a linearly independent set is linearly independent.

(18) Let V be a vector space over a field F . Then the set $\{\alpha_1, \dots, \alpha_n\}$ of vectors is linearly dependent if and only if some vector from the set can be written as a linear combination of remaining vectors.

1.13 Terminal questions

(1) Show that a field K is a vector space over any subfield F of K .

(2) Let $F[x]$ denote the set of all polynomials in an indeterminate x over a field F . Show that $F[x]$ is a vector space over the field F with respect to addition of two polynomials as addition of vectors and the product of a polynomial by an element of F as scalar multiplication.

(3) Show that the set of all convergent real sequences is a vector space over the field of real numbers.

(4) Let $(x_1, y_1), (x_2, y_2) \in \mathbb{R}^2$ and $c \in \mathbb{R}$. Define

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

$$c(x_1, y_1) = (cx_1, 0)$$

Is $\mathbb{R}^2$ a vector space over $\mathbb{R}$ under these operations? [Ans: No]

(5) Show that the set $\mathbb{C}$ of complex numbers is a vector space over $\mathbb{R}$ under the following operations

$$(x + iy) + (u + iv) = (x + u) + i(y + v)$$

$$a(x + iy) = ax + iay$$

for all $(x + iy), (u + iv) \in \mathbb{C}$ and $a \in \mathbb{R}$.

(6) Show that a non-empty subset W of a vector space V is a subspace of V if and only if for each pair of vectors α, β in W and each scalar a in F , the vector $a\alpha + \beta$ is again in W , i.e. $a \in F$ and $\alpha, \beta \in W \Rightarrow a\alpha + \beta \in W$.

(7) Show that $W = \{(a, 2b, 3c) : a, b, c \in \mathbb{R}\}$ is a subspace of $\mathbb{R}^3$.

(8) Let $W = \{A \in M_{2 \times 2}(\mathbb{R}) : \det(A) = 0\}$. Is W a subspace of $M_{2 \times 2}(\mathbb{R})$?

[Ans: No]

(9) Prove that arbitrary intersection of subspaces of a vector space is a subspace.

(10) Prove that (i) S is a subspace of a vector space $V \Leftrightarrow L(S) = S$

(ii) $L(L(S)) = L(S)$

(11) Show that the vectors $(1, 1, 1)$, $(0, 1, 1)$ and $(0, 1, -1)$ generate $\mathbb{R}^3$.

(12) Is the vector $(2, -5, 3)$ in the subspace of $\mathbb{R}^3$ spanned by the vectors $(1, -3, 2)$, $(2, -4, -1)$, $(1, -5, 7)$? [Ans: No]

(13) Show that the subset $\{e_1, e_2, \dots, e_n\}$ of the vector space $\mathbb{R}^n$ is linearly independent. Where $e_1 = (1, 0, \dots, 0)$, $e_2 = (0, 1, \dots, 0)$, ..., $e_n = (0, 0, \dots, 1)$.

(14) Show that the subset $\{(1, 3, 2), (1, -7, -8), (2, 1, -1)\}$ of $\mathbb{R}^3$ is linearly dependent.

(15) If α, β, γ are linearly independent vectors of a vector space V over the subfield F of the complex numbers, then prove that $\alpha + \beta, \beta + \gamma, \gamma + \alpha$ are also linearly independent.

(16) Let $\alpha = (1 + i, 2i)$ and $\beta = (1, 1 + i)$. Show that α and β are linearly independent in the vector space $\mathbb{C}^2$ over the field $\mathbb{R}$ and linearly dependent in the vector space $\mathbb{C}^2$ over the field $\mathbb{C}$.

1.14 Further readings

(1) Hoffman, K. and Kunze, R. (2015): Linear Algebra, Pearson India Education Services Ltd.

(2) Kumaresan, S. (2014): Linear Algebra: A Geometric approach, New Delhi, Prentice Hall of India.

(3) Strang, G. (2006): Lineart Algebra and its applications, New Delhi, Cengage Learning India Pvt Ltd.

(4) Lang, S. (2014): Introduction to Linear Algebra, New Delhi, Springer India Pvt. Ltd.

Unit-2: Basis, Dimension and Quotient Spaces

Structure

- 3.1 Introduction**
- 3.2 Objectives**
- 3.3 Basis and Dimension**
 - 2.3.1 Basis of a vector space**
 - 2.3.2 Finite dimensional vector spaces**
 - 2.3.3 Dimension of a vector space**
- 3.4 Coordinates**
- 3.5 Illustrative examples**
- 3.6 linear sum and direct sum of subspaces**
- 3.7 Quotient spaces**
- 3.8 Illustrative examples**
- 3.9 Summary**
- 3.10 Terminal questions**
- 3.11 Further readings**

2.1 Introduction

Consider the vector space V of all the vectors (directed segments) in three dimensional Euclidean space. Let $S = \{\hat{i}, \hat{j}, \hat{k}\} \subset V$ be the set of unit vectors along coordinate axes. Then the set S will be linearly independent as

$$a\hat{i} + b\hat{j} + c\hat{k} = \mathbf{0} \Rightarrow a = 0, b = 0, c = 0$$

Also we see that every vector $\vec{d} \in V$ can be expressed as a linearly combination of $\hat{i}, \hat{j}$ and $\hat{k}$, i.e. there exist $a_1, a_2, a_3 \in \mathbb{R}$ such that

$$\vec{d} = a_1\hat{i} + a_2\hat{j} + a_3\hat{k}$$

Hence $L(S) = V$. Thus S is a linearly independent subset of V which generates V . Such a subset is called a basis of the vector space. The interesting thing about a basis is that you can create the entire space by its vectors. In this unit, we shall discuss the concept of basis in details. We shall see how to characterize a basis for a given vector space. Some vector spaces are generated by their finite subsets. Such spaces are called finite dimensional. We shall see that there always exists a basis for each finite dimensional vector space and any two bases have the same number of elements. This facilitates the definition of dimension of a finitely generated vector space as the number of elements in any of its basis. Some other concepts such as coordinates, linear sum and direct sum of subspaces and quotient spaces will be discussed.

2.2 Objectives

After reading this unit, you should be able to:

- Define the basis of a vector space;
- Characterize the basis as a minimal generating set and as a maximal linearly independent subset of the given vector space;
- Discuss various properties of bases;
- Determine whether a given subset of a vector space is a basis of the vector space or not;
- Construct a basis of a finite-dimensional vector space;
- Determine the dimensions of subspaces;
- Define coordinates, linear sum and direct sum of subspaces and quotient spaces
- Determine the dimensions of the sum of two subspaces, intersection of two subspaces and quotient spaces

2.3 Basis and Dimension

2.3.1 Basis of a vector space

As we have seen that there exist some linearly independent subsets for a given vector space which generate the entire vector space. Such subsets are called bases of the given space. Let us give a formal definition of this concept:

Definition Let V be a vector space over a field F . A subset S of V is said to be a **basis** of V if S is linearly independent and spans the space V , i.e. $L(S) = V$.

Example 2.3.1 The vectors $e_1 = (1, 0, \dots, 0)$, $e_2 = (0, 1, \dots, 0)$, ..., $e_n = (0, 0, \dots, 1)$ form a basis of $\mathbb{R}^n$.

Let $a_1, \dots, a_n \in \mathbb{R}$ such that

$$\begin{aligned} a_1 e_1 + a_2 e_2 + \dots + a_n e_n &= \mathbf{0} \\ \Rightarrow a_1(1, 0, \dots, 0) + a_2(0, 1, \dots, 0) + \dots + a_n(0, 0, \dots, 1) &= (0, 0, \dots, 0) \\ \Rightarrow (a_1, \dots, a_n) &= (0, 0, \dots, 0) \\ \Rightarrow a_1 = 0, \dots, a_n &= 0 \end{aligned}$$

Hence the vectors are linearly independent.

Also if $\alpha \in \mathbb{R}^n$ such that $\alpha = (c_1, \dots, c_n)$ then

$$\alpha = c_1 e_1 + c_2 e_2 + \dots + c_n e_n$$

Therefore the vectors $e_1, \dots, e_n$ generate $\mathbb{R}^n$. Hence $\{e_1, \dots, e_n\}$ is a basis of $\mathbb{R}^n$. This is called the **standard basis** of $\mathbb{R}^n$.

Example 2.3.2 Let $M_{n \times n}(\mathbb{R})$ be the vector space of all square $n \times n$ matrices with real entries. Let $E_{ij} \in M_{n \times n}(\mathbb{R})$ such that $E_{ij} = [a_{pq}]$ where

$$a_{pq} = \begin{cases} 1, & \text{if } p = i, q = j \\ 0, & \text{if } p \neq i, q \neq j \end{cases}$$

Then the set $\{E_{ij} : 1 \leq i, j \leq n\}$ forms a basis of $M_{n \times n}(\mathbb{R})$.

Example 2.3.3 Let F be any field and let V be the vector space of all polynomial functions in one variable with coefficients belonging to F . Then the infinite set $\{1, x, \dots, x^k, \dots\}$ is a basis for V .

2.3.2 Finite dimensional vector spaces

Definition A vector space V is called **finite dimensional** or **finitely generated** if it is the zero space $\{0\}$, or if there exists a finite subset S of V which spans V , i.e. $L(S) = V$.

Theorem 1 There exists a basis for every finite dimensional vector space.

Proof Suppose V is a finite dimensional vector space over a field F and S is the finite subset of V generating it, i.e. $L(S) = V$. Let

$$S = \{\alpha_1, \alpha_2, \dots, \alpha_n\}$$

If S is linearly independent there is nothing to prove, i.e. the subset S will form a basis of V . If S is linearly dependent then there exist scalars $a_1, \dots, a_n$, not all of which are 0, such that

$$a_1\alpha_1 + a_2\alpha_2 + \dots + a_n\alpha_n = 0$$

Let k be the largest integer such that $a_k \neq 0$. Hence $a_{k+1} = 0, \dots, a_n = 0$. Thus we have

$$a_1\alpha_1 + a_2\alpha_2 + \dots + a_k\alpha_k = 0$$

Here $k \geq 2$. For if $k = 1$, we have $a_1\alpha_1 = 0$. Since $a_1 \neq 0$, this gives $\alpha_1 = 0$. This is a contradiction. Therefore $2 \leq k \leq n$. Now we can write

$$\begin{aligned} a_k\alpha_k &= -(a_1\alpha_1 + a_2\alpha_2 + \dots + a_{k-1}\alpha_{k-1}) \\ \text{or } a_k &= -a_k^{-1}(a_1\alpha_1 + a_2\alpha_2 + \dots + a_{k-1}\alpha_{k-1}) \\ \text{or } \alpha_k &= (-a_k^{-1}a_1)\alpha_1 + (-a_k^{-1}a_2)\alpha_2 + \dots + (-a_k^{-1}a_{k-1})\alpha_{k-1} \\ &= b_1\alpha_1 + b_2\alpha_2 + \dots + b_{k-1}\alpha_{k-1} \text{ where } b_i = -a_k^{-1}a_i \end{aligned}$$

i.e. α_k is a linear combination of its predecessors. If we omit this vector α_k from S , the subset $\{\alpha_1, \alpha_2, \dots, \alpha_{k-1}, \alpha_{k+1}, \dots, \alpha_n\}$ also generates V . For if $\alpha \in V$, then there exists scalars $c_1, \dots, c_n$ such that

$$\begin{aligned} \alpha &= c_1\alpha_1 + c_2\alpha_2 + \dots + c_n\alpha_n \\ &= c_1\alpha_1 + \dots + c_k(b_1\alpha_1 + \dots + b_{k-1}\alpha_{k-1}) + \dots + c_n\alpha_n \\ &= \text{A linear combination of vectors } \alpha_1, \dots, \alpha_{k-1}, \alpha_{k+1}, \dots, \alpha_n \end{aligned}$$

If the subset $\{\alpha_1, \alpha_2, \dots, \alpha_{k-1}, \alpha_{k+1}, \dots, \alpha_n\}$ is linearly independent, then it will form a basis of V . Otherwise we may continue this process of omitting vectors until we get a linearly independent subset of S which generates V . Thus we get a subset $\{\beta_1, \beta_2, \dots, \beta_m\} \subset S$ which is linearly independent and spans V , i.e. we get a basis of V .

Remark In light of above theorem we can say that *any finite spanning set in a vector space contains a basis*.

Theorem 2 (Characterization of a Basis) Let V be a vector space over a field F . Let B be a non-empty subset of V . Then the following conditions on B are equivalent:

- B is a basis for V .
- Every element of V can be uniquely expressed as a linear combination of elements of B .
- B is a minimal generating set, i.e. B spans V but any proper subset of B does not span V .
- B is maximal linearly independent, i.e. B is linearly independent but any proper superset of B is not linearly independent.

Proof: (i) $\Rightarrow$ (ii) Assume (i). Then B is a linearly independent subset of V such that $L(B) = V$. Let $0 \neq \alpha \in V$ and

$$\alpha = a_1\alpha_1 + a_2\alpha_2 + \cdots + a_n\alpha_n = b_1\beta_1 + b_2\beta_2 + \cdots + b_m\beta_m \quad \dots\dots (1)$$

Where $\alpha_1, \alpha_2, \dots, \alpha_n \in B$ are distinct and $\beta_1, \beta_2, \dots, \beta_m \in B$ are distinct. Also a_i 's and b_j 's are nonzero scalars. The sets $\{\alpha_1, \alpha_2, \dots, \alpha_n\}$ and $\{\beta_1, \beta_2, \dots, \beta_m\}$ cannot be disjoint, for otherwise (1) implies

$$a_1\alpha_1 + a_2\alpha_2 + \cdots + a_n\alpha_n + (-b_1)\beta_1 + \cdots + (-b_m)\beta_m = 0$$

This makes B linearly dependent and hence not possible. Suppose

$$\{\alpha_1, \alpha_2, \dots, \alpha_n\} \cap \{\beta_1, \beta_2, \dots, \beta_m\} = \{\alpha_{i_1}, \alpha_{i_2}, \dots, \alpha_{i_k}\}$$

Then after rearranging the terms, we can write

$$\{\alpha_1, \alpha_2, \dots, \alpha_n\} = \{\alpha_{i_1}, \alpha_{i_2}, \dots, \alpha_{i_k}, \alpha_{i_{k+1}}, \dots, \alpha_{i_n}\}$$

$$\text{and} \quad \{\beta_1, \beta_2, \dots, \beta_m\} = \{\alpha_{i_1}, \alpha_{i_2}, \dots, \alpha_{i_k}, \beta_{i_{k+1}}, \dots, \beta_{i_m}\}$$

Hence (1) can be written as

$$(a_{i_1} - b_{i_1})\alpha_{i_1} + \cdots + (a_{i_k} - b_{i_k})\alpha_{i_k} + a_{i_{k+1}}\alpha_{i_{k+1}} + \cdots + a_{i_n}\alpha_{i_n} \\ + (-b_{i_{k+1}})\beta_{i_{k+1}} + \cdots + (-b_{i_m})\beta_{i_m} = 0$$

Now B is a basis, i.e. B is linearly independent, therefore we must have $n = m = k$ and $\alpha_{i_r} = \beta_{i_r}$ for each r , i.e. $\alpha_j = \beta_j$ for all j . Therefore we have

$$(a_{i_1} - b_{i_1})\alpha_{i_1} + \cdots + (a_{i_n} - b_{i_n})\alpha_{i_n} = 0$$

After rearrangement we can write

$$(a_1 - b_1)\alpha_1 + \cdots + (a_n - b_n)\alpha_n = 0$$

Since the B is linearly independent, hence we have

$$a_1 = b_1, \dots, a_n = b_n$$

Hence (i) $\Rightarrow$ (ii).

(ii) $\Rightarrow$ (iii) Suppose (ii) holds. Obviously B spans V as any $\alpha \in V$ can be uniquely expressed as a linear combination of elements of B . Assume that (iii) does not hold, i.e. B is not a minimal generating set. Then there is a proper subset B' of B such that $L(B') = V$. Let $\alpha \in B - B'$. Then

$$\alpha = a_1\alpha_1 + a_2\alpha_2 + \cdots + a_n\alpha_n$$

Where $\alpha_1, \alpha_2, \dots, \alpha_n$ are distinct elements of B' and a_i 's are nonzero scalars.

Now $B' \subset B \Rightarrow \alpha_1, \alpha_2, \dots, \alpha_n \in B$. Also since $\alpha \in B$, hence we can write

$$\alpha = 1\alpha$$

Thus we get two distinct representations of α . This contradicts (ii). Hence (iii) must hold, i.e. (ii) $\Rightarrow$ (iii).

(iii) $\Rightarrow$ (iv) Suppose (iii) holds. Then B must be linearly independent, for otherwise there exist distinct elements $\alpha_1, \alpha_2, \dots, \alpha_n \in B$ and nonzero scalars $a_1, a_2, \dots, a_n \in F$ such that

$$a_1\alpha_1 + a_2\alpha_2 + \cdots + a_n\alpha_n = 0$$

There will be no loss of generality if we assume $a_1 \neq 0$. Hence we can write

$$\alpha_1 = (-a_1^{-1}a_2)\alpha_2 + \cdots + (-a_1^{-1}a_n)\alpha_n$$

$\Rightarrow \alpha_1$ is a linear combination of the elements of $B - \{\alpha_1\}$

$\Rightarrow B \subset L(B - \{\alpha_1\})$

Since $L(B) = V$, therefore we have $L(B - \{\alpha_1\}) = V$. This contradicts (iii).

Hence B is linearly independent.

Now suppose that $S \supset B$. Let $0 \neq \alpha \in S - B$. Since $L(B) = V$, hence there exist distinct elements $\gamma_1, \gamma_2, \dots, \gamma_n \in B$ and nonzero scalars $c_1, c_2, \dots, c_n \in F$ such that

$$\begin{aligned}\alpha &= c_1\gamma_1 + \dots + c_n\gamma_n \\ \Rightarrow 1\alpha + (-c_1)\gamma_1 + \dots + (-c_n)\gamma_n &= 0\end{aligned}$$

Therefore S is linearly dependent. Hence B is maximal linearly independent.

(iv) $\Rightarrow$ (i) Suppose (iv) holds. Then B must generate V , for if not, there is a vector $\alpha \in V - B$ that is not a linear combination of the elements of B , i.e. $\alpha \notin L(B)$. Therefore $B \cup \{\alpha\}$ is linearly independent superset of B . This contradicts (iv). Hence B generates V . Thus B is a linearly independent set generating V , i.e. B is a basis for V .

Theorem 3 If a basis of a vector space V contains n elements, then any linearly independent subset of V contains no more than n elements.

Proof: Let $B = \{\alpha_1, \alpha_2, \dots, \alpha_n\}$ be a basis of a vector space V containing n elements. Suppose $S = \{\beta_1, \beta_2, \dots, \beta_m\}$ be any subset of V such that $m > n$. Since B is a basis of V , hence $L(B) = V$. Therefore for any $\beta_j \in S$ there exist scalars $a_{1j}, a_{2j}, \dots, a_{nj} \in F$ such that

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

Now for any $b_1, b_2, \dots, b_m \in F$, we have

$$\begin{aligned}\sum_{j=1}^m b_j\beta_j &= \sum_{j=1}^m b_j \left(\sum_{i=1}^n a_{ij}\alpha_i \right) \\ &= \sum_{i=1}^n \left(\sum_{j=1}^m a_{ij}b_j \right) \alpha_i\end{aligned}$$

Therefore $\sum_{j=1}^m b_j\beta_j = 0$ whenever $\sum_{i=1}^n \left(\sum_{j=1}^m a_{ij}b_j \right) \alpha_i = 0$. Since α_i 's are linearly independent, hence

$$\sum_{i=1}^n \left(\sum_{j=1}^m a_{ij}b_j \right) \alpha_i = 0 \Rightarrow \sum_{j=1}^m a_{ij}b_j = 0 \text{ for } 1 \leq i \leq n$$

So we have n equations in m variables $b_1, b_2, \dots, b_m$. Since $m > n$, hence the system of equations will have a non-trivial solution, i.e. there exist scalars $b_1, b_2, \dots, b_m$ not all zero such that

$$\sum_{j=1}^m b_j\beta_j = 0$$

Consequently, the set $S = \{\beta_1, \beta_2, \dots, \beta_m\}$ is linearly dependent.

Corollary If V is a finite dimensional vector space over a field F , then any two bases of V contain same number of elements.

Proof: Let $\{\alpha_1, \alpha_2, \dots, \alpha_n\}$ and $\{\beta_1, \beta_2, \dots, \beta_m\}$ be any two bases of V . Since the vectors $\beta_1, \beta_2, \dots, \beta_m$ are linearly independent, hence by Theorem 3 we have $m \leq n$. Similarly $\{\beta_1, \beta_2, \dots, \beta_m\}$ is a basis and $\alpha_1, \alpha_2, \dots, \alpha_n$ are linearly independent, hence we have $n \leq m$. Therefore we have $m = n$.

Remark A vector space V is finite-dimensional if it is the zero space, or if V has a finite basis. All other vector spaces are called **infinite-dimensional**. Now we are in position to define the dimension of a finite dimensional vector space.

2.3.3 Dimension of a vector space

Definition The **dimension** of a non-zero finite-dimensional vector space V is the number of elements in a basis for V . The dimension of the zero space is zero.

Thus if a vector space V has a basis consisting of n vectors, we say that V is n -dimensional and we write $\dim V = n$.

For example, the vector space $\mathbb{R}^n$ has dimension n . The subspace W_1 of $\mathbb{R}^n$ given by $W_1 = \{(x_1, \dots, x_n) : x_1, \dots, x_n \in \mathbb{R} \text{ such that } x_1 + \dots + x_n = 0\}$ has dimension $n - 1$ as the set $\{(1, -1, \dots, 0), (1, 0, -1, \dots, 0), \dots, (1, 0, \dots, 0, -1)\}$ is a basis of W_1 .

Can you guess the dimension of the following subspace of $\mathbb{R}^n$?

$$W_2 = \{(x_1, \dots, x_n) : x_1, \dots, x_n \in \mathbb{R} \text{ such that } x_1 = \dots = x_m = 0\}$$

for any fixed $m, 1 \leq m < n$.

The dimension of W_2 is $n - m$ (why?).

Remark: In view of Theorem 3, we can say that if V is a finite-dimensional vector space such that $\dim V = n$, then any linearly independent subset of V contains no more than n elements, i.e. any subset of V which contains more than n vectors is linearly dependent.

Theorem 4 Let V be a finite-dimensional vector space such that $\dim V = n$. Then any linearly independent set of n -vectors in V forms a basis for V .

Proof Suppose V is a finite-dimensional vector space such that $\dim V = n$, i.e. the number of elements in any of the bases of V is n . Let $S = \{\alpha_1, \alpha_2, \dots, \alpha_n\}$ be a linearly independent set. Assume $\alpha \in V$. Then by theorem 3, the set $\{\alpha, \alpha_1, \alpha_2, \dots, \alpha_n\}$ is linearly dependent. Therefore there exist scalars $a_0, a_1, \dots, a_n$, not all of which are 0, such that

$$a_0\alpha + a_1\alpha_1 + a_2\alpha_2 + \cdots + a_n\alpha_n = 0 \quad \dots \dots (2)$$

Now $a_0 \neq 0$. For if $a_0 = 0$, then (2) gives

$$a_1\alpha_1 + a_2\alpha_2 + \cdots + a_n\alpha_n = 0$$

Since α 's are linearly independent, we have $a_1 = 0, a_2 = 0, \dots, a_n = 0$. This contradicts the fact that not all of the scalars $a_0, a_1, \dots, a_n$ are zero.

Hence from (2), we have

$$\begin{aligned} \alpha &= -a_0^{-1}(a_1\alpha_1 + a_2\alpha_2 + \cdots + a_n\alpha_n) \\ \text{or } \alpha &= (-a_0^{-1}a_1)\alpha_1 + (-a_0^{-1}a_2)\alpha_2 + \cdots + (-a_0^{-1}a_n)\alpha_n \\ \text{or } \alpha &= b_1\alpha_1 + b_2\alpha_2 + \cdots + b_n\alpha_n \text{ where } b_i = -a_0^{-1}a_i \end{aligned}$$

Therefore the set S generates the vector space V , i.e. $L(S) = V$.

Thus S is a linearly independent set generating V . Hence S is a basis of V .

Theorem 5 Let W be a subspace of a finite-dimensional vector space V . Then every linearly independent subset of W can be extended to a basis of W .

Proof: Let $S = \{\alpha_1, \alpha_2, \dots, \alpha_m\}$ be a linearly independent subset of W . Let $\dim V = n$. Then any linearly independent subset of V cannot contain more than n elements. Hence any linearly independent subset of W containing S cannot have more than n elements.

If S generates W , then S is the required basis of W . Suppose S does not span W . Let $\beta_1 \in W$ such that $\beta_1 \notin L(S)$. If we adjoin β_1 to the set S , then the set $S_1 = S \cup \{\beta_1\} = \{\alpha_1, \alpha_2, \dots, \alpha_m, \beta_1\}$ is also linearly independent subset of W . Since if

$$a_1\alpha_1 + a_2\alpha_2 + \cdots + a_m\alpha_m + b_1\beta_1 = 0 \quad \dots \dots (3)$$

Then $b_1 = 0$. For if $b_1 \neq 0$, then (3) gives

$$\begin{aligned} \beta_1 &= (-b_1^{-1}a_1)\alpha_1 + (-b_1^{-1}a_2)\alpha_2 + \cdots + (-b_1^{-1}a_m)\alpha_m \\ \Rightarrow \beta_1 &\in L(S) \end{aligned}$$

This contradicts our assumption that $\beta_1 \notin L(S)$. Hence $b_1 = 0$. Therefore (3) becomes

$$\begin{aligned} a_1\alpha_1 + a_2\alpha_2 + \cdots + a_m\alpha_m &= 0 \\ \Rightarrow a_1 = 0, a_2 = 0, \dots, a_m = 0 &\text{ as } S \text{ is linearly independent.} \end{aligned}$$

Thus $a_1 = 0, a_2 = 0, \dots, a_m = 0, b_1 = 0$. Hence the set S_1 is linearly independent. If S_1 generates W , it will be a basis of W . If $L(S_1) \neq W$, then there exists vector $\beta_2 \in W$ such that $\beta_2 \notin L(S_1)$. Then proceeding as above, we can show that the set $S_2 = S_1 \cup \{\beta_2\} = \{\alpha_1, \alpha_2, \dots, \alpha_m, \beta_1, \beta_2\}$ is linearly independent. Continuing the process, we will have a linearly independent set

$$S_k = \{\alpha_1, \alpha_2, \dots, \alpha_m, \beta_1, \beta_2, \dots, \beta_k\}$$

such that $L(S_k) = W$, where $m + k \leq n$. Therefore S_k is the required extended linearly independent set which forms a basis of W .

Corollary 1. If V is a finite dimensional vector space over a field F , then every non-empty linearly independent set of vectors is a part of a basis of V .

Corollary 2. Let W be a proper subspace of a finite-dimensional vector space V . Then W is finite-dimensional and $\dim W < \dim V$.

Theorem 6 Let W be a subspace of a finite-dimensional vector space V . Then every generating set S of W contains a basis of W , i.e. if $L(S) = W$ then S contains a basis of W .

Proof Suppose S_1 is the maximal linearly independent subset of S . We shall show that S_1 is a basis of W . Let $\alpha \in S$. If $\alpha \in S_1$, then $\alpha = 1\alpha \in L(S_1)$.

Also if $\alpha \notin S_1$, then the set $S_1 \cup \{\alpha\}$ is linearly dependent (as S_1 is the maximal linearly independent subset of S). Hence There exists a finite linearly dependent subset $\{\alpha_1, \alpha_2, \dots, \alpha_m, \alpha\}$ of $S_1 \cup \{\alpha\}$. There must exist some vector in $S_1 \cup \{\alpha\}$ which can be written as a linear combination of preceding vectors. Since $\alpha_1, \alpha_2, \dots, \alpha_m$ are linearly independent vectors of S_1 , hence we have $\alpha = a_1\alpha_1 + a_2\alpha_2 + \dots + a_m\alpha_m$ for some scalars $a_1, a_2, \dots, a_m$, i.e. $\alpha \in L(S_1)$. Thus $S \subseteq L(S_1)$.

Now $S \subseteq L(S_1) \Rightarrow L(S) \subseteq L(L(S_1))$

$$\Rightarrow L(S) \subseteq L(S_1) \text{ as } L(L(S_1)) = L(S_1)$$

$$\Rightarrow W \subseteq L(S_1) \text{ as } L(S) = W$$

$$\Rightarrow W = L(S_1) \text{ as } S_1 \subseteq W$$

Since S_1 is linearly independent, hence S_1 is a basis of W .

Corollary If W is a subspace of a finite-dimensional vector space V such that $\dim W = m$, then any generating set with m vectors forms a basis of W .

Theorem 7 Let W_1 and W_2 be any two subspaces of a finite-dimensional vector space V such that $W_1 \subseteq W_2$. Then $\dim W_1 \leq \dim W_2$. Also $\dim W_1 = \dim W_2$ if and only if $W_1 = W_2$.

Proof: Let B be a basis of W_1 . Since $W_1 \subseteq W_2$, hence B is a linearly independent subset of W_2 . Now B can be extended to a basis of W_2 . Therefore $\dim W_1 \leq \dim W_2$.

If $W_1 = W_2$, then obviously $\dim W_1 = \dim W_2$.

Conversely, suppose $\dim W_1 = \dim W_2 = k$ (say). If B is a basis of W_1 , then $L(B) = W_1$. Since $W_1 \subseteq W_2$, hence B is a linearly independent subset of W_2 containing k vectors. Therefore B is a basis of W_2 . Hence we have $L(B) = W_2$ and so $W_1 = L(B) = W_2$. Thus $\dim W_1 = \dim W_2$ if and only if $W_1 = W_2$.

Corollary If $\dim V = n$, then V itself is the only subspace of V of dimension n .

2.4 Coordinates

Consider a vector $\alpha = (2, 3, -5)$ in the space $\mathbb{R}^3$. Obviously

$$(2, 3, -5) = 2(1, 0, 0) + 3(0, 1, 0) + (-5)(0, 0, 1)$$

In general, a vector $\alpha = (x_1, \dots, x_n) \in \mathbb{R}^n$ can be expressed as a unique linear combination of vectors in standard basis $\{e_1, \dots, e_n\}$ in the following way

$$\alpha = x_1 e_1 + x_2 e_2 + \dots + x_n e_n$$

Thus the i^{th} coordinate x_i of α is the coefficient of e_i in above linear combination. So we can define a coordinate system in $\mathbb{R}^n$ with respect to the standard basis $\{e_1, \dots, e_n\}$. To generalize this concept to any n -dimensional vector space, we need the following definition.

Definition Let V be a finite-dimensional vector space over a field F . An ordered basis for V is a finite sequence of vectors which is independent and generates V .

If $\alpha_1, \alpha_2, \dots, \alpha_n$ is such finite sequence, then we say that $\{\alpha_1, \alpha_2, \dots, \alpha_n\}$ is an ordered basis for V .

Let $\alpha \in V$. Suppose, if possible, that

$$\alpha = x_1 \alpha_1 + x_2 \alpha_2 + \dots + x_n \alpha_n \text{ and } \alpha = y_1 \alpha_1 + y_2 \alpha_2 + \dots + y_n \alpha_n$$

Where x 's and y 's are scalars. Then we have

$$x_1 \alpha_1 + x_2 \alpha_2 + \dots + x_n \alpha_n = y_1 \alpha_1 + y_2 \alpha_2 + \dots + y_n \alpha_n$$

$$\Rightarrow (x_1 - y_1) \alpha_1 + (x_2 - y_2) \alpha_2 + \dots + (x_n - y_n) \alpha_n = 0$$

Since $\alpha_1, \alpha_2, \dots, \alpha_n$ are linearly independent, hence

$$x_1 - y_1 = 0, x_2 - y_2 = 0, \dots, x_n - y_n = 0$$

$$\Rightarrow x_1 = y_1, x_2 = y_2, \dots, x_n = y_n$$

Hence for $\alpha \in V$, there exists a unique n -tuple $(x_1, \dots, x_n)$ of scalars such that

$$\alpha = \sum_{i=1}^n x_i \alpha_i$$

We may assume x_i as functions defined on V . The functions $x_i: V \rightarrow F$ are called coordinate functions with respect to the given ordered basis and $x_i(\alpha)$ is called the i^{th} coordinate of α relative to the ordered basis $\{\alpha_1, \alpha_2, \dots, \alpha_n\}$.

If $(x_1, \dots, x_n)$ is the n -tuple of coordinates of $\alpha \in V$ with respect to an ordered basis $B = \{\alpha_1, \alpha_2, \dots, \alpha_n\}$, then the coordinate matrix of α relative to the ordered basis B is defined as follows

$$[\alpha]_B = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

This notation is helpful when we change the basis and obtain the corresponding coordinates. We shall discuss it in details in the next unit.

2.5 Illustrative Examples

Example 2.5.1 Show that the vectors $\{(1,0,-1), (1,2,1), (0,-3,2)\}$ forms a basis for $\mathbb{R}^3$.

Solution : Let $a, b, c \in \mathbb{R}$. Then

$$\begin{aligned} & a(1,0,-1) + b(1,2,1) + c(0,0,1) = (0,0,0) \\ \Rightarrow & (a+b+0c, 0a+2b+0c, -a+b+c) = (0,0,0) \\ \Rightarrow & a+b+0c = 0, \quad 0a+2b+0c = 0, \quad -a+b+c = 0 \\ \Rightarrow & a = 0, b = 0, c = 0 \end{aligned}$$

Hence the given vectors are linearly independent. Since $\dim \mathbb{R}^3 = 3$, hence any set with three linearly independent vectors will form a basis of $\mathbb{R}^3$. Therefore $\{(1,0,-1), (1,2,1), (0,-3,2)\}$ forms a basis for $\mathbb{R}^3$.

Example 2.5.2 Determine whether or not the vectors $(1,1,2)$, $(1,2,5)$, and $(5,3,4)$ form a basis for $\mathbb{R}^3$.

Solution Let $a, b, c \in \mathbb{R}$. Then

$$\begin{aligned} & a(1,1,2) + b(1,2,5) + c(5,3,4) = (0,0,0) \\ \Rightarrow & (a+b+5c, a+2b+3c, 2a+5b+4c) = (0,0,0) \\ \Rightarrow & a+b+5c = 0, \quad a+2b+3c = 0, \quad 2a+5b+4c = 0 \end{aligned}$$

The coefficient matrix is

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

Now

$$|A| = \begin{vmatrix} 1 & 1 & 5 \\ 1 & 2 & 3 \\ 2 & 5 & 4 \end{vmatrix} = 0$$

Since the coefficient matrix is singular, the system of equations will have a nonzero solution. Therefore the given vectors are linearly dependent.

Example 2.5.3 Show that the set $\{1, x, x^2, \dots, x^n\}$ is a basis of the vector space P_n of all polynomials in x of degree at most n over the field of real numbers.

Solution Let $S = \{1, x, x^2, \dots, x^n\}$. Let $a_0, a_1, \dots, a_n \in \mathbb{R}$ such that

$$a_0 1 + a_1 x + a_2 x^2 + \cdots + a_n x^n = 0 \text{ (zero polynomial)}$$

$$\Rightarrow a_0 1 + a_1 x + a_2 x^2 + \cdots + a_n x^n = 0 \cdot 1 + 0x + 0x^2 + \cdots + 0x^n$$

By the definition of equality of two polynomials, we have

$$a_0 = 0, a_1 = 0, \dots, a_n = 0$$

Hence the set S is linearly independent subset of P_n .

Obviously if $\alpha \in P_n$, then there exists $c_0, c_1, \dots, c_n \in \mathbb{R}$ such that

$$\alpha = c_0 1 + c_1 x + \cdots + c_n x^n$$

Thus S spans P_n . Therefore S is a basis of P_n .

Example 2.5.4 If $\{\alpha, \beta, \gamma\}$ is a basis of $\mathbb{R}^3$, show that the set $\{\alpha + \beta, \beta + \gamma, \gamma + \alpha\}$ is also a basis of $\mathbb{R}^3$.

Solution Let $S = \{\alpha + \beta, \beta + \gamma, \gamma + \alpha\}$. Since $\dim \mathbb{R}^3 = 3$, hence any linearly independent subset of $\mathbb{R}^3$ with three vectors will form a basis of $\mathbb{R}^3$. Hence in order to prove that the set S is a basis of $\mathbb{R}^3$, we simply show that the set S is linearly independent.

Let $a, b, c \in \mathbb{R}$. Then

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

$$\Rightarrow (a + c)\alpha + (a + b)\beta + (b + c)\gamma = 0$$

Since $\{\alpha, \beta, \gamma\}$ is a basis of $\mathbb{R}^3$, the vectors α, β, γ are linearly independent. Hence we have $a + c = 0, a + b = 0, b + c = 0$. Solving these equations we get $a = 0, b = 0, c = 0$, i.e. the vectors $\alpha + \beta, \beta + \gamma, \gamma + \alpha$ are linearly independent and therefore form a basis of $\mathbb{R}^3$.

Example 2.5.5 Let $W = \{(x, y, z) \in \mathbb{R}^3 : x = y\}$. Show that $\dim W = 2$.

Solution We know that the standard basis of $\mathbb{R}^3$ consists of vectors $(1, 0, 0), (0, 1, 0)$ and $(0, 0, 1)$. Suppose $\alpha = (x, y, z) \in W$. Then $x = y$. Therefore

$$\alpha = (x, x, z) = x(1, 0, 0) + x(0, 1, 0) + z(0, 0, 1)$$

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

Therefore the set $\{(1, 1, 0), (0, 0, 1)\}$ generates W . Also we see that for $a, b \in \mathbb{R}$, we have

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

$$\Rightarrow (a, a, b) = (0, 0, 0)$$

$$\Rightarrow a = 0, b = 0$$

Thus the set $\{(1, 1, 0), (0, 0, 1)\}$ is a linearly independent subset of W which generates W , i.e. $\{(1, 1, 0), (0, 0, 1)\}$ is a basis of W . Hence $\dim W = 2$.

Example 2.5.6 Prove that the set $\{(1, 2), (4, 3)\}$ is a basis of $\mathbb{R}^2$.

Solution Let $\alpha_1 = (1, 2)$ and $\alpha_2 = (4, 3)$. Then for $a_1, a_2 \in \mathbb{R}$, we have

$$a_1 \alpha_1 + a_2 \alpha_2 = 0$$

$$\Rightarrow a_1(1, 2) + a_2(4, 3) = (0, 0)$$

$$\Rightarrow (a_1 + 4a_2, 2a_1 + 3a_2) = (0, 0)$$

$$\Rightarrow a_1 + 4a_2 = 0, 2a_1 + 3a_2 = 0$$

$$\Rightarrow a_1 = 0, a_2 = 0$$

i.e. the set $\{(1,2), (4,3)\}$ is linearly independent.

Now suppose $(p, q) \in \mathbb{R}^2$. If for $a, b \in \mathbb{R}$, we have

$$(p, q) = a(1,2) + b(4,3)$$

Then

$$\begin{aligned}(p, q) &= (a + 4b, 2a + 3b) \\ \Rightarrow p &= a + 4b, q = 2a + 3b \\ \Rightarrow a &= \frac{4q - 3p}{5}, b = \frac{2p - q}{5}\end{aligned}$$

Hence

$$(p, q) = \frac{4q - 3p}{5}(1,2) + \frac{2p - q}{5}(4,3)$$

Thus every vector in $\mathbb{R}^2$ can be expressed as a linear combination of $(1,2)$ and $(4,3)$, i.e. $\{(1,2), (4,3)\}$. Therefore $\{(1,2), (4,3)\}$ is a basis of $\mathbb{R}^2$.

Example 2.5.7 Expand the linearly independent subset $\{(3,4,0)\}$ of $\mathbb{R}^3$ to a basis of $\mathbb{R}^3$.

Solution In constructing a basis for $\mathbb{R}^3$, we shall follow the proof of Theorem 5. Let $S = \{(3,4,0)\}$. Then $L(S) = \{k(3,4,0) : k \in \mathbb{R}\}$. Let $\alpha \in \mathbb{R}^3$ such that $\alpha \notin L(S)$. Obviously $\alpha \neq (3k, 4k, 0)$ for any $k \in \mathbb{R}$. We may take $\alpha = (1,2,3)$. Then the set $S_1 = S \cup \{(1,2,3)\} = \{(3,4,0), (1,2,3)\}$ is a linearly independent subset of $\mathbb{R}^3$. Now

$L(S_1) = \{h(3,4,0) + k(1,2,3) : h, k \in \mathbb{R}\} = \{(3h + k, 4h + 2k, 3k) : h, k \in \mathbb{R}\}$
If we take $h = 1, k = 1$, then $(4,6,3) \in L(S_1)$. Now we require $\beta \in \mathbb{R}^3$, such that $\beta \notin L(S_1)$. Obviously $\beta = (4,6,0) \notin L(S_1)$. Therefore the set $S_2 = S_1 \cup \{(4,6,0)\} = \{(3,4,0), (1,2,3), (4,6,0)\}$ is a linearly independent subset of $\mathbb{R}^3$. Now $\dim \mathbb{R}^3 = 3$, hence S_2 is a basis of $\mathbb{R}^3$.

Here you will notice that we have other choices for the vectors α and β as well. Therefore, we can extend $\{(3,4,0)\}$ in many ways to obtain different bases of $\mathbb{R}^3$.

Example 2.5.8 Find the coordinates of the vector $(2,1,-6)$ of $\mathbb{R}^3$ with respect to the ordered basis $\{(1,1,2), (3,-1,0), (2,0,-1)\}$

Solution Let (p, q, r) be the coordinates of $(2,1,-6)$ with respect to the given basis. Then

$$\begin{aligned}(2,1,-6) &= p(1,1,2) + q(3,-1,0) + r(2,0,-1) \\ \Rightarrow (2,1,-6) &= (p + 3q + 2r, p - q, 2p - r) \\ \Rightarrow p + 3q + 2r &= 2, p - q = 1, 2p - r = -6 \\ \Rightarrow p &= -7/8, q = -15/8, r = 17/4\end{aligned}$$

Therefore the coordinates of $(2,1,-6)$ with respect to the given basis are $(-7/8, -15/8, 17/4)$.

2.6 Linear Sum and Direct sum of Subspaces

We can define addition of subsets of a vector space. Suppose $S_1, S_2, \dots, S_m$ are some subsets of a vector space V defined over a field F . Then by $S_1 + S_2 + \dots + S_m$ we mean the set of all sums $\alpha_1 + \alpha_2 + \dots + \alpha_m$ of vectors $\alpha_1 \in S_1, \alpha_2 \in S_2, \dots, \alpha_m \in S_m$.

Definition If $W_1, W_2, \dots, W_m$ are subspaces of a vector space V , then their **linear sum** $W_1 + W_2 + \dots + W_m$ is defined as the set

$$\left\{ \sum_{k=1}^m \alpha_k : \alpha_1 \in W_1, \alpha_2 \in W_2, \dots, \alpha_m \in W_m \right\}$$

For example, the linear sum of subspaces $W_1 = \{(x, y, 0) \in \mathbb{R}^3 : x, y \in \mathbb{R}\}$ and $W_2 = \{(0, y, z) \in \mathbb{R}^3 : y, z \in \mathbb{R}\}$ is $\mathbb{R}^3$, i.e. $W_1 + W_2 = \mathbb{R}^3$.

Since $\alpha_1 + \alpha_2 \in W_1 + W_2 \Rightarrow \alpha_1 \in W_1, \alpha_2 \in W_2$
 $\Rightarrow \alpha_1 \in \mathbb{R}^3, \alpha_2 \in \mathbb{R}^3$ as $W_1 \subset \mathbb{R}^3$ and $W_2 \subset \mathbb{R}^3$
 $\Rightarrow \alpha_1 + \alpha_2 \in \mathbb{R}^3$

Thus $W_1 + W_2 \subseteq \mathbb{R}^3$. Also if $\alpha = (a, b, c) \in \mathbb{R}^3$, then we can write

$$\alpha = (a, b, c) = (a, b/2, 0) + (0, b/2, c) \in W_1 + W_2$$

Hence $\mathbb{R}^3 \subseteq W_1 + W_2$. Therefore $W_1 + W_2 = \mathbb{R}^3$.

Here you will notice that if $\alpha_1 + \alpha_2 \in W_1 + W_2$, then $\alpha_1 \in W_1, \alpha_2 \in W_2$ and therefore $\alpha_1, \alpha_2 \in W_1 \cup W_2$. So we must have some relation involving $W_1 + W_2$ and $W_1 \cup W_2$. Theorem 9 explores this relationship further. First we shall show that the linear sum of two subspaces is also a subspace.

Theorem 8 Let W_1 and W_2 be any two subspaces of a vector space V . Then $W_1 + W_2$ is also a subspace of V .

Proof Let $\alpha, \beta \in W_1 + W_2$ and $a, b \in F$. Then $\alpha = \alpha_1 + \alpha_2$ and $\beta = \beta_1 + \beta_2$ where $\alpha_1, \beta_1 \in W_1$ and $\alpha_2, \beta_2 \in W_2$. Since W_1 and W_2 are subspaces, hence

$$a\alpha_1, b\beta_1 \in W_1, a, b \in F \Rightarrow a\alpha_1 + b\beta_1 \in W_1$$

and

$$a\alpha_2, b\beta_2 \in W_2, a, b \in F \Rightarrow a\alpha_2 + b\beta_2 \in W_2$$

Therefore $a\alpha + b\beta = a(\alpha_1 + \alpha_2) + b(\beta_1 + \beta_2)$

$$= (a\alpha_1 + b\beta_1) + (a\alpha_2 + b\beta_2) \in W_1 + W_2$$

Hence $W_1 + W_2$ is a subspace of V .

Theorem 9 Let W_1 and W_2 be any two subspaces of a vector space V . Then we have $W_1 + W_2 = L(W_1 \cup W_2)$.

Proof Let $\alpha \in W_1 + W_2$. Then $\alpha = \alpha_1 + \alpha_2$, where $\alpha_1 \in W_1$ and $\alpha_2 \in W_2$. Hence $\alpha_1, \alpha_2 \in W_1 \cup W_2$. Now

$$\alpha = \alpha_1 + \alpha_2 = 1\alpha_1 + 1\alpha_2 \in L(W_1 \cup W_2)$$

i.e. $\alpha \in W_1 + W_2 \Rightarrow \alpha \in L(W_1 \cup W_2)$

Therefore $W_1 + W_2 \subseteq L(W_1 \cup W_2)$.

Now suppose that $\alpha \in L(W_1 \cup W_2)$. Then α is a linear combination of elements of $W_1 \cup W_2$, i.e. linear combination of some elements of W_1 and some elements of W_2 . Assume that $\beta_1, \beta_2, \dots, \beta_k \in W_1$ and $\gamma_1, \gamma_2, \dots, \gamma_m \in W_2$ such that $\alpha = b_1\beta_1 + \dots + b_k\beta_k + c_1\gamma_1 + \dots + c_m\gamma_m$ for some $b_1, \dots, b_k, c_1, \dots, c_m \in F$. Obviously $b_1\beta_1 + \dots + b_k\beta_k \in W_1$ and $c_1\gamma_1 + \dots + c_m\gamma_m \in W_2$, and therefore $\alpha \in W_1 + W_2$. Thus $L(W_1 \cup W_2) \subseteq W_1 + W_2$ and so we have $W_1 + W_2 = L(W_1 \cup W_2)$.

Note: $L(W_1 \cup W_2) \subseteq W_1 + W_2$ is also obvious from the fact that $W_1 + W_2$ is a subspace containing $W_1 \cup W_2$ (since $W_1 \subseteq W_1 + W_2$, $W_2 \subseteq W_1 + W_2$) and $L(W_1 \cup W_2)$ is the smallest subspace containing $W_1 \cup W_2$.

So we have seen that $W_1 + W_2$ is a subspace generated by $W_1 \cup W_2$ and that an element of $W_1 + W_2$ can be written as a sum of an element of W_1 and an element of W_2 . Also we have seen that if $W_1 = \{(x, y, 0) \in \mathbb{R}^3 : x, y \in \mathbb{R}\}$ and $W_2 = \{(0, y, z) \in \mathbb{R}^3 : y, z \in \mathbb{R}\}$, then $W_1 + W_2 = \mathbb{R}^3$. Suppose $(2, 3, 4) \in \mathbb{R}^3$. Then $(2, 3, 4) = (2, 3/2, 0) + (0, 3/2, 4) \in W_1 + W_2$.

Also we can write

$$(2, 3, 4) = (2, 1, 0) + (0, 2, 4) \in W_1 + W_2$$

$$(2, 3, 4) = (2, -1, 0) + (0, 4, 4) \in W_1 + W_2$$

and so on.

Here you see that the sum is not unique and there are many ways to express $\alpha \in W_1 + W_2$ as a sum $\alpha = \alpha_1 + \alpha_2$, where $\alpha_1 \in W_1$ and $\alpha_2 \in W_2$. However in some cases you will find uniqueness in the expression $\alpha = \alpha_1 + \alpha_2$. For example, if we take $W_3 = \{(0, 0, z) \in \mathbb{R}^3 : z \in \mathbb{R}\}$, then $W_1 + W_3 = \mathbb{R}^3$ and every $(a, b, c) \in \mathbb{R}^3$ can be uniquely expressed as

$$(a, b, c) = (a, b, 0) + (0, 0, c) \in W_1 + W_3$$

This sum $W_1 + W_3$ is then called the direct sum of the subspaces W_1 and W_3 and is denoted by $W_1 \oplus W_3$. Hence we have the following definition:

Definition Let $W_1, W_2, \dots, W_m$ be the subspaces of a vector space V . Then the sum $W_1 + W_2 + \dots + W_m$ is said to be **direct** if any vector in the sum can be expressed in a unique way as $\alpha_1 + \alpha_2 + \dots + \alpha_m$ with $\alpha_1 \in W_1, \alpha_2 \in W_2, \dots, \alpha_m \in W_m$. We denote the direct sum of subspaces $W_1, W_2, \dots, W_m$ as $W_1 \oplus W_2 \oplus \dots \oplus W_m$.

Theorem 10 Let W_1 and W_2 be any two subspaces of a vector space V . Then the following statements are equivalent:

- (i) The sum $W_1 + W_2$ of subspaces W_1 and W_2 is direct,
- (ii) $W_1 \cap W_2 = \{0\}$

(iii) If $0 \neq \alpha \in W_1$ and $0 \neq \beta \in W_2$, then α and β are linearly independent,

(iv) If $\alpha \in W_1$ and $\beta \in W_2$ such that $\alpha + \beta = 0$, then $\alpha = 0$ and $\beta = 0$.

Proof (i) $\Rightarrow$ (ii) Suppose the sum $W_1 + W_2$ of subspaces W_1 and W_2 is direct. Let $\alpha \in W_1 \cap W_2$. Then $\alpha \in W_1$ and $\alpha \in W_2$.

Now $\alpha = \alpha + 0 \in W_1 + W_2$ where $\alpha \in W_1$ and $0 \in W_2$

Also $\alpha = 0 + \alpha \in W_1 + W_2$ where $0 \in W_1$ and $\alpha \in W_2$

Since $W_1 + W_2$ is direct, hence α cannot be written in two different ways as a sum of elements of W_1 and W_2 . Therefore we must have $\alpha = 0$. Thus $W_1 \cap W_2 \subseteq \{0\}$. Also $0 \in W_1$ and $0 \in W_2 \Rightarrow 0 \in W_1 \cap W_2$, i.e. $\{0\} \subseteq W_1 \cap W_2$. Therefore $W_1 \cap W_2 = \{0\}$.

(ii) $\Rightarrow$ (iii) Suppose $W_1 \cap W_2 = \{0\}$. Let $0 \neq \alpha \in W_1$ and $0 \neq \beta \in W_2$ and $a, b \in F$. Then $a\alpha + b\beta = 0 \Rightarrow a\alpha = -b\beta$. Now $a\alpha \in W_1$ and $-b\beta \in W_2$, therefore $a\alpha = -b\beta \in W_1 \cap W_2$. Hence $a\alpha = 0$ and $b\beta = 0$. Since $\alpha \neq 0$ and $\beta \neq 0$, hence $a = 0, b = 0$. Thus α and β are linearly independent.

(iii) $\Rightarrow$ (iv) Suppose $\alpha \in W_1$ and $\beta \in W_2$ such that $\alpha + \beta = 0$. Then α and β are linearly dependent. By (iii), if $\alpha \neq 0$ and $\beta \neq 0$, then α and β are linearly independent. Therefore we must have $\alpha = 0$ and $\beta = 0$.

(iv) $\Rightarrow$ (i) Suppose $\alpha \in W_1 + W_2$ such that $\alpha = \alpha_1 + \alpha_2 = \beta_1 + \beta_2$ where $\alpha_1, \beta_1 \in W_1$ and $\alpha_2, \beta_2 \in W_2$. This gives $(\alpha_1 - \beta_1) + (\alpha_2 - \beta_2) = 0$. By (iv), we have $\alpha_1 - \beta_1 = 0$ and $\alpha_2 - \beta_2 = 0$, i.e. $\alpha_1 = \beta_1, \alpha_2 = \beta_2$. Hence the sum $W_1 + W_2$ of subspaces W_1 and W_2 is direct.

Definition Let W be a subspace of a vector space V . Then a subspace W' is said to be a complement of W if $W \oplus W' = V$.

For example, $W = \{(x, y, 0) \in \mathbb{R}^3 : x, y \in \mathbb{R}\}$ is a complement of $W' = \{(0, 0, z) \in \mathbb{R}^3 : z \in \mathbb{R}\}$, since $W \oplus W' = \mathbb{R}^3$.

Clearly, if W' is a complement of W , then W is a complement of W' . Then we say that W and W' are complementary subspaces.

Definition Let W_1 be a complement of W_2 and $\alpha \in V$. Let $\alpha = \alpha_1 + \alpha_2$ such that $\alpha_1 \in W_1$ and $\alpha_2 \in W_2$. Then α_1 is called the projection of α into W_1 along W_2 .

Obviously, if α_1 is the projection of α into W_1 along W_2 , then $\alpha - \alpha_1$ is the projection of α into W_2 along W_1 .

For example, Let $V = \mathbb{R}^2$. Then $W_1 = \{(x, 0) \in \mathbb{R}^2 : x \in \mathbb{R}\}$ is a complement of $W_2 = \{(x, y) \in \mathbb{R}^2 : x = y\}$. The projection of $(1, -1) \in \mathbb{R}^2$ into W_1 along W_2 is $(2, 0)$ since $(1, -1) = (2, 0) + (-1, -1)$.

Here you will notice that $\alpha = \alpha_1 + \alpha_2$ exists and unique. Therefore the projection of α into W_1 along W_2 is well defined.

Theorem 11 Let W_1 and W_2 be any two finite-dimensional subspaces of a vector space V . Then

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

Proof Let $\dim (W_1 \cap W_2) = k$ and $S = \{\alpha_1, \alpha_2, \dots, \alpha_k\}$ be a basis of $W_1 \cap W_2$. Then S is a linearly independent subset of W_1 as well as W_2 . Therefore S can be extended to a basis $\{\alpha_1, \alpha_2, \dots, \alpha_k, \beta_1, \beta_2, \dots, \beta_m\}$ of W_1 . Similarly S can be extended to a basis $\{\alpha_1, \alpha_2, \dots, \alpha_k, \gamma_1, \gamma_2, \dots, \gamma_n\}$ of W_2 . Hence $\dim W_1 = k + m$ and $\dim W_2 = k + n$.

Now the set $S' = \{\alpha_1, \alpha_2, \dots, \alpha_k, \beta_1, \beta_2, \dots, \beta_m, \gamma_1, \gamma_2, \dots, \gamma_n\}$ generates the subspace $W_1 + W_2$, since if $\alpha \in W_1 + W_2$, then

$$\begin{aligned} \alpha &= \text{an element of } W_1 + \text{an element of } W_2 \\ &= \text{A linear combination of vectors } \alpha_1, \alpha_2, \dots, \alpha_k, \beta_1, \beta_2, \dots, \beta_m \\ &\quad + \text{A linear combination of vectors } \alpha_1, \alpha_2, \dots, \alpha_k, \gamma_1, \gamma_2, \dots, \gamma_n \\ &= \text{A linear combination of vectors} \\ &\quad \alpha_1, \alpha_2, \dots, \alpha_k, \beta_1, \beta_2, \dots, \beta_m, \gamma_1, \gamma_2, \dots, \gamma_n \end{aligned}$$

Also the set S' is linearly independent. For suppose

$$\sum_{i=1}^k a_i \alpha_i + \sum_{j=1}^m b_j \beta_j + \sum_{r=1}^n c_r \gamma_r = 0 \dots \dots (4)$$

Where $a_i, b_j, c_r \in F$ for all i, j, r . Then

$$\sum_{i=1}^k a_i \alpha_i + \sum_{j=1}^m b_j \beta_j = - \sum_{r=1}^n c_r \gamma_r$$

Since $\sum_{i=1}^k a_i \alpha_i + \sum_{j=1}^m b_j \beta_j \in W_1$ and $-\sum_{r=1}^n c_r \gamma_r \in W_2$, therefore

$$-\sum_{r=1}^n c_r \gamma_r \in W_1 \cap W_2$$

Hence we can write

$$\begin{aligned} -\sum_{r=1}^n c_r \gamma_r &= \sum_{i=1}^k d_i \alpha_i \\ \Rightarrow \sum_{i=1}^k d_i \alpha_i + \sum_{r=1}^n c_r \gamma_r &= 0 \end{aligned}$$

Since the vectors $\alpha_1, \alpha_2, \dots, \alpha_k, \gamma_1, \gamma_2, \dots, \gamma_n$ form a basis of W_2 , hence we have $c_1 = 0, c_2 = 0, \dots, c_n = 0$; $d_1 = 0, \dots, d_k = 0$.

From (4), we have

$$\sum_{i=1}^k a_i \alpha_i + \sum_{j=1}^m b_j \beta_j = \mathbf{0}$$

Also the vectors $\alpha_1, \alpha_2, \dots, \alpha_k, \beta_1, \beta_2, \dots, \beta_m$ form a basis of W_1 and therefore

$$a_1 = 0, a_2 = 0, \dots, a_k = 0; b_1 = 0, \dots, b_m = 0$$

Thus the set S' is linearly independent and therefore forms a basis of $W_1 + W_2$.

$$\begin{aligned} \therefore \dim(W_1 + W_2) &= k + m + n \\ \Rightarrow \dim(W_1 + W_2) + \dim(W_1 \cap W_2) &= (k + m + n) + k \\ &= (k + m) + (k + n) \\ &= \dim W_1 + \dim W_2 \end{aligned}$$

Corollary If the sum $W_1 + W_2$ of subspaces W_1 and W_2 is direct, then $W_1 \cap W_2 = \{\mathbf{0}\}$. Therefore $\dim(W_1 \cap W_2) = 0$ and hence

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

2.7 Quotient Space

In the section 2.6 we have defined the linear sum of subsets of a given vector space V . If $A \subseteq V$ and $B \subseteq V$, then we define $A + B = \{a + b : a \in A, b \in B\}$. Suppose $A = \{\alpha\}$ is a singleton. Then we have

$$\{\alpha\} + B = \{\alpha + b : b \in B\}$$

We write $\alpha + B$ for $\{\alpha\} + B$. The set $\alpha + B$ is called a translate of B . If we take B as a subspace of V , then we get special sets called cosets. So we have the following definition:

Definition Let V be a vector space over a field F and W be a subspace of V . Then a coset of W in V is a subset $\alpha + W$ of V defined as

$$\alpha + W := \{\alpha + \gamma : \gamma \in W\}$$

For example, consider the subspace $W = \{(x, y, 0) \in \mathbb{R}^3 : x, y \in \mathbb{R}\}$ of $\mathbb{R}^3$. Let

$\alpha = (a, b, c) \in \mathbb{R}^3$. Then

$$\begin{aligned} \alpha + W &= \{(a + x, b + y, c) \in \mathbb{R}^3 : x, y \in \mathbb{R}\} \\ &= \{(x', y', c) \in \mathbb{R}^3 : x', y' \in \mathbb{R}\} \end{aligned}$$

Now W represents xy -plane and $\alpha + W$ or $(a, b, c) + W$ represents the plane parallel to xy -plane through $\alpha = (a, b, c)$.

Interestingly, we can add any two cosets to obtain a new coset. Also we can multiply a coset by a scalar. These two new operations defined on the collection of cosets facilitate the construction of a new type of vector space called a quotient space. Before defining a quotient space let us first discuss some properties of cosets.

Theorem 12 Let W be a subspace of a vector space V . Then for $\alpha, \beta \in V$, we have

$$(i) \alpha \in W \Leftrightarrow \alpha + W = W$$

$$(ii) \alpha + W = \beta + W \Leftrightarrow \alpha - \beta \in W$$

$$(iii) \text{ Either } \alpha + W = \beta + W \text{ or } (\alpha + W) \cap (\beta + W) = \emptyset,$$

Proof (i) Let $\alpha \in W$. If $\gamma \in \alpha + W$, then $\gamma = \alpha + \mu$ for some $\mu \in W$. Since $\alpha, \mu \in W$, hence $\gamma = \alpha + \mu \in W$. Thus $\alpha + W \subseteq W$.

Now if $\delta \in W$, then $\delta = \alpha + (\delta - \alpha) \in \alpha + W$ as $\delta - \alpha \in W$. Thus $W \subseteq \alpha + W$. This gives $\alpha + W = W$

Conversely, suppose that $\alpha + W = W$. Since $0 \in W$, hence

$$\alpha + 0 \in \alpha + W = W$$

$$\Rightarrow \alpha \in W$$

$$(ii) \text{ We have } \alpha = \alpha + 0 \in \alpha + W$$

$$\text{Therefore } \alpha + W = \beta + W \Rightarrow \alpha \in \beta + W$$

$$\Rightarrow \alpha = \beta + \gamma \text{ for some } \gamma \in W$$

$$\Rightarrow \gamma = \alpha - \beta \in W$$

$$\text{Thus } \alpha + W = \beta + W \Rightarrow \alpha - \beta \in W$$

Conversely, suppose that $\alpha - \beta \in W$. Let $\delta \in \alpha + W$. Then $\delta = \alpha + \gamma$ for some $\gamma \in W$. Now

$$\delta = \beta + [(\alpha - \beta) + \gamma]$$

$$\text{But } \alpha - \beta \in W, \gamma \in W \Rightarrow (\alpha - \beta) + \gamma \in W$$

$$\Rightarrow \delta = \beta + [(\alpha - \beta) + \gamma] \in \beta + W$$

$$\text{Thus } \delta \in \alpha + W \Rightarrow \delta \in \beta + W, \text{ i.e. } \alpha + W \subseteq \beta + W.$$

Similarly, we can show that $\beta + W \subseteq \alpha + W$. Therefore $\alpha + W = \beta + W$.

(iii) Suppose that $(\alpha + W) \cap (\beta + W) \neq \emptyset$. Then there exists $\delta \in V$ such that $\delta \in \alpha + W$ and $\delta \in \beta + W$. Hence there exists $\gamma \in W$ such that $\delta = \alpha + \gamma$. Also there exists $\mu \in W$ such that $\delta = \beta + \mu$. Thus $\delta = \alpha + \gamma = \beta + \mu$, i.e. $\alpha - \beta = \mu - \gamma \in W$. Hence by (ii), we have $\alpha + W = \beta + W$. ■

Now we are in position to define the addition (vector addition) of two cosets and the multiplication (scalar multiplication) of a coset by a scalar. Suppose V/W denotes the collection of all cosets of W in V , i.e.

$$V/W = \{\alpha + W : \alpha \in V\}$$

Let $\alpha + W, \beta + W \in V/W$ and $c \in F$. Define

$$(\alpha + W) + (\beta + W) = (\alpha + \beta) + W$$

$$c(\alpha + W) = (c\alpha) + W$$

Since different vectors in V can produce same coset of W in V , a natural question arises. Are the vector addition and scalar multiplication well defined? In other words, if $\alpha + W = \alpha' + W$ and $\beta + W = \beta' + W$, do we have $(\alpha + W) + (\beta + W) = (\alpha' + W) + (\beta' + W)$ and $c(\alpha + W) = c(\alpha' + W)$?

The answer is yes. Let us see how.

$$\alpha + W = \alpha' + W \Rightarrow \alpha - \alpha' \in W$$

and

$$\beta + W = \beta' + W \Rightarrow \beta - \beta' \in W$$

Now $\alpha - \alpha' \in W, \beta - \beta' \in W \Rightarrow (\alpha - \alpha') + (\beta - \beta') \in W$

$$\Rightarrow (\alpha + \beta) - (\alpha' + \beta') \in W$$

$$\Rightarrow (\alpha + \beta) + W = (\alpha' + \beta') + W$$

$$\Rightarrow (\alpha + W) + (\beta + W) = (\alpha' + W) + (\beta' + W)$$

Therefore the vector addition is well defined in V/W .

Now $c \in F, \alpha - \alpha' \in W \Rightarrow c(\alpha - \alpha') \in W$

$$\Rightarrow c\alpha - c\alpha' \in W$$

$$\Rightarrow c\alpha + W = c\alpha' + W$$

$$\Rightarrow c(\alpha + W) = c(\alpha' + W)$$

Hence the scalar multiplication is also well defined in V/W .

In the next theorem, you will observe how V/W forms a vector space over the field F with these two operations, i.e. vector addition and scalar multiplication defined above.

Theorem 13 Let W be a subspace of a vector space V over a field F . Let V/W denote the set of all distinct cosets of W in V . Then V/W is a vector space over F with the vector addition and scalar multiplication defined as follows

$$(\alpha + W) + (\beta + W) = (\alpha + \beta) + W$$

$$\text{and } c(\alpha + W) = (c\alpha) + W$$

Where $\alpha + W, \beta + W \in V/W$ and $c \in F$.

Proof We have already proved that the operations of vector addition and scalar multiplication are well-defined.

(1) Let $\alpha + W, \beta + W \in V/W$, then

$$(\alpha + W) + (\beta + W) = (\alpha + \beta) + W$$

Since $\alpha, \beta \in V \Rightarrow \alpha + \beta \in V$, hence $(\alpha + \beta) + W \in V/W$

(2) Let $\alpha + W, \beta + W, \gamma + W \in V/W$. Since we have

$$\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma \text{ for all } \alpha, \beta, \gamma \in V$$

Therefore

$$\begin{aligned} (\alpha + W) + [(\beta + W) + (\gamma + W)] &= (\alpha + W) + [(\beta + \gamma) + W] \\ \Rightarrow &= [\alpha + (\beta + \gamma)] + W \\ \Rightarrow &= [(\alpha + \beta) + \gamma] + W \\ \Rightarrow &= [(\alpha + \beta) + W] + (\gamma + W) \\ \Rightarrow &= [(\alpha + W) + (\beta + W)] + (\gamma + W) \end{aligned}$$

Hence the associativity of addition on V/W is proved.

(3) Since $0 \in V$, hence $0 + W \in V/W$. Now $0 \in W$, therefore $0 + W = W$. You will observe that W acts as the zero (identity element) of V/W . For any $\alpha + W \in V/W$, we have

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

Also

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

(4) Now $\alpha \in V \Rightarrow -\alpha \in V$, such that $\alpha + (-\alpha) = \mathbf{0} = (-\alpha) + \alpha$. Thus for every $\alpha + W \in V/W$, we have $(-\alpha) + W \in V/W$ such that

$$(\alpha + W) + [(-\alpha) + W] = [\alpha + (-\alpha)] + W = \mathbf{0} + W = W$$

and

$$[(-\alpha) + W] + (\alpha + W) = [(-\alpha) + \alpha] + W = \mathbf{0} + W = W$$

Hence the additive inverse of $\alpha + W$ is $(-\alpha) + W$.

(5) Since V is a vector space, hence

$$\alpha + \beta = \beta + \alpha \text{ for all } \alpha, \beta \in V$$

For any $\alpha + W, \beta + W \in V/W$, we have

$$\begin{aligned}(\alpha + W) + (\beta + W) &= (\alpha + \beta) + W \\&= (\beta + \alpha) + W \\&= (\beta + W) + (\alpha + W)\end{aligned}$$

Thus V/W is an abelian group under the addition defined above.

(6) Now for $c \in F$ and $\alpha \in V$, we have $c\alpha \in V$. Therefore for any $c \in F$ and $\alpha + W \in V/W$, we have $c(\alpha + W) = (c\alpha) + W \in V/W$

(7) For any $a, b \in F$ and $\alpha + W, \beta + W \in V/W$, we have

$$\begin{aligned}\text{(i) } a[(\alpha + W) + (\beta + W)] &= a[(\alpha + \beta) + W] \\&= a(\alpha + \beta) + W \\&= (a\alpha + a\beta) + W \\&= (a\alpha + W) + (a\beta + W) \\&= a(\alpha + W) + a(\beta + W)\end{aligned}$$

$$\begin{aligned}\text{(ii) } (a + b)(\alpha + W) &= (a + b)\alpha + W \\&= (a\alpha + b\alpha) + W \\&= (a\alpha + W) + (b\alpha + W) \\&= a(\alpha + W) + b(\alpha + W)\end{aligned}$$

$$\begin{aligned}\text{(iii) } (ab)(\alpha + W) &= (ab)\alpha + W \\&= a(b\alpha) + W = a[(b\alpha + W)] = a[b(\alpha + W)]\end{aligned}$$

$$\text{(iv) } 1.(\alpha + W) = (1.\alpha) + W = \alpha + W$$

Thus V/W is a vector space over F . This vector space is called the **quotient space** of V by W .

Let us now obtain the dimension of quotient space V/W .

Theorem 14 Let W be a subspace of a finite-dimensional vector space V over a field F . Then $\dim V/W = \dim V - \dim W$.

Proof Let $\dim W = k$. Then any basis of W has k elements. Let $\{\alpha_1, \alpha_2, \dots, \alpha_k\}$ be a basis of W . We can extend this to a basis of V . Let $\{\alpha_1, \alpha_2, \dots, \alpha_k, \beta_1, \beta_2, \dots, \beta_m\}$ be a basis of V . Then $\dim V = k + m$. Therefore $\dim V - \dim W = k + m - k = m$.

We claim that $\{\beta_1 + W, \beta_2 + W, \dots, \beta_m + W\}$ is a basis of V/W . To prove that $\{\beta_1 + W, \beta_2 + W, \dots, \beta_m + W\}$ generates V/W , consider $\alpha + W \in V/W$. Then $\alpha \in V$ and we can write

$$\begin{aligned}\alpha &= a_1\alpha_1 + \dots + a_k\alpha_k + b_1\beta_1 + \dots + b_m\beta_m \\ \Rightarrow \alpha + W &= (a_1\alpha_1 + \dots + a_k\alpha_k + b_1\beta_1 + \dots + b_m\beta_m) + W \\ \Rightarrow \alpha + W &= [(a_1\alpha_1 + \dots + a_k\alpha_k) + W] + [(b_1\beta_1 + \dots + b_m\beta_m) + W]\end{aligned}$$

Now $a_1\alpha_1 + \dots + a_k\alpha_k \in W$. Hence $(a_1\alpha_1 + \dots + a_k\alpha_k) + W = W$.

$$\begin{aligned}\therefore \alpha + W &= W + [(b_1\beta_1 + \dots + b_m\beta_m) + W] \\ &= (b_1\beta_1 + \dots + b_m\beta_m) + W \\ &= b_1(\beta_1 + W) + \dots + b_m(\beta_m + W)\end{aligned}$$

Thus every element of V/W can be written as a linear combination of the vectors $\beta_1 + W, \beta_2 + W, \dots, \beta_m + W$.

To show that the vectors $\beta_1 + W, \beta_2 + W, \dots, \beta_m + W$ are linearly independent, let $c_1, c_2, \dots, c_m \in F$. Then

$$\begin{aligned}c_1(\beta_1 + W) + \dots + c_m(\beta_m + W) &= W \\ \Rightarrow (c_1\beta_1 + W) + \dots + (c_m\beta_m + W) &= W \\ \Rightarrow (c_1\beta_1 + \dots + c_m\beta_m) + W &= W \\ \Rightarrow c_1\beta_1 + \dots + c_m\beta_m &\in W\end{aligned}$$

Since $\{\alpha_1, \alpha_2, \dots, \alpha_k\}$ is a basis of W , hence we can express $c_1\beta_1 + \dots + c_m\beta_m$ as a linear combination of vectors $\alpha_1, \alpha_2, \dots, \alpha_k$, i.e. there exist scalars $d_1, d_2, \dots, d_k \in F$ such that

$$c_1\beta_1 + \dots + c_m\beta_m = d_1\alpha_1 + \dots + d_k\alpha_k$$

$$\Rightarrow c_1\beta_1 + \dots + c_m\beta_m - d_1\alpha_1 - \dots - d_k\alpha_k = 0$$

Now $\{\alpha_1, \alpha_2, \dots, \alpha_k, \beta_1, \beta_2, \dots, \beta_m\}$ is a basis of V , hence $c_1 = 0, \dots, c_m = 0$, $d_1 = 0, \dots, d_k = 0$. Thus

$$c_1(\beta_1 + W) + \dots + c_m(\beta_m + W) = W \Rightarrow c_1 = 0, \dots, c_m = 0$$

Hence the set $\{\beta_1 + W, \beta_2 + W, \dots, \beta_m + W\}$ is linearly independent. Thus $\{\beta_1 + W, \beta_2 + W, \dots, \beta_m + W\}$ forms a basis of V/W and we have $\dim V/W = m$. Hence $\dim V/W = \dim V - \dim W$.

2.8 Illustrative Examples

Example 2.8.1 Let $W_1 = \{(x, y, 0) \in \mathbb{R}^3 : x, y \in \mathbb{R}\}$ and $W_2 = \{(0, y, z) \in \mathbb{R}^3 : y, z \in \mathbb{R}\}$ be the subspaces of $\mathbb{R}^3$. Find the dimension of $W_1 \cap W_2$.

Solution We know that the set $\{(1, 0, 0), (0, 1, 0)\}$ is a basis of W_1 and the set $\{(0, 1, 0), (0, 0, 1)\}$ is a basis of W_2 . Hence $\dim W_1 = 2$ and $\dim W_2 = 2$. Now $W_1 + W_2 = \mathbb{R}^3$, therefore $\dim (W_1 + W_2) = \dim \mathbb{R}^3 = 3$.

From Theorem 11, we have

$$\begin{aligned}\dim (W_1 \cap W_2) &= \dim W_1 + \dim W_2 - \dim (W_1 + W_2) \\ &= 2 + 2 - 3 = 1\end{aligned}$$

Example 2.8.2 Let $W = \{(x, 0, 0) \in \mathbb{R}^3 : x \in \mathbb{R}\}$. Find the dimension of $\mathbb{R}^3/W$.

Solution We have $\dim \mathbb{R}^3/W = \dim \mathbb{R}^3 - \dim W$. Since $\dim \mathbb{R}^3 = 3$ and $\dim W = 1$. Hence $\dim \mathbb{R}^3/W = 3 - 1 = 2$.

Example 2.8.3 Let $W_1 = \{(x, y, z) \in \mathbb{R}^3 : x = y = z\}$ and $W_2 = \{(0, y, z) \in \mathbb{R}^3 : y, z \in \mathbb{R}\}$ be the subspaces of $\mathbb{R}^3$. Show that $W_1 \oplus W_2 = \mathbb{R}^3$.

Solution Let $(x, y, z) \in \mathbb{R}^3$. Then we can write

$$(x, y, z) = (x, x, x) + (0, y - x, z - x)$$

Since $(x, x, x) \in W_1$ and $(0, y - x, z - x) \in W_2$, hence $W_1 + W_2 = \mathbb{R}^3$.

Now $W_1 \cap W_2 = \{0\}$. Therefore, by theorem 10, we have $W_1 \oplus W_2 = \mathbb{R}^3$.

Example 2.8.4 Let V be a vector space over a field F . Find V/V and $V/\{0\}$

Solution we have $V/V = \{\alpha + V : \alpha \in V\}$. By theorem 12 (i), we have

$$\alpha \in V \Leftrightarrow \alpha + V = V$$

Hence $V/V = V$.

Now $V/\{0\} = \{\alpha + \{0\} : \alpha \in V\} = \{\alpha : \alpha \in V\} = V$.

Example 2.8.5 Let $\mathbb{P}$ be the vector space of all real polynomials in x and $W = \{f \in \mathbb{P} : f(0) = 0\}$. Find $\mathbb{P}/W$.

Solution we have $\mathbb{P}/W = \{f + W : f \in \mathbb{P}\}$. If $f \in \mathbb{P}$, then we can write

$$f(x) = a_0 + a_1x + \dots + a_nx^n, a_i \in \mathbb{R}$$

Suppose $g(x) = a_1x + \dots + a_nx^n$. Then $g(0) = 0$ and hence $g \in W$.

Thus we have $f = a_0 + g$ and

$$f + W = (a_0 + g) + W = a_0 + (g + W) = a_0 + W$$

Hence $\mathbb{P}/W = \{a_0 + W : a_0 \in \mathbb{R}\}$.

2.9 Summary

In this unit we have covered the following concepts

- (1) We have defined and studied the concept of a basis for a vector space.
- (2) We have proved various results concerning the characterization of a basis, existence of a basis and construction of a basis.
- (3) We have defined the concept of dimension for a vector space and proved results for finite-dimensional vector spaces.
- (4) We have defined and discussed the concept of coordinates. We also defined linear sum and direct sum of subspaces.
- (5) We have defined and discussed the concept of cosets and a quotient space.

2.10. Terminal questions

- (1) Show that the set $S = \{(1,0), (i,0), (0,1), (0,i)\}$ is a basis of $\mathbb{C}^2(\mathbb{R})$.
- (2) Show that the set $S = \{(1,2,1), (2,1,0), (1,-1,2)\}$ is a basis of $\mathbb{R}^3$.
- (3) Let $M_{2 \times 2}(\mathbb{R})$ be the vector space of all square 2×2 matrices with real entries. Show that the set $S = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \right\}$ forms a basis of $M_{2 \times 2}(\mathbb{R})$.
- (4) Show that the dimension of the space $\mathbb{Q}(\sqrt{2}, \sqrt{3})$ over $\mathbb{Q}$ is 4.
[HINT: $\mathbb{Q}(\sqrt{2}, \sqrt{3}) = \{a + b\sqrt{2} + c\sqrt{3} + d\sqrt{2}\sqrt{3} : a, b, c, d \in \mathbb{Q}\}$]
- (5) Prove that the space $\mathbb{R}$ over $\mathbb{Q}$ is infinite dimensional.
- (6) Let W be a subspace of $\mathbb{R}^4$ spanned by the vectors $(1, -2, 5, -3)$, $(2, 3, 1, -4)$ and $(3, 8, -3, -5)$. Show that $\dim W = 2$.
- (7) Let $M_{n \times n}(\mathbb{R})$ be the vector space of all square $n \times n$ matrices with real entries. Let W_1 and W_2 be the subspaces of symmetric and skew-symmetric matrices respectively. Show that $M_{n \times n}(\mathbb{R}) = W_1 \oplus W_2$.
- (8) Find the coordinates of the vector $(3 + 4i, 6i, 3 + 7i)$ in $\mathbb{C}^3(\mathbb{C})$ with respect to the ordered basis $\{(1,0,0), (1,1,0), (1,1,1)\}$. [Ans: $(3 - 2i, -3 - i, 3 + 7i)$]
- (9) If W is a subspace of a vector space V , the union of all the cosets of W in V is V .

(10) Let $\mathbb{P}_n$ be the vector space of all real polynomials in x with degree less than or equal to n . Find $\mathbb{P}_3/\mathbb{P}_2$. [Ans: $\{a_3x^3 + \mathbb{P}_2 : a_3 \in \mathbb{R}\}$]

2.11 Further readings

(1) Hoffman, K. and Kunze, R. (2015): Linear Algebra, Pearson India Education Services Ltd.

(2) Kumaresan, S. (2014): Linear Algebra: A Geometric approach, New Delhi, Prentice Hall of India.

(3) Strang, G. (2006): Lineart Algebra and its applications, New Delhi, Cengage Learning India Pvt Ltd.

(4) Lang, S. (2014): Introduction to Linear Algebra, New Delhi, Springer India Pvt. Ltd.

UNIT -3 Linear Transformation -I

Structure

- 3.1 Introduction**
- 3.2 Objectives**
- 3.3 Linear Transformations**
- 3.4 Illustrative Examples**
- 3.5 properties of Linear Transformations**
- 3.6 Null space and Range of a linear Transformation**
- 3.7 Rank -nullity theorem (Sylvester law)**
- 3.8 Illustrative examples**
- 3.9 Isomorphism of vector spaces**
- 3.10 Fundamental theorem of vector space homomorphism**
- 3.11 Invert isle and non-singular trans formations**
- 3.12 Summary**
- 3.13 Terminal Questions**
- 3.14 Further readings**

3.1 Introduction

In this unit, we are going to introduce a Concept this establishes a relationship between two given vector spaces defines over the same field. You are already familiar with the concept of group homomorphism. Here we translate this idea to the vector spaces and define linear transformations as the homomorphism of vector spaces.

By a transformation of a variable x . We mean a function f which yields a new variable y such that $y = f(x)$. If the variables x and y are related to each other by some linear relations, then the transformations are called linear.

To illustrate the concept further, Consider two Variables X and Y in $\mathbb{R}^2$. These two variables may be thought of as ordered pairs of real numbers, i.e. $X=(x_1, x_2), Y=(y_1, y_2)$. Sometimes the matrix representation is helpful so we take X and Y as $X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} Y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$

Sometimes the two vectors X and Y in Euclidean plane $\mathbb{R}^2$ are Connected by the following linear relations.

$$2x_1 + 4x_2 = y_1$$

$$3x_1 - 5x_2 = y_2$$

In matrix form, we can write

$$\begin{bmatrix} 2 & 4 \\ 3 & -5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \text{ or } AX=Y \dots\dots\dots(1)$$

$$\text{where } A = \begin{bmatrix} 2 & 4 \\ 3 & -5 \end{bmatrix}$$

Equation (1) gives a transformation which moves the point X to a point Y in $\mathbb{R}^2$. Since the relationship is linear, This is a linear transformation. we may Consider it as a mapping $T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ given by.

$$T(X) = AX \forall X \in \mathbb{R}^2$$

suppose we have two points X_1 and X_2 in $\mathbb{R}^2$ Such that

$$X_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \text{ And } X_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \text{ then}$$

$$X_1 + X_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$

$$\text{Therefore } T(X_1 + X_2) = A(X_1 + X_2) = \begin{bmatrix} 2 & 4 \\ 3 & -5 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 14 \\ -12 \end{bmatrix}$$

$$\text{Also we have } T(X_1) = AX_1 = \begin{bmatrix} 2 & 4 \\ 3 & -5 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 10 \\ -7 \end{bmatrix}$$

$$T(X_2) = AX_2 = \begin{bmatrix} 2 & 4 \\ 3 & -5 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ -5 \end{bmatrix}$$

$$\text{Now } T(X_1) + T(X_2) = \begin{bmatrix} 10 \\ -7 \end{bmatrix} + \begin{bmatrix} 4 \\ -5 \end{bmatrix} = \begin{bmatrix} 14 \\ -12 \end{bmatrix}$$

Thus we see that

$$T(X_1 + X_2) = T(X_1) + T(X_2)$$

$$\text{Also if we take } C=2, \text{ then } CX_1 = 2 \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$$

$$\therefore T(CX_1) = A(CX_1) = \begin{bmatrix} 2 & 4 \\ 3 & -5 \end{bmatrix} \begin{bmatrix} 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 20 \\ -14 \end{bmatrix}$$

$$\text{And } CT(X_1) = 2 \begin{bmatrix} 10 \\ -7 \end{bmatrix} = \begin{bmatrix} 20 \\ -14 \end{bmatrix}$$

$$\text{Hence } T(CX_1) = CT(X_1)$$

So we see that such transformations satisfying the following properties

$$T(X_1 + X_2) = T(X_1) + T(X_2)$$

$$\text{and } T(CX) = CT(X)$$

Therefore the linear transformations preserve the operation of addition and multiplication by scalar. We can generalize this idea to arbitrary vector space and think of linear transformations as mappings between vector spaces defined over the same fields which preserve vector addition and scalar multiplication. In this unit, we shall study linear Transformation in details and discuss related Concepts such as null space, range, rank and nullity.

3.2 Objectives

After reading this unit, you will be able to

- define a linear transformation.
- discuss various properties of vector spaces.
- define null-space, range, rank and nullity for a given linear transformation.

- define isomorphism of vector spaces and prove the fundamental theorem of vector space homomorphism.
- discuss invertible and non-singular transformations

3.3 Linear Transformations

Definition :- Let V and W be any two vector spaces over the field F . Then a linear transformation from V into W is a mapping $T: V \rightarrow W$ such that

- I. $T(\alpha + \beta) = T(\alpha) + T(\beta)$
- II. $T(C\alpha) = CT(\alpha)$

for all $\alpha, \beta \in V$ And all $C \in F$.

These two conditions may be combined into a single condition

$$i.e. T(a\alpha + b\beta) = aT(\alpha) + bT(\beta) \dots \dots \dots (2)$$

for all $\alpha, \beta \in V$ And all $C \in F$.

A linear transformation from V to itself is called a linear operator on V .

We have already seen that the mapping $T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ given by (1) is an example of linear transformation.

You can verify yourself that the identity transformation,

$$I: V \rightarrow V \text{ Given by } I(\alpha) = \alpha \quad \forall \alpha \in V \text{ is a}$$

linear transformation.

Also the map $\hat{0}: V \rightarrow W$ defined by $\hat{0}(\alpha) = 0 \in W$

is a linear transformation. This is called the Zero transformation or null transformation.

Let us discuss some more examples of linear transformations, in order to understand this concept better.

3.4 Illustrative Examples

Example 3.4.1 : Show that the mapping $T: \mathbb{R}^3 \rightarrow \mathbb{R}^2$ Defined by

$$T(a, b, c) = (a, b)$$

is a linear transformation of $\mathbb{R}^3$ onto $\mathbb{R}^2$

Solution : Let $\alpha, \beta \in \mathbb{R}^3$ such that

$$\alpha = (a_1, b_1, c_1) \text{ and}$$

$$\beta = (a_2, b_2, c_2). \text{ Suppose } a, b \in \mathbb{R}, \quad \text{then}$$

$$a\alpha + b\beta = a(a_1, b_1, c_1) + b(a_2, b_2, c_2)$$

$$(aa_1, ab_1, ac_1) + (ba_2, bb_2, bc_2)$$

$$(aa_1 + ba_2, ab_1 + bb_2, ac_1 + bc_2)$$

$$\text{Now } T(\alpha) = (a_1, b_1), \quad T(\beta) = (a_2, b_2) \text{ and}$$

$$T(a\alpha + b\beta) = (aa_1 + ba_2, ab_1 + bb_2)$$

$$= (aa_1, ab_1) + (ba_2, bb_2)$$

$$= a(a_1, b_1) + b(a_2, b_2)$$

$$= aT(\alpha) + bT(\beta)$$

Thus T is a linear transformation. Since for every $(a, b) \in \mathbb{R}^2$,

there exists $(a, b, 0) \in \mathbb{R}^3$ such that $T(a, b, 0) = (a, b)$,

hence the mapping T is onto $\mathbb{R}^2$, Therefore T is a linear transformation of $\mathbb{R}^3$ onto $\mathbb{R}^2$

Example 3.4.2 : Prove that rotation R_θ of the Euclidean plane by an angle θ about the origin is a linear operator on $\mathbb{R}^2$

Solution : Let (x, y) be the coordinates of a point P with respect to a rectangular Cartesian coordinate system with origin at O . Suppose the plane is rotated about the origin by an angle θ in counter clock-wise direction. Let the point $P(x, y)$ move to the point $Q(\xi, \eta)$ given by

$$\xi = x \cos \theta - y \sin \theta$$

$$\eta = x \sin \theta + y \cos \theta$$

The rotation R_θ may be considered as a mapping R_θ : given by

$$R_\theta(x, y) = (\xi, \eta)$$

$$= (x \cos \theta - y \sin \theta, x \sin \theta + y \cos \theta)$$

To show that R_θ is a linear operator on let

$(x_1, y_1), (x_2, y_2) \in \mathbb{R}^2$... Then for $a, b \in \dots$, we have

$$\begin{aligned} R_\theta (a(x_1, y_1) + b(x_2, y_2)) &= R_\theta(ax_1 + bx_2, ay_1 + by_2) \\ &= ((ax_1 + bx_2) \cos \theta - (ay_1 + by_2) \sin \theta, (ax_1 + bx_2) \sin \theta + (ay_1 + by_2) \cos \theta) \\ &= (a(x_1 \cos \theta - y_1 \sin \theta) + b(x_2 \cos \theta - y_2 \sin \theta), a(x_1 \sin \theta + y_2 \cos \theta) \\ &\quad + b(x_2 \sin \theta + y_2 \cos \theta)) \\ &= a(x_1 \cos \theta - y_1 \sin \theta, x_1 \sin \theta + y_1 \cos \theta) + b(x_2 \cos \theta - y_2 \sin \theta, x_2 \sin \theta + y_2 \cos \theta) \\ &= a R_\theta(x_1, y_1) + b R_\theta(x_2, y_2) \end{aligned}$$

Thus we see that the rotation is a linear operation on $\mathbb{R}^2$

Note- Here you will notice that in matrix form, we have

$$R_\theta \left(\begin{bmatrix} x \\ y \end{bmatrix} \right) = \begin{bmatrix} \xi \\ \eta \end{bmatrix} = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

$$\text{or } R_\theta(X) = AX$$

$$\text{Where } A = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \text{ and } X = \begin{bmatrix} x \\ y \end{bmatrix}$$

we have already discussed a similar transformation in the section 3.1 and we know that such type of transformations are linear transformations. The matrix A given above is called the rotation matrix.

Example 3.4.3 You can verify for yourself that the conjugation map

$T: C \rightarrow C$ given by $T(Z) = \bar{Z}, \forall Z \in C$ is a linear operator

Example 3.4.4 Let F be a field and let V be the space of polynomial functions F defined from F into F .

Let $f \in V$ Such that $f(x) = a_0 + a_1x + \dots + a_nx^n$

where $a_0, a_1, \dots, a_n \in F$

Let $D: V \rightarrow V$ be a mapping given by

$$(Df)(x) = a_1 + 2a_2x + \dots + na_nx^{n-1}$$

Then you can verify that D is a linear transformation from V into V . This transformation is called the differentiation transformation. All you have to do is to show that

$$D(f + g) = Df + Dg$$

$$\text{and } D(cf) = cD(f)$$

for all $f, g \in V$ and all $C \in F$

Example 3.4.5 Now we give an important example from matrices. let F be a field. For any positive integers m and n , let $F^{m \times n}$ denote the set of all $m \times n$ matrices over F and $A \in F^{m \times n}$

Define a mapping $T: F^{n \times 1} \rightarrow F^{m \times 1}$ by

$$T(X) = AX, \forall X \in F^{n \times 1}$$

Now from the theory of matrices we know that

$$A(X + Y) = AX + AY \text{ and } A(cX) = c(AX)$$

For any $X, Y \in F^{n \times 1}$ and any $C \in F$

$$\text{Therefore } T(X + Y) = T(X) + T(Y)$$

$$\text{And } T(CX) = CT(X)$$

i.e. T is a linear transformation from $F^{n \times 1}$ into $F^{m \times 1}$

Example 3.4.6 Let us generalize the mapping given in example 3.4.1 Consider the mapping $T: R^m \rightarrow R^n (m \geq n)$ defined by

$$T(x_1, x_2, \dots, x_m) = (x_1, \dots, x_n) \text{ defined by}$$

Then you can verify that T is a linear transformation called the natural projection of R^m into R^n

3.5 Properties of linear Transformations

Theorem 3.5.1 let $T: V \rightarrow W$ be a linear transformation.

- I. then $T(O_V) = O_W$, Where O_V and O_W are the zero vectors in V and W respectively.
- II. $T(-\alpha) = -T(\alpha), \forall \alpha \in V$
- III. $T(\alpha - \beta) = T(\alpha) - T(\beta) \dots \forall \alpha, \beta \in V$

$$\text{IV. } T[\sum_{i=0}^n a_i \alpha_i] = \sum_{i=0}^n a_i T(\alpha_i)$$

Where $\alpha_1, \alpha_2, \dots, \alpha_n \in V$ and $a_1, a_2, \dots, a_n \in F$

Proof: -(1) Let $\alpha \in V$. Then $T(\alpha) \in W$. we have

$$T(\alpha) + 0_W = T(\alpha)$$

$$= T(\alpha + 0_V)$$

$$= T(\alpha) + T(0_V)$$

$$\therefore T(\alpha) + 0_W = T(\alpha) + T(0_V)$$

$0_W = T(0_V)$, by cancellation property for addition

(2) We have

$$T(0_V) = 0_W$$

$$= T[\alpha + (-\alpha)] = 0_W$$

$$= T(\alpha) + T(-\alpha) = 0_W$$

$$\text{Consequently } T(-\alpha) = -T(\alpha)$$

(3) We have

$$T(\alpha - \beta) = T(\alpha) + T(-\beta) \dots \because \alpha - \beta = \alpha + (-\beta)$$

$$= T(\alpha) + [-T(\beta)] \dots \text{By (2)}$$

$$= T(\alpha) - T(\beta)$$

we prove the result by induction on n

For $n = 1$, the result is obviously true. Suppose that the result holds for $(n = k, i.e.)$

now

$$T\left[\sum_{i=1}^k a_i \alpha_i\right] = \sum_{i=1}^k a_i T(\alpha_i) \dots (3)$$

$$\text{now } T\left[\sum_{i=1}^{k+1} a_i \alpha_i\right] = T\left[\sum_{i=1}^k a_i \alpha_i + a_{k+1} \alpha_{k+1}\right]$$

$$= T\left[\sum_{i=1}^k a_i \alpha_i\right] + T(a_{k+1} \alpha_{k+1})$$

$$\sum_{i=1}^k a_i T(\alpha_i) + a_{k+1} T(\alpha_{k+1}) \text{ using (3)}$$

$$= \sum_{i=1}^{k+1} a_i T(\alpha_i)$$

Hence the result holds for α also. Therefore by mathematical induction, the result is true for all α .

Theorem 3.5.2 Let V be a finite dimensional vector space over the field F . Let $B = \{\alpha_1, \alpha_2, \dots, \alpha_n\}$ be an ordered basis for V . Let W be a vector space over the same field F . Then for any sequence of n vectors $\beta_1, \beta_2, \dots, \beta_n$ in W there exists a unique linear transformation $T: V \rightarrow W$ such that $T(\alpha_i) = \beta_i$, $i=1, 2, \dots, n$.

Proof- Let $\alpha \in V$ then there exists a unique set of scalars $a_1, a_2, \dots, a_n$ such that

$$\alpha = a_1 \alpha_1 + \dots + a_n \alpha_n$$

Define $T: V \rightarrow W$ such that

$$T(\alpha) = a_1 \beta_1 + \dots + a_n \beta_n$$

This mapping is well defined. For if $\alpha, \beta \in V$ such that

$$\alpha = a_1 \alpha_1 + \dots + a_n \alpha_n, \quad \beta = b_1 \alpha_1 + \dots + b_n \alpha_n,$$

$$\text{then } \alpha = \beta \Rightarrow a_1 \alpha_1 + \dots + a_n \alpha_n = b_1 \alpha_1 + \dots + b_n \alpha_n$$

$$\Rightarrow (a_1 - b_1) \alpha_1 + \dots + (a_n - b_n) \alpha_n = 0_V$$

$$\Rightarrow a_1 - b_1 = 0, \dots, a_n - b_n = 0$$

$$\Rightarrow a_1 = b_1, \dots, a_n = b_n$$

$\therefore \alpha_1, \alpha_2, \dots, \alpha_n$ are linearly independent

To show that T is unique, suppose T' is a linear transformation from V to W such that $T'(\alpha_i) = \beta_i$

$$\text{then } T'(\alpha) = T'(a_1 \alpha_1 + \dots + a_n \alpha_n)$$

$$= a_1 T'(\alpha_1) + \dots + a_n T'(\alpha_n)$$

$$= a_1 \beta_1 + \dots + a_n \beta_n$$

$$= T(\alpha)$$

So That $T' = T$

If V and W are any two given vector spaces defined over the same field F , We can define a number of mappings from V into W . Theorem 3.5.2 is

helpful in constructing linear mappings from V into W . Here we notice that any such linear mapping is completely determined by its action on the given ordered basis of V .

To illustrate it further, Suppose we are interested in finding a linear transformation

$$T: \mathbb{R}^3 \rightarrow \mathbb{R}^3 \text{ such that } T(1,0,0) = (1,0,-1) \\ T(0,1,0) = (1,2,2) \text{ and } T(0,0,1) = (0,0,0).$$

Here we have taken $\{(1,0,0), (0,1,0), (0,0,1)\}$

the ordered basis for $\mathbb{R}^3$ and $\beta_1 = (1,0,0), \beta_2 = (0,1,0)$ and $\beta_3 = (0,0,1)$.

to find an explicit expression for T ,

let $\alpha \in \mathbb{R}^3$ such that $\alpha = (a, b, c)$. then

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

$$T(\alpha) = T[a(1,0,0) + bT(0,1,0) + c(0,0,1)]$$

$$= aT(1,0,0) + bT(0,1,0) + cT(0,0,1)$$

$$= a(1,0,-1) + b(1,2,2) + c(0,0,0)$$

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

We also have the following observation:

Theorem 3.5.3 -Let S and T be two linear transformations from V into W . Let $\{\alpha_1, \alpha_2, \dots, \alpha_n\}$ be a

basis of V . if $S(\alpha_i) = T(\alpha_i)$ for $i = 1, 2, \dots, n$, then $S(\alpha) = T(\alpha) \forall \alpha \in V$

Proof- Let $\alpha \in V$. Then α can be uniquely expressed as a linear combination of the vectors $\alpha_1, \alpha_2, \dots, \alpha_n$. Hence there exist scalars $a_1, a_2, \dots, a_n$ such that

$$\alpha = a_1\alpha_1 + \dots + a_n\alpha_n$$

$$\text{now } S(\alpha) = S(a_1\alpha_1 + \dots + a_n\alpha_n)$$

$$= a_1S(\alpha_1) + \dots + a_nS(\alpha_n), \text{ by theorem 3.5.1(4)}$$

$$= a_1T(\alpha_1) + \dots + a_nT(\alpha_n), \{\because S(\alpha_i) = T(\alpha_i)\}$$

$$T(a_1\alpha_1 + \dots + a_n\alpha_n) = T(\alpha)$$

Hence if S and T agree on the basis of V , these linear transformation will be the same.

Now we introduce some important subspaces related to linear transformations.

3.6 NULL SPACE AND RANGE OF A LINEAR TRANSFORMATION

Let $T: V \rightarrow W$ be a linear transformation. Suppose we are interested in those elements of V which are mapped onto the zero O_w of W . So we have the subset $N(T)$ of V such that

$$N(T) = \{\alpha \in V : T(\alpha) = O_w\}$$

You can verify that $N(T)$ is indeed a subspace of V .

We have $T(O_v) = O_w$, So that $N(T)$ is non-empty. Let $\alpha, \beta \in N(T)$ and $a, b \in F$, then $T(\alpha) = O_w$ and $T(\beta) = O_w$

$$\text{Now } T(a\alpha + b\beta) = aT(\alpha) + bT(\beta) = aO_w + bO_w = O_w$$

$$a\alpha + b\beta \in N(T)$$

Therefore $N(T)$ is a subspace of V . So we have the following definition.

Definition - Let V and W be any two vector spaces over a field F and let $T: V \rightarrow W$ be a linear transformation. The subspace $\{\alpha \in V : T(\alpha) = O_w\}$, called the null space of T (or kernel of T). It is denoted by $\ker(T)$. The dimension of the null space is called the nullity of T .

Also there is another important Subspace associated with the linear transformation $T: V \rightarrow W$. We call it the range of T . If we collect the images of the elements of V under T , we get a subset $R(T)$ as

$$R(T) = \{T(\alpha) \in W : \alpha \in V\}$$

Since $O_v \in V$, hence $T(O_v) = O_w$
 $\in R(T)$. Thus $R(T)$ is non-empty. let β_1, β_2
 $\in R(T)$ and $a, b \in F$,

By definition, there exist

$$\alpha_1, \alpha_2 \in V \text{ such that } \beta_1 = T(\alpha_1) \text{ and } \beta_2 = T(\alpha_2)$$

$$\text{now } a\beta_1 + b\beta_2 = aT(\alpha_1) + bT(\alpha_2)$$

$$T(a\alpha_2 + b\alpha_1) \in R(T), \text{ Since } a\alpha = a\alpha_1 + b\alpha_2 \in V$$

Hence $R(T)$ is a subspace of W .

So we have the following definition .

Definition- Let $T: V \rightarrow W$ be a linear transformation from V into W , then the subspace $\{T(\alpha) \in W: \alpha \in V\}$ is called the range of T . The dimension of the range of T is called the rank of T . Now we prove an important result called the rank-nullity theorem or the Sylvester law.

3.7 Rank- Nullity theorem (Sylvester law)

Theorem 3.7 Let T be a linear transformation of a finite dimensional vector space V into a vector space W , both defined over a field F . Then

$$\text{rank}(T) + \text{nullity}(T) = \dim V$$

Proof- Suppose $\text{nullity}(T) = K$ and $(\alpha_1, \alpha_2, \dots, \alpha_K)$ a basis of $\ker(T)$. We can extend this linearly independent set to form a basis of V .

$$\text{let } \{\alpha_1, \alpha_2, \dots, \alpha_K, \alpha_{K+1}, \dots, \alpha_n\}$$

$$\text{be a basis of } V. \text{ Then } \dim V = n$$

The vectors $T(\alpha_1), \dots, T(\alpha_n)$ belong to the range $R(T)$ of T . Hence if $\beta \in R(T)$, then there exists $\alpha \in V$ such that $\beta = T(\alpha)$

$$\text{Now } \alpha = a_1\alpha_1 + \dots + a_n\alpha_n \text{ for some } a_1, \dots, a_n \in F.$$

$$\begin{aligned} \therefore \beta &= T(a_1\alpha_1 + \dots + a_K\alpha_K + a_{K+1}\alpha_{K+1} + \dots + a_n\alpha_n) \\ &= a_1T(\alpha_1) + \dots + a_KT(\alpha_K) + a_{K+1}T(\alpha_{K+1}) + \dots + a_nT(\alpha_n) \end{aligned}$$

$$\text{since } \alpha_1, \dots, \alpha_K \in \ker(T), \text{ hence } T(\alpha_1) = 0, \dots, T(\alpha_K) = 0$$

$$\text{thus } \beta = a_{K+1}T(\alpha_{K+1}) + \dots + a_nT(\alpha_n)$$

$$\text{Hence the vectors } T(\alpha_{K+1}), \dots, T(\alpha_n)$$

$$\text{span the range of } T. \text{ We show}$$

$$\text{that } \{T(\alpha_{K+1}), \dots, T(\alpha_n)\} \text{ is indeed a basis of } R(T).$$

To show that these vectors are linearly independent,

$$\text{Let } c_{K+1}, \dots, c_n \in F \text{ such that}$$

$$C_{k+1}T(\alpha_{k+1}) + \dots + C_n T(\alpha_n) = O_w$$

$$\Rightarrow T(C_{k+1}\alpha_{k+1} + \dots + C_n \alpha_n) = O_w$$

$$C_{k+1}\alpha_{k+1} + \dots + C_n \alpha_n \in \ker(T).$$

Therefore there exist scalars $C_1 \dots \dots C_k \in F$ Such that

$$C_{k+1}\alpha_{k+1} + \dots + C_n \alpha_n = C_1 \alpha_1 + \dots \dots C_k \alpha_k$$

$$\Rightarrow C_1 \alpha_1 + \dots \dots + C_k \alpha_k - C_{k+1}\alpha_{k+1} - \dots - C_n \alpha_n = O_v$$

Since the vectors $\alpha_1, \dots \dots, \alpha_1$

are linearly independent, We

$$\text{must have } C_1 = 0, \dots \dots C_k = 0, C_{k+1} = 0, \dots \dots C_n = 0$$

$$\text{i.e. the vectors } T(\alpha_{k+1}), \dots \dots T(\alpha_n)$$

are linearly independent and therefore form a basis of $R(T)$.

$$\therefore \text{rank}(T) = \dim R(T) = n - k$$

$$\text{Hence rank}(T) + \text{nullity}(T) = n - k + k = n = \dim V$$

3.8 Illustrative Examples

Example 3.8.1 Find the range, rank, null space and nullity of the linear transformation $T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ given by $T(x, y) = (x + y, x - y, y)$

Solution We know that the set $\{(1,0)(0,1)\}$

is a basis of $\mathbb{R}^2$ now

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

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

obviously the vectors $T(1,0)$ $T(0,1)$ span the range of T therefore, the range of T is subspace of $\mathbb{R}^3$ spanned by the vectors $(1,1,0)$ and $(1,-1,1)$. you can verify that the vectors $(1,1,0)$ and $(1,-1,1)$. You can verify that the vector $(1,1,0)$ and $(1,-1,1)$ are linearly independent. Hence the set $\{(1,1,0), (1,-1,1)\}$ forms a basis for the range of T .

$$\therefore \text{rank}(T) = 2$$

Let $N(T)$ be the null space of T . Suppose

$(x, y) \in N(T)$, then we have

$$T(x, y) = (0, 0, 0)$$

$$\Rightarrow (x + y, x - y, y) = (0, 0, 0)$$

$$\Rightarrow x + y = 0, x - y = 0, y = 0$$

$$\Rightarrow x = 0, y = 0$$

Hence $N(T) = \{(0, 0)\}$ i.e the zerospace of $\mathbb{R}^2$

$$\therefore \text{Nullity}(T) = 0$$

Example 3.8.2 For the linear transformation $T: \mathbb{R}^3 \rightarrow \mathbb{R}^2$

$$\text{given by } T(x, y, z) = (x + y, y + z)$$

Show that (i) range $(T) = \mathbb{R}^2$ (ii) rank $(T) = 2$ (iii) nullity $(T) = 1$

Solution: -The set $\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$

is a basis of $\mathbb{R}^3$.

$$\text{now } T(1, 0, 0) = (1 + 0, 0 + 0) = (1, 0)$$

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

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

Hence the range of T is spanned

by the vectors $(1, 0), (1, 1), (0, 1)$ these vectors are linearly

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

Hence $\{(1, 0), (0, 1)\}$ is the linearly

independent set generating range (T)

i.e. a basis of range (T) .

(ii) and (iii): -Hence $\dim\{\text{range}(T)\} = 2$

Therefore $\dim \text{range}(T) = \dim \Rightarrow \text{rank}(T) = \dim \mathbb{R}^2$

Also $\text{range}(T) \subseteq \mathbb{R}^2$

$\therefore \text{range}(T) = \mathbb{R}^2$

(iii) Let $(x, y, z) \in N(T)$. Then $T(x, y, z) = (0, 0)$.

$$\text{Hence } (x + y, y + z) = (0, 0)$$

$$= x + y = 0, y + z = 0$$

So we have two homogeneous linear equations in three unknowns. Hence the system of equations will have non-zero solutions.

$$\text{Let } z = k, \text{ then } y = -z = -k$$

$$\text{and } x = -y = k$$

There're the general solution for this system of equations is

$$x = k, y = -k, z = k.$$

$$\therefore N(T) = \{(k, -k, k) : k \in \mathbb{R}\}$$

$$= \{k(1, -1, 1) : k \in \mathbb{R}\}$$

Hence the nullspace $N(T)$ is a subspace

of $\mathbb{R}^3$ spanned by the

vector $(1, -1, 1)$. Thus $\{(1, -1, 1)\}$ is

a basis of $N(T)$ and

$$\text{nullity}(T) = \dim N(T) = 1$$

You can also deduce this resultant by using rank-nullity theorem

from rank-nullity theorem, we have

$$\text{nullity}(T) = \dim \mathbb{R}^3 - \text{rank}(T) = 3 - 2 = 1$$

$$\therefore \dim \mathbb{R}^3 = 3 \text{ \& rank}(T) = 2$$

Example 3.8.3- Let V and W be any two vector spaces defined over a field F . Let $T: V \rightarrow W$ be a linear transformation. then T is one-to-one if and only if $\ker(T) = \{0_v\}$.

Solution;- First suppose that T is one-to-one.

Let $\alpha \in \ker(T)$. Then

$$T(\alpha) = 0_w$$

or $T(\alpha) = T(O_v)$, Since $T(O_v) = 0$

$= \alpha = O_v$, since T is one - to - one

$$\therefore \text{Ker}(T) = \{O_v\}$$

Conversely, Suppose $\text{Ker}(T) = \{O_v\}$. Let $\alpha, \beta \in V$.

$$\text{Then } T(\alpha) = T(\beta) \Rightarrow T(\alpha) - T(\beta) = O_w$$

$$\Rightarrow T(\alpha - \beta) = O_w$$

$$\Rightarrow \alpha - \beta \in \text{ker}(T)$$

$$\Rightarrow \alpha - \beta = O_v$$

$$\Rightarrow \alpha = \beta$$

Hence, T is one - to - one .

Example 3.8.4 In example 3.4.5 we have seen that if $F^{m \times n}$ denotes

the set of all $m \times n$ matrices over a field F and $A \in F^{m \times n}$,

then the mapping $T: F^{n \times 1} \rightarrow F^{m \times 1}$ given by

$T(X) = AX$ is a linear transformation. For example.

Consider $T: R^{2 \times 1} \rightarrow R^{3 \times 1}$ given by $T(x) = AX$ where $X = \begin{bmatrix} x \\ y \end{bmatrix}$ and A

$$= \begin{bmatrix} 2 & 3 \\ 1 & 2 \\ 0 & 1 \end{bmatrix}. \text{ Suppose we are interested}$$

in finding the nullity and

rank of this transformation.

If $x = \begin{bmatrix} x \\ y \end{bmatrix} \in \text{ker}(T)$, Then

$$T(X) = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 2 & 3 \\ 1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow 2x + 3y = 0, x + 2y = 0, y = 0$$

$$\Rightarrow x = 0, y = 0$$

$$\text{Hence Ker } (T) = \left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\}$$

So That T is one – to – one and nullity $(T) = 0$.

$$\begin{aligned} \text{By Theorem 3.7 rank } (T) &= \dim \mathbb{R}^{2 \times 1} - \text{nullity}(T) \\ &= 2 - 0 = 2 \end{aligned}$$

3.9 Isomorphism of vector spaces

Definition: Let V and W be two vector spaces defined over a field F . Then a linear transformation $T: V \rightarrow W$ is called an isomorphism of V onto W if

- i. T is one-one
- ii. T is onto

In other words, a one-one linear transformation of V onto W is an isomorphism of V onto W .

If such an isomorphism of V and W exists, we say that the space V is isomorphism to the space W .

Suppose $\mathbb{C}(\mathbb{R})$ be the space of all complex numbers over the field of reals $\mathbb{R}$. Then you can verify that the mapping $T: \mathbb{C} \rightarrow \mathbb{R}^2$ given by $T(a+ib) = (a, b)$ is an isomorphism.

Let $\alpha, \beta \in \mathbb{C}$ such that $\alpha = a_1 + ib_1, \beta = a_2 + ib_2$.

$$\text{Then } T(\alpha) = T(\beta) \Rightarrow (a_1, b_1) = (a_2, b_2)$$

$$\Rightarrow a_1 = a_2, \quad b_1 = b_2$$

$$= a_1 + ib_1 = a_2 + ib_2$$

$$\alpha = \beta$$

Hence T is one – one.

Also T is onto. if $(a, b) \in \mathbb{R}^2$, Then there exists $\alpha = a_1 + ib_1 \in \mathbb{C}$ such that $T(a + ib) = (a, b)$.

To check whether T is a linear transformation, let

$\alpha = a_1 + ib_1, \beta = a_2 + ib_2 \in \mathbb{C}$ and $a, b \in \mathbb{R}$. Then

$$\begin{aligned} T(a\alpha + b\beta) &= T[a(a_1 + ib_1) + b(a_2 + ib_2)] \\ &= T[(aa_1 + ba_2) + i(ab_1 + bb_2)] \\ &= (aa_1 + ba_2, ab_1 + bb_2) \\ &= (aa_1, ab_1) + (ba_2, bb_2) \\ &= a(a_1, b_1) + b(a_2, b_2) \\ &= aT(\alpha) + bT(\beta) \end{aligned}$$

Thus T is an isomorphism of $\mathbb{C}$ onto $\mathbb{R}^2$. Now we prove an important result which shows that every n -dimensional vector space over the field F is abstractly identical with $\dots F^n$.

Theorem 3.9.1 Every n -dimensional vector space V over the field F is isomorphic to the space $\dots F^n$.

Proof: — Let V be an n

— dimensional vector space over the field F and let $B = \{\alpha_1, \dots, \alpha_n\}$ be an ordered basis of V .

Then every vector $\alpha \in V$

can be uniquely expressed as.

$$\alpha = x_1\alpha_1 + \dots + x_n\alpha_n$$

Where $x_1, x_2, \dots, x_n \in F$.

Obviously the ordered n — tuple $(x_1, \dots, x_n) \in F^n$.

So we can define a mapping $T: V \rightarrow F^n$

$$\text{by } T(\alpha) = (x_1, x_2, \dots, x_n)$$

The mapping is well — defined as for given

$\alpha \in V$ there exists unique n — tuple $(x_1, x_2, \dots, x_n) \in F^n$ such that

$$\alpha = x_1\alpha_1 + \dots + x_n\alpha_n \text{ and } \beta = y_1\alpha_1 + \dots + y_n\alpha_n.$$

$$\text{then } T(\alpha) = T(\beta) \Rightarrow (x_1, \dots, x_n) = (y_1, \dots, y_n)$$

$$\Rightarrow x_1 = y_1, \dots, x_n = y_n$$

$$\Rightarrow x_1\alpha_1 + \dots + x_n\alpha_n = y_1\alpha_1 + \dots + y_n\alpha_n$$

$$\Rightarrow \alpha = \beta$$

Also T is onto. for if $(x_1, \dots, x_n) \in F^n$ then there exists $\alpha = x_1\alpha_1 + \dots + x_n\alpha_n \in V$ such that $T(\alpha) = (x_1, \dots, x_n)$

To show that T is a linear transformation, let $\alpha, \beta \in V$

such that $\alpha = x_1\alpha_1 + \dots + x_n\alpha_n$ and

$\beta = y_1\alpha_1 + \dots + y_n\alpha_n$. Let $a, b \in F$, then

$$\begin{aligned}
a\alpha + b\beta &= a(x_1\alpha_1 + \cdots x_n\alpha_n) + b(y_1\alpha_1 + \cdots + y_n\alpha_n) \\
&= a(ax_1 + by_1)\alpha_1 + \cdots + (ax_n + by_n)\alpha_n \\
\therefore T(a\alpha + b\beta) &= (ax_1 + by_1, \dots, ax_n + by_n) \\
&= a(x_1 \dots x_n) + b(y_1, \dots y_n) \\
&= aT(\alpha) + bT(\beta)
\end{aligned}$$

Therefore T is an isomorphism of V onto F^n , i.e. V is isomorphic to F^n .

It appears that an isomorphism must be a dimension preserving mapping. The following result confirms this fact.

Theorem 3.9.2. Two finite dimensional vector spaces over the same field are isomorphic if and only if they are of the same dimension.

Proof: Let V and W be two vector spaces over the same field F such that $\dim V = \dim W = n$ (say). Suppose

$\{\alpha_1, \dots, \alpha_n\}$ and $\{\beta_1, \dots, \beta_n\}$ are the bases of V and W respectively.

If $\alpha \in V$, then α can be uniquely expressed as

$$\alpha = x_1\alpha_1 + \cdots x_n\alpha_n$$

where $x_1, \dots, x_n \in F$.

Let us define a mapping $T: V \rightarrow W$ by

$$T(\alpha) = x_1\beta_1 + \cdots + x_n\beta_n$$

Obviously, The mapping is well-defined. We leave it to you as an exercise to verify that T is indeed an isomorphism of V onto W . Thus V is isomorphic to W . Conversely, Suppose that V is isomorphic to W and T is an isomorphism of V onto W .

Let $\dim V = n$ and $B = \{\alpha_1, \alpha_2, \dots, \alpha_n\}$

be a basis of V . We claim that $\{T(\alpha_1), \dots, T(\alpha_n)\}$ is a basis of W .

To show that $\{T(\alpha_1), \dots, T(\alpha_n)\}$ is linearly independent

let $x_1, \dots, x_n \in F$ such that

$$x_1T(\alpha_1) + \cdots + x_nT(\alpha_n) = 0_W$$

$$\Rightarrow T(x_1\alpha_1 + \cdots + x_n\alpha_n) = 0_w$$

$$x_1\alpha_1 + \cdots + x_n\alpha_n \in \ker(T)$$

since T is an isomorphism, it is one. Therefore

$$x_1\alpha_1 + \cdots + x_n\alpha_n = 0_v \quad \{\because \ker(T) = \{0_v\}\}$$

The vectors $\alpha_1, \dots, \alpha_n$ are linearly independent, hence we have

$$x_1 = 0, \dots, x_n = 0$$

It remains to prove that the vectors $T(\alpha_1), \dots, T(\alpha_n)$

span the space W .

Let $\beta \in W$. Since T is onto W , hence there exists $\alpha \in V$ such that

$$\beta = T(\alpha)$$

Now α can be uniquely expressed as

$$\alpha = c_1\alpha_1 + \cdots + c_n\alpha_n$$

for some $c_1, \dots, c_n \in F$

$$\therefore \beta = T(\alpha)$$

$$= T(c_1\alpha_1 + \cdots + c_n\alpha_n)$$

$$= c_1T(\alpha_1) + \cdots + c_nT(\alpha_n)$$

thus $\{T(\alpha_1), \dots, T(\alpha_n)\}$ is a basis of W and

$$\dim W = n = \dim V$$

Now we Characterize the isomorphism by the following important result.

Theorems 3.9.3 Let V and W be the finite-dimensional vector spaces over F . Then the following are equivalent:

- I. T is one-one
- II. T is onto
- III. T is an isomorphism

Proof:- Let $\dim V = \dim W = n$.

(I) $\Rightarrow$ (II) Suppose T is one-one. Then $\ker(T) = \{0\}$

and therefore $\text{nullity}(T) = \dim\{\ker(T)\} = 0$.

by rank-nullity theorem we have

$$\dim R(T) = \dim V - \text{nullity}(T) = n - 0 = n$$

$$\Rightarrow \dim R(T) = \dim W$$

Since $R(T)$ is a vector subspace of W , hence we must have $R(T) = W$. This (II) follows

(II) $\Rightarrow$ (III) Now suppose that T is onto. Then

$$\dim R(T) = \dim W = n$$

$$\Rightarrow \text{nullity}(T) = \dim V - \dim R(T) = n - n = 0$$

consequently, $\ker(T) = \{0_V\}$, i.e. T is one-one.

Hence T is a one-one linear transformation from V onto W , i.e. T is an isomorphism.

(III) $\Rightarrow$ (I) Suppose that T is an isomorphism. Then obviously T is one-one, i.e. (I) holds.

This completes the proof (I) $\Rightarrow$ (II) $\Rightarrow$ (III) $\Rightarrow$ (I)

Remark:- Theorem 3.9.3 is false in case of infinite dimensional vector spaces. Also the result fails when $\dim V \neq \dim W$.

3.10 Fundamental Theorem of vector space homomorphism

now we shall prove an important result which relates some quotient space to the range of a linear transformation by means of homomorphism.

Theorem 3.10.1: Let V and W be vector spaces over a field F and $T: V \rightarrow W$ be a linear transformation with kernel K . Then $V/K \cong R(T)$.

Proof:- The quotient space V/K consists of cosets of K in V .

since the kernel K is a subspace

of V , hence the quotient space V/K is well
— defined vector over the field F .

Let us define a map $\theta: V/K \rightarrow R(T)$ by

$$\theta(\alpha + K) = T(\alpha).$$

You can verify that θ is a well-defined mapping.

Let $\alpha, \beta \in V$ such that $\alpha + K = \beta + K$. Then $\alpha - \beta \in k$.

By the definition of kernels of T , We have

$$\begin{aligned} T(\alpha - \beta) &= 0_w \\ \Rightarrow T(\alpha) - T(\beta) &= 0_w \\ \Rightarrow T(\alpha) &= T(\beta) \\ \Rightarrow \theta(\alpha + K) &= \theta(\beta + K) \end{aligned}$$

Now we show that θ is a linear transformation.

Let $\alpha + k, \beta + k \in V/K$ and $a, b \in F$. Then

$$\begin{aligned} \theta[a(\alpha + k) + b(\beta + k)] &= \theta[a(\alpha + k) + b(\beta + k)] \\ &= \theta[(a\alpha + b\beta) + k] \\ &= T(a\alpha + b\beta) \\ &= aT(\alpha) + bT(\beta) \\ &= a\theta(\alpha + k) + b\theta(\beta + k) \end{aligned}$$

Thus θ is a linear transformation .

$$\text{Now } \ker(\theta) = \{\alpha + k \in V/K : \theta(\alpha + k) = 0_w\}$$

$$\text{Since } \theta(\alpha + k) = 0_w \Rightarrow T(\alpha) = 0_w$$

$$\Rightarrow \alpha \in k$$

$$\Rightarrow \alpha + k = k$$

hence $\ker(\theta) = \{K\}$, i. e. θ is one - one.

To show that θ is onto, let $T(\alpha) \in R(T)$.

Then $\alpha \in V$ and therefore $\alpha + k \in V/K$

such that $T(\alpha) = \theta(\alpha + k)$.

thus θ is an isomorphism of V/K onto $R(T)$ and have

$$V/K \simeq R(T).$$

To illustrate this theorem, Consider $W = \{(x, 0, 0) : x \in \mathbb{R}\}$

we know that W is a subspace of $\mathbb{R}^3$. Now define a linear

transformation $T: \mathbb{R}^3 \rightarrow \mathbb{R}^2$ by $T(x, y, z) = (y, z)$.

obviously, $\ker(T) = \{\alpha \in \mathbb{R}^3 : T(\alpha) = (0, 0)\}$

$$= \{(x, y, z) \in \mathbb{R}^3 : T(x, y, z) = (0, 0)\}$$

$$\text{Now } T(x, y, z) = (0, 0)$$

$$\Rightarrow (y, z) = (0, 0)$$

$$\Rightarrow y = 0, z = 0$$

$$\therefore \ker(T) = \{(x, 0, 0) : x \in \mathbb{R}\} = W$$

The map T is onto. For if $(x, y) \in \mathbb{R}^2$,

Then there exist $(0, x, y) \in \mathbb{R}^3$, such that

$$T(0, x, y) = (x, y)$$

Therefore T is a homomorphism from $\mathbb{R}^3$ onto $\mathbb{R}^2$ with kernel W .

Hence by the fundamental theorem of homomorphism $\mathbb{R}^3/W \simeq \mathbb{R}^2$

The subspace W of $\mathbb{R}^3$ P - 23

More generally, We have $\mathbb{R}^n/\mathbb{R}^m \simeq \mathbb{R}^{n-m}$ for any $n \geq m$.

You can also verify that $\mathbb{C}^n/\mathbb{C}^m \simeq \mathbb{C}^{n-m}$ for any $n \geq m$.

Now we have an important corollary to theorem 3.10.1

corollary: Let V be vector space over the field F and A, B be subspaces of V .

$$\text{Then } (A+B)/B \simeq A/(A \cap B).$$

we have $A + B = \{a + b : a \in A, b \in B\}$

Obviously $a = a + 0 \in A + B$. Hence $a + b \in \frac{A + B}{B}$.

So we can define a mapping $T: A \rightarrow \frac{A + B}{B}$ by

$T(a) = a + B$. This mapping is well – defined.

Now you can verify that this is a linear transformation.

Let $x, y \in F$ and $a_1, a_2 \in A$.

Then $T(xa_1 + ya_2) = (xa_1 + ya_2) + B$

$$= x(a_1 + B) + y(a_2 + B)$$

$$= xT(a_1) + yT(a_2)$$

Since $a + b \in A + B$ for $a \in A, b \in B$, hence

$$a + b + B \in A + B/B$$

$$\text{now } a + b + B = (a + B) + (b + B)$$

$$= (a + B) + B$$

$$= a + b$$

$$= T(a)$$

Hence for every $a + b + B \in A + B/B$

we have $a \in A$ such that $T(a) = a + b + B$.

Thus T is onto.

Let us find the Kernel of T

So that we can apply Theorem 3.10.1.

$$\text{Now } \ker(T) = \{a \in A : T(a) = B\}$$

Note that the zero element of $A + B/B$ is B .

$$\therefore \ker(T) = \{a \in A : a + B = B\}$$

$$\text{But } a + B = B \Rightarrow a \in B.$$

Thus $a \in A$ and $a \in B$, i.e. $a \in A \cap B$

$$\therefore \ker(T) \subseteq A \cap B$$

Also if $a \in A \cap B$, then $a \in A$ and $a \in B$.

$$\therefore a + B = B$$

$$\Rightarrow T(a) = B$$

$$\text{Hence } A \cap B \subseteq \ker(T)$$

$$\text{This gives } \ker(T) = A \cap B$$

Thus $T: A \rightarrow \frac{A+B}{B}$ given by $T(a) = a + B$ is a linear transformation of

A onto $\frac{A+B}{B}$ with kernel $A \cap B$.

$$\text{Since } T \text{ is onto, hence } R(T) = \frac{A+B}{B}.$$

Therefore by theorem 3.10.1

$$\text{we have } A/A \cap B \cong A+B/B$$

3.11 Invertible and Non – Singular Transformations

You know from the basic set theory that if f is a one-one and onto mapping, then a mapping $f^{-1}: B \rightarrow A$ exists such that

Let V and W be vector spaces over a field F .

Let $T: V \rightarrow W$ be a one – one and onto linear

transformation. Then T^{-1} exists,.

Now the question is whether T^{-1} is also a linear transformation. Let us investigate.

Let $\beta_1, \beta_2 \in W$ and $\alpha, \beta \in F$. Suppose

$$\alpha_1, \alpha_2 \in V \text{ such that } T^{-1}(\beta_1) = \alpha_1 \text{ and}$$

$$T^{-1}\left(\frac{\alpha}{\beta_2}\right) = \frac{\beta}{\alpha} \text{ Then } T(\alpha_1) = \beta_2 \text{ and } T(\alpha_2) = \beta_2$$

$$\text{Now } a\beta_1 + b\beta_2 = aT(\alpha_1) + bT(\alpha_2)$$

$$= T(a\alpha_1 + b\alpha_2)$$

$$\Rightarrow T^{-1}(a\beta_1 + b\beta_2) = a\alpha_1 + b\alpha_2$$

$$aT^{-1}(\beta_1) + bT^{-1}(\beta_2)$$

This shows that $T^{-1}: W \rightarrow V$ is a linear transformation.

Now $TT^{-1}: W \rightarrow W$ is the composition of the transformations

$$T^{-1}: W \rightarrow V \text{ and } T: V \rightarrow W.$$

Let $\beta \in W$. Since T is onto, there exists $\alpha \in V$

such that $\beta = T(\alpha)$. Then by definition of T^{-1} , We have $T^{-1}(\beta) = \alpha$.

$$\text{Now } (TT^{-1})(\beta) = T[T^{-1}(\beta)]$$

$$= T(\alpha)$$

$$= \beta$$

$$= I(\beta)$$

$$\therefore TT^{-1} = I(\text{the identity transformation})$$

Also $T^{-1}T: V \rightarrow V$ is the composition of

the transformations $T: V \rightarrow W$ and $T^{-1}: W \rightarrow V$.

You can verify for yourself that $T^{-1}T$ is

the identity transformation

from V onto V , i.e.

$$T^{-1}T = I$$

$$\therefore TT^{-1} = I = T^{-1}T$$

Hence we have the following definition:

DEFINITION:—Let V and W be vector spaces over the field F .

*A linear transformation $T: V \rightarrow W$ is called invertible if
there exists a map on W and $T^{-1}: W \rightarrow V$
such that TT^{-1} is the identity map on W
and $T^{-1}T$ is the identity map on V*

Also T is invertible if and only if T is one-one and onto.

Furthermore, $T^{-1}: W \rightarrow V$ is also a linear transformation.

Now you will observe that one-one linear transformations have interesting properties. In example 3.8.3, we see that a linear transformation $T: V \rightarrow W$ is one – one if and only if $\ker(T) = \{O_v\}$

In other words, if $T(\alpha) = O_w$, then $\alpha = O_v$.

We call such linear transformation non – singular.

So we have the following definition:

DEFINITION: Let V and W be vector spaces over the field F . A linear transformation $T: V \rightarrow W$ is called non-singular if

$$T(\alpha) = O_w \Rightarrow \alpha = O_v$$

i.e. if the kernel of T is $\{O_v\}$.

Obviously, T is one – one if and if T is non – singular.

*Now we prove that a non – singular linear
transformation preserves linear independence.*

Theorem 3.11.1 Let V and W be vector spaces over the field F . Then a linear transformation $T: V \rightarrow W$ is

*non – singular if and only if T carries
each linearly independent subset of V
onto a linearly independent subset of W .*

Proof:- First suppose that T is non-singular. Let $S = \{\alpha_1 \alpha_2 \dots \alpha_n\}$
be a linearly independent subset of V . we shall show

that the vectors

$T(\alpha_1), \dots, T(\alpha_n)$ are linearly independent.

Let $a_1 \dots a_n \in F$ such that

$$a_1 T(\alpha_1) + \dots + a_n T(\alpha_n) = 0_w$$

$$\text{Then } T(a_1 \alpha_1 + \dots + a_n \alpha_n) = 0_w$$

Since T is non-singular, hence

$$a_1 \alpha_1 + \dots + a_n \alpha_n = 0_v$$

$$\Rightarrow a_1 = 0, \dots, a_n = 0 \quad \{\because \alpha_1, \dots, \alpha_n \text{ are linearly independent}\}$$

Thus the image of a linearly independent subset of V under T is linearly independent.

Conversely, suppose that T carries every linearly independent subset of V onto a linearly independent subset of W .

consider $S = \{\alpha\} \subseteq V$ where $\alpha \neq 0_v$.

Obviously S is linearly independent. By our assumption, $\{T(\alpha)\}$

must be a linearly independent subset of W .

$$\text{Therefore } T(\alpha) \neq 0_w$$

$$\text{thus } 0_v \neq \alpha \in V \Rightarrow T(\alpha) \neq 0_w$$

$$\text{i.e. } T(\alpha) = 0_w \Rightarrow \alpha = 0_v$$

$$\Rightarrow \ker(T) = \{0_v\}$$

Hence T is non-singular.

Now we prove an important result for finite-dimensional vector spaces.

Theorem 3.11.2 Let V and W be finite-dimensional vector spaces over the field F such that $\dim V = \dim W$. If T is a linear transformation from V into W , the following are equivalent:

- I. T is invertible

- II. T is non-singular
 III. $R(T)=W$, i.e. T is onto

Proof: (I) $\Rightarrow$ (II) The implication is obvious. If T is invertible, then T is one-one. Thus T is non-singular.

(II) $\Rightarrow$ (III) Suppose that T is non-singular.

Let $\{\alpha_1, \dots, \alpha_n\}$ be a basis of V .

By Theorem 3.11.1, the set $\{T(\alpha_1), \dots, T(\alpha_n)\}$ is a linearly independent subset of W . Since $\dim V = \dim W$, hence $\dim W = n$. Therefore $\{T(\alpha_1), \dots, T(\alpha_n)\}$ forms a basis of W .

Let $\beta \in W$. Then for some $a_1, \dots, a_n \in F$, we have

$$\begin{aligned}\beta &= a_1 T(\alpha_1) + \dots + a_n T(\alpha_n) \\ &= T(a_1 \alpha_1 + \dots + a_n \alpha_n)\end{aligned}$$

Now $a_1 \alpha_1 + \dots + a_n \alpha_n \in V$. Hence

$$\beta = T(a_1 \alpha_1 + \dots + a_n \alpha_n) \in R(T)$$

Thus $W = R(T)$, i.e. T is onto.

(III) $\Rightarrow$ (I) Suppose that T is onto, i.e. $W = R(T)$.

Let $n = \dim V = \dim W$.

Since $W = R(T)$, Hence $\text{rank}(T) = n$.

By rank - nullity theorem, we have

$$\text{rank}(T) + \text{nullity}(T) = n$$

$$\Rightarrow n + \text{nullity}(T) = n$$

$$\Rightarrow \text{nullity}(T) = 0$$

$$\text{i.e. } \ker(T) = \{0_v\}$$

Thus T is non-singular and hence one - one

and onto linear transformation. Therefore T is invertible.

Let us illustrate this theorem with an example.

Suppose $T: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ be a linear operator on $\mathbb{R}^3$ given by T

$(x, y, z) \mapsto (3x, x - y, 2x + y + z)$ you can verify for yourself that T is one-one. Therefore T is non-singular. By theorem 3.11.2, T must be invertible.

To find the rule for T^{-1} , Let $T(x, y, z) = (p, q, r)$.

Then $T^{-1}(p, q, r) = (x, y, z)$.

Now $T(x, y, z) = (p, q, r) \Rightarrow (3x, x - y, 2x + y + z) = (p, q, r)$.

$$\Rightarrow 3x = p, x - y = q, 2x + y + z = r$$

$$\therefore x = \frac{p}{3}, y = \frac{p}{3} - q, z = r - p + q$$

$$\text{Hence } T^{-1}(p, q, r) = \left(\frac{p}{3}, \frac{p}{3} - q, r - p + q\right)$$

3.12 .SUMMARY

In this unit, We have studied the following facts. (1) A linear transformation from a vector space V over F to a vector space W over F is a function $T: V \rightarrow W$ such that

$$T(a\alpha + b\beta) = aT(\alpha) + bT(\beta)$$

for all $\alpha, \beta \in V$ and all $a, b \in F$

(2) If $T: V \rightarrow W$ is a linear transformation, then

- I. $T(O_v) = O_w$
- II. $T(-\alpha) = -T(\alpha) \forall \alpha \in V$
- III. $T(\alpha - \beta) = T(\alpha) - T(\beta) \forall \alpha, \beta \in V$
- IV. $T[\sum_{i=1}^n a_i \alpha_i] = \sum_{i=1}^n a_i T(\alpha_i)$
Where $a_1, \dots, a_n \in F$ and $\alpha_1, \dots, \alpha_n \in V$

(3) Let V be a finite dimensional vector space over the field F . Let $B = \{\alpha_1, \dots, \alpha_n\}$ be an ordered basis for V . Let W be a vector space over the same field F . Then for any sequence of n vectors $\beta_1, \beta_2, \dots, \beta_n$ in W , there exists a unique linear transformation

$$T: V \rightarrow W \text{ such that } T(\alpha_i) = \beta_i, i = 1, \dots, n$$

(4) Let $T: V \rightarrow W$ be a linear transformation. the null space or kernel for T is a subspace $\ker(T)$ of V given by

$$\ker(T) = \{\alpha \in V: T(\alpha) = O_w\}$$

(5) The range $R(T)$ of a linear transformation $T: V \rightarrow W$ is a subspace of W given by

$$R(T) = \{T(\alpha): \alpha \in V\}$$

(6) The dimension of $R(T)$ is called the rank of T and the dimension of $\ker(T)$ is called the nullity of T , i.e.

$$\text{rank}(T) = \dim R(T)$$

$$\text{nullity}(T) = \dim \ker(T)$$

(7) Let T be a linear transformation of a finite dimensional vector space V a vector space W , both defined over a field F . Then

$$\text{rank}(T) + \text{nullity}(T) = \dim V$$

(8) A linear transformation $T: V \rightarrow W$ is called an isomorphism if T is one-one and onto. Then V and W are called isomorphic spaces and we write $V \cong W$.

(9) Every n -dimensional vector space V over the field F is isomorphic to the space F^n

(10) Two finite dimensional vector spaces over the same field are isomorphic if and only if they are of the same dimension.

(11) A linear transformation $T:V \rightarrow W$ is one-one if and only if $\ker(T)=\{O_v\}$

(12) A linear transformation $T:V \rightarrow W$ is onto if and only if $R(T)=W$.

(13) Let V and W be the finite-dimensional vector spaces over the same field F . Let $T:V \rightarrow W$ be a linear transformation, then the following statements are equivalent:

- I. T is one-one
- II. T is onto
- III. T is an isomorphism.

(14) Let V and W be the vector spaces over a field F and $T:V \rightarrow W$ be a linear transformation with kernel K . Then $V/K=R(T)$.

(15) A linear transformation $T:V \rightarrow W$ is called invertible if there exists a map $T^{-1}:W \rightarrow V$ such that $T^{-1}T$ is the identity map on V and TT^{-1} is that identity map on W .

(16) If $T:V \rightarrow W$ is invertible linear transformation, then $T^{-1}:W \rightarrow V$ is also a linear transformation

(17) $T:V \rightarrow W$ is invertible if and only if T is non-singular if T is one-one and onto.

(18) A linear transformation $T:V \rightarrow W$ if $T(\alpha)=O_w \Rightarrow \alpha = O_v$

i.e. $\ker(T)=\{O_v\}$

(19) A non-singular linear transformation $T:V \rightarrow W$ carries each linearly independent subset of V onto a linearly independent subset of W .

(20) Let V and W be finite-dimensional vector spaces over the field F such that $\dim V = \dim W$. If T is a linear transformation from V onto W , the following are equivalent;

- I. T is invertible
- II. T is non-singular
- III. $R(T) = W$, i.e. T is onto

3.13 Terminal Questions

1. Show that the mapping

$T: \mathbb{R}^2 \rightarrow \mathbb{R}^3$ given by $T(x, y) = (x, y, 0)$ is a linear transformation.

2. Show that the function $\pi_k: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ given

by $\pi_k(x_1, \dots, x_{k-1}, x_k, \dots, x_n) = x_k$
is a linear transformation.

This linear transformation is called the projection

(on the K^{th} coordinate for $K = 1, \dots, n$)

3. Let $T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be defined by

$$T(x, y) = (x, -y) \forall (x, y) \in \mathbb{R}^2$$

show that T is a linear transformation.

[This is the reflection in the x -axis]

4. For a fixed vector $(a, b, c) \in \mathbb{R}^3$, show that the mapping

$$T(x, y, z) = ax + by + cz$$

is a linear transformation.

5. Show that the mapping $T: \mathbb{R}^3 \rightarrow \mathbb{R}^2$ given

$$\text{by } T(x, y, z) = (3x - 2y + z, x - 3y - 2z)$$

is a linear transformation from $\mathbb{R}^3$ into $\mathbb{R}^2$.

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

Consider the map $T: V \rightarrow \frac{V}{W}$ given by $T(\alpha) = \alpha + W$.

Then T is called the quotient map or the natural map.

Show that T is a linear transformation.

7. Let $M_n(F)$ be the vector space of all $n \times n$

matrices over the field F and let B be a fixed $n \times n$

matrix. Define $T: M_n(F) \rightarrow M_n(F)$ by

$$T(A) = AB - BA$$

Show that T is a linear operator on $M_n(F)$.

8. Show that the following mappings are not linear

1. $T: \mathbb{R}^3 \rightarrow \mathbb{R}^2$ defined by $T(x, y, z) = (1, 1, 0)$

2. $T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ defined by $T(x, y) = (1 + x, y)$

3. $T: \mathbb{R}^3 \rightarrow \mathbb{R}$ defined by $T(x, y) = xy$

9. Let $I: V \rightarrow V$ be the identity transformation.

$$\text{Show that } R(I) = V \text{ and } \ker(I) = \{0\}$$

10. Let $T: \mathbb{R}^2 \rightarrow \mathbb{R}$ be defined by

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

find $R(T)$ and $\ker(T)$.

$$\text{Ans: } R(T) = \mathbb{R}, \ker(T) = \{(x, y, z) \in \mathbb{R}^3 : 3x + y + 2z = 0\}$$

11. Let $T: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ be the linear transformation defined by

$$T(x, y, z) = (x + 2y - z, y + z, x + y - 2z) \text{ find}$$

I. A basis of $R(T)$

II. null space (i.e. $\ker(T)$)

III. rank (T) and nullity (T)

$$[\text{Ans: (I) } (1, 0, 1), (2, 1, 1), \text{ (II) } \ker(T) = \{(3k, -k, k) : k \in \mathbb{R}\} \text{ (III) rank } (T) = 2, \text{ nullity}(T) = 1]$$

12. Let $T: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ be defined by

$$T(x, y, z) = (x - y + 2z, 2x + y, -x - 2y + 2z)$$

Find rank (T) and nullity (T) .

$$[\text{Ans: rank}(T) = 2, \text{ nullity}(T) = 1]$$

13. Find $R(T)$ and $\ker(T)$ for the operator $T: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ given by

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

$$\left[\begin{array}{l} \text{Ans: } R(T) = \{(x, x, x) \in \mathbb{R}^3 : x \in \mathbb{R}\} \\ \ker(T) = \{(x, y, z) : x + y + z = 0\} \end{array} \right]$$

14. Find the nullity of the map $L: \mathbb{R}^3 \rightarrow \mathbb{R}$ given

$$\text{by } L(x, y, z) = x + y + z \quad [\text{Ans: nullity}(L) = 2]$$

15. let V be a vector space over a field F . Let S and T

be linear operators from V to V . prove that

- I. $\text{rank}(ST) \leq \min(\text{rank}(S), \text{rank}(T))$
- II. $\text{nullity}(ST) \geq \max(\text{nullity}(S), \text{nullity}(T))$

16. Let $T: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ be given by

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

Is T an isomorphism why, Define T^{-1} , if it exists.

[Ans: yes. T^{-1} exists. $T^{-1}(x, y, z) = (x + y, y, z)$]

17. Let V be the real vector space of all polynomials

Let $D: V \rightarrow V$ be defined by

$$D(a_0 + a_1x + \dots + a_nx^n) = a_1 + \dots + na_nx^{n-1}$$

show that D is onto but not 1-1

18. let $T: C \rightarrow C$ be a linear operator given $T(x, y) = (\alpha x + \beta y, \gamma x + \delta y)$
where $\alpha, \beta, \gamma, \delta$ are fixed elements of C Prove that T is
invertible if and only if $\alpha\delta - \beta\gamma \neq 0$

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

20. Describe explicitly the linear transformation

$T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ such that $T(1, 0) = (a, b)$ and

$T(0, 1) = (c, d)$ [Ans ; $T(x, y) = (ax + cy, bx + dy)$]

21. Describe explicitly a linear transformation

$T: \mathbb{R}^3 \rightarrow \mathbb{R}^4$ Which has its range the subspace spanned

by the vector $(1, 2, 0, -4), (2, 0, -1, -3)$

$$\text{Ans : } T(x, y, z) = (x + 2y, 2x, -y, -4x - 3y)$$

22. Let V be a finite dimensional space and let S

and T be linear operators on V . Show that

- I. $\text{rank}(ST) = \text{rank}(T) - \dim(R(T) \cap \ker(S))$
- II. $\text{nullity}(ST) = \text{nullity}(T) + \dim(R(T) \cap \ker(S))$

23. Using the fundamental theorem of

homomorphism, show that $C^{n-m} \simeq C^n / C^m$ for $n \geq m$

24. Let A and B be subspaces of a vector space V .

show that $A + B/B \simeq A$

25. Let A and B be subspaces of a vector space V .

Then for any subspace U of V containing W , SHOW THAT

$$\frac{\frac{V}{W}}{\frac{U}{W}} \cong \frac{V}{U}$$

26. Let T be a linear transformation from $\mathbb{R}^3$ into $\mathbb{R}^2$ and let

S be a linear transformation from $\mathbb{R}^2$ into $\mathbb{R}^3$ prove that the

transformation (ST) is not invertible

27. Let V and W be finite – dimensional vector spaces over

the field F such that $\dim V$

$= \dim W$. If T is a linear transformation from V into W , Prove

that the following are equivalent:

- I. T is non – singular
- II. If $\{\alpha_1, \dots, \alpha_n\}$ is a basis for V .
- III. There is some basis $\{\alpha_1, \dots, \alpha_n\}$ for V such that $\{T(\alpha_1), \dots, T(\alpha_n)\}$ is a basis for W .

3.14 FURTHER READINGS

1. Hoffman, K. and Kunze, R. (2015): linear Algebra, Pearson India Education Services Ltd.
2. Kumaresan, S. (2014): Linear Algebra: A geometric approach, New Delhi, Prentice Hall of India
3. String, G (2006): Linear Algebra and its applications, New Delhi, Engage learning India Pvt. Ltd.
4. Lang, S. (2014): Introduction to Linear Algebra, New Delhi, Springer India Pvt. Ltd.

Unit 04 Dual Vector Space

Structure

- 4.1. Introduction
- 4.2. Objectives
- 4.3. $\text{Hom}(V, W)$ as a vector space
- 4.4. Dual vector space
- 4.5. Dual basis of a vector space
- 4.6. Transpose of a linear transformation
- 4.7. Annihilator of a subset of a vector space
- 4.8. Rank of transpose of a linear transformation
- 4.9. Summary
- 4.10 Terminal Questions and Answers

4.1. INTRODUCTION

This unit we introduce $\text{Hom}(V, W)$ as a vector space, dual vector space, dual basis of a vector space. We also define Transpose of a linear transformation, Annihilator of a subset of a vector space and rank of transpose of a linear transformation.

4.2. Objectives

After studying this unit we should be able to:

- Define a $\text{Hom}(V, W)$ as a vector space;
- Define Dual vector space and dual basis of a vector space
- Define Transpose of a linear transformation
- Annihilator of a subset of a vector space;
- Rank of transpose of a linear transformation

4.3. $\text{Hom}(V, W)$ as a vector space

Let V and W be two vector spaces over the same field F .

We consider the set $\text{Hom}(V, W) = \{T \mid T: V \rightarrow W \text{ is a homomorphism}\}$.

That is $\text{Hom}(V, W)$ denotes the set of all homomorphism from $V \rightarrow W$.

Claim: $\text{Hom}(V, W)$ is a vector space over the same field F .

Proof: We define addition in $\text{Hom}(V, W)$ as if $T_1, T_2 \in \text{Hom}(V, W)$ then $T_1 + T_2: V \rightarrow W$ be a map defined by $(T_1 + T_2)(x) = T_1(x) + T_2(x) \quad \forall x \in V$.

We define scalar multiplication on $\text{Hom}(V, W)$ as if $T_1 \in \text{Hom}(V, W), \alpha \in F$, then $\alpha T_1: V \rightarrow W$ be a map given by $(\alpha T_1)(x) = \alpha T_1(x), \quad \forall x \in V$

We can easily check that $(T_1 + T_2)$ and (αT_1) are linear maps and they are homomorphism from $V \rightarrow W$. Also, we can easily prove that $(\text{Hom}(V, W), +, \cdot)$ is a vector space over the field F .

Theorem: If V and W be two finite dimensional vector spaces over the same field F , then $\dim(\text{Hom}(V, W)) = \dim V \cdot \dim W$

Proof: Let $\dim V = m$, and $\dim W = n$.

We can define mn linear transformations from $V \rightarrow W$ which will generate $\text{Hom}(V, W)$ and will be linearly independent.

Let $S = \{x_1, x_2, \dots, x_m\}$ and $S' = \{y_1, y_2, \dots, y_n\}$ be bases of V and W respectively. Let $T_{ij}: V \rightarrow W$ be a map where $i = 1, 2, 3, \dots, m$ and $j =$

$1, 2, \dots, n$ defined by $T_{ij}(x_k) = \delta_{ik}y_j$, for $k = 1, 2, 3, \dots, m$. Where $\delta_{ik} = 1$, if $i = k$ and $\delta_{ik} = 0$, if $i \neq k$. That is if $i = k = 1, j = 2$ then $T_{12}(x_1) = \delta_{11}y_2 = y_2$ and if $i = 1, k = 2, j = 2$ then $T_{12}(x_2) = \delta_{12}y_2 = 0y_2 = 0$.

Thus the image of (x_k) are either 0 or the basis element of W .

We know that each map T_{ij} from the basis $\{x_k\}$ of V to W will determine a unique linear transformation from $V \rightarrow W$. Also, these T_{ij} linear transformations are mn .

Now we want to show that these mn linear transformations are linearly independent.

Let $\alpha_{ij} \in F$ where $i = 1, 2, 3, \dots, m$ and $j = 1, 2, \dots, n$ Now,

$\sum_{i,j} \alpha_{ij} T_{ij} = 0$ (zero linear transformation)

$$= \sum_{i,j} \alpha_{ij} T_{ij}(x_k) = 0 \Rightarrow \sum_{i,j} \alpha_{ij} \delta_{ik} y_j = 0$$

$= \sum_{i,j} \alpha_{kj} \delta_{kk} y_j = 0 \Rightarrow \sum_j \alpha_{kj} y_j = 0 \Rightarrow \alpha_{kj} = 0$ where $k = 1, 2, 3, \dots, m$ and $j = 1, 2, \dots, n$ [since, $\{y_j\}$ is linearly independent]

This result is true for $k = 1, 2, 3, \dots, m$. Thus all the mn coefficients α_{ij} where $i = 1, 2, 3, \dots, m$ and $j = 1, 2, \dots, n$ are zeros.

Hence the set $\{T_{ij}\}$ is linearly independent.

Now, we want to prove that $\text{Hom}(V, W)$ is generated by $\{T_{ij}\}$.

Let $T \in \text{Hom}(V, W)$ be given by

$T(x_k) = \sum_{j=1}^n \beta_{kj} y_j$ where $k = 1, 2, 3, \dots, m$

Now, $(\sum_{j=1}^n \beta_{ij} T_{ij})(x_k) = \sum_j (\sum_{i=1}^m \beta_{ij} \delta_{ik} y_j) = \sum_j \beta_{kj} y_j = T(x_k)$ for $k = 1, 2, 3, \dots, m$

Thus it can be proved that $\forall x \in V, (\sum_{i,j} \beta_{ij} T_{ij})(x) = T(x)$.

Hence, $T = \sum_{i,j} \beta_{ij} T_{ij}$.

Thus, $\text{Hom}(V, W)$ is generated by $\{T_{ij}\}$, which is the basis of $\text{Hom}(V, W)$.

Hence, $\dim(\text{Hom}(V, W)) = mn = \dim V \cdot \dim W$

Definition: Let V, U, W be vector spaces over the same field F .

Let $T_1: V \rightarrow U$ and $T_2: U \rightarrow W$ be linear maps. Then the composite map $T_2 \circ T_1: V \rightarrow W$ is defined as the product $T_2 T_1$ of the two linear transformations T_2 and T_1 . We can easily verify that the composite map $T_2 \circ T_1: V \rightarrow W$ is also a linear map.

Theorem: The set $\text{Hom}(V, V)$ of all endomorphisms on a vector space V , is a ring with respect to addition and multiplication of linear transformations given as We define addition in $\text{Hom}(V, W)$ as if $T_1, T_2 \in \text{Hom}(V, W)$ then $T_1 + T_2: V \rightarrow W$ be a map defined by $(T_1 + T_2)(x) = T_1(x) + T_2(x) \forall x \in V$ and Let $T_1: V \rightarrow U$ and $T_2: U \rightarrow W$ be linear maps. Then the composite map $T_2 \circ T_1: V \rightarrow W$ is defined as the product $T_2 T_1$ of the two linear transformations T_2 and T_1 . The ring has identity for multiplication and is not in general commutative.

Note: The set $\text{Hom}(V, V)$ of all endomorphisms on a vector space V , is a vector space with respect to addition and scalar multiplication of linear transformations given as We define addition in $\text{Hom}(V, W)$ as if $T_1, T_2 \in \text{Hom}(V, W)$ then $T_1 + T_2: V \rightarrow W$ be a map defined by $(T_1 + T_2)(x) = T_1(x) + T_2(x) \forall x \in V$ and We define scalar multiplication on $\text{Hom}(V, W)$ as if $T_1 \in \text{Hom}(V, W), \alpha \in F$, then $\alpha T_1: V \rightarrow W$ be a map given by $(\alpha T_1)(x) = \alpha T_1(x), \forall x \in V$.

Theorem: Let V, U, W be finite dimensional vector spaces over the same field F .

Let $T: V \rightarrow U$ and $S: U \rightarrow W$ be linear maps. Then the composite map $ST: V \rightarrow W$. Then (i). $\text{Rank} ST$ can not exceed $\text{Rank} S$ or $\text{Rank} T$ and (ii). $\text{Nullity} ST \geq \text{Nullity} T$.

Proof: (i). Since, the composite map $ST: V \rightarrow W$ is a linear map.

Now, $T(V) \subseteq U \Rightarrow S(T(V)) \subseteq S(U) \Rightarrow ST(V) \subseteq S(U)$

Now, $ST(V)$ and $S(U)$ are both subspaces of W . So, $\dim(ST(V)) \leq \dim S(U)$.

Therefore, $\text{Rank} ST \leq \text{Rank} S$. Also, $T(V)$ is a subspace of U . Hence, the restriction of S to $T(V)$ may be considered as a linear map $S: (V) \rightarrow W$. We have $\dim S(T(V)) \leq \dim T(V)$

Or, $\dim(ST)(V) \leq \dim T(V)$. Hence, $\text{Rank} ST \leq \text{Rank} T$

(ii). Since, $x \in \text{Ker} T \Rightarrow T(x) = 0$

$\Rightarrow S(T(x)) = S(0) = 0 \Rightarrow ST(x) = 0 \Rightarrow x \in \text{Ker} ST$.

Hence, $\text{Ker} T \subseteq \text{Ker} ST$.

We know that $\text{Ker} ST$ and $\text{Ker} T$ are subspaces of V .

Hence, $\text{Ker} T$ is a subspace of $\text{Ker} ST$.

$\dim(\text{Ker} T) \leq \dim(\text{Ker} ST)$

So, $\text{Nullity} T \leq \text{Nullity}(ST)$

Definition: A linear transformation $T: V \rightarrow U$ is said to be singular if there exists a non-zero vector $v \in V$ such that $T(v) = 0$. On the other hand $T(v) = (0) \Rightarrow v = 0$. That is $\text{Ker}T = \{0\}$, then T is called non-singular. It means that T is non-singular if and only if T is injective. Also, we can say that a linear transformation $T: V \rightarrow U$ is non-singular if $\text{Rank}T = \dim V$.

Check your progress

1. Determine all non-singular linear transformation $T: R^4 \rightarrow R^3$.
2. Prove that a linear map $T: V \rightarrow U$ is non-singular if and only if the image of a linearly independent set is linearly independent.
3. Let $T: V \rightarrow V$ and $S: V \rightarrow V$ be linear maps. Then the composite map $ST: V \rightarrow V$, where V is a finite dimensional vector space, S is non-singular. Prove that $\text{Rank}(ST) = \text{Rank}(TS) = \text{Rank}T$.
4. If T, T_1 and T_2 are linear transformations on a vector space V such that $TT_1 = T_2T = I$, then T is invertible and $T^{-1} = T_1 = T_2$.
5. If V be an n - dimensional vector space, then a linear transformation T on V is invertible if and only if there exists a linear transformation T_1 on V such that (i). $TT_1 = I$ or (ii). $T_1T = I$.
6. If T_1 and T_2 are invertible linear transformations on a vector space V , then prove that (i). T_2T_1 is invertible and $(T_2T_1)^{-1} = T_1^{-1}T_2^{-1}$
(ii). $\beta \neq 0 \in F, \beta T_1$ is invertible and $(\beta T_1)^{-1} = \beta^{-1}T_1^{-1}$.
1. Let T be a non-singular linear transformation on an n - dimensional vector space V . If $S = \{x_1, x_2, \dots, x_n\}$ be a basis of V . then prove that $S' = \{T(x_1), T(x_2), \dots, T(x_n)\}$ is a basis of V .
2. If A is a linear transformation on V , such that $A^2 - A + I = 0$, then prove that A is invertible.
3. If k be a non-zero scalar, show that the linear map $T: V \rightarrow U$ is singular if and only if kT is singular.

Definition: Let V be a vector space over a field F . a map $f: V \rightarrow F$ is said to be a linear functional on V if for all $x, y \in V$ and $\forall \alpha, \beta \in F$ such that $f(\alpha x + \beta y) = \alpha f(x) + \beta f(y)$.

Note: The linear functional f on V is linear transformation from V to the one dimensional vector space F over the field F .

Example: 1. We define a map $\pi_i: F^n \rightarrow F$ by $f(x_1, x_2, \dots, x_n) = x_i$ where $x_i \in F$ then π_i is a linear functional on F^n .

2. We define a map $f: F^n \rightarrow F$ by $f(x_1, x_2, \dots, x_n) = \alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_n x_n$ where $x_i \in F$ and $\alpha_1, \alpha_2, \dots, \alpha_n$ scalars then f is a linear functional on F^n .

4.4. Dual vector space: The set $\text{Hom}(V, F)$ of all linear functional on V , being a vector space over F . This vector space is called the dual vector space or the conjugate vector space of V and is denoted by V^* .

Theorem: Let V be a finite dimensional vector space over F with a basis

$S = \{e_1, e_2, \dots, e_n\}$ then a map $f: \{e_1, e_2, \dots, e_n\} \rightarrow F$ can be extended to a unique linear functional ϕ on V such that $\phi(e_i) = f(e_i)$ for $i = 1, 2, 3, \dots, n$.

4.5 (a). Dual basis of a vector space

Theorem: Let V be a finite dimensional vector space over F with a basis $S = \{e_1, e_2, \dots, e_n\}$. There exists a uniquely determined basis $\{e^*_1, e^*_2, \dots, e^*_n\}$ of its dual space V^* such that $e^*_i(e_j) = \delta_{ij}$ for $j = 1, 2, 3, \dots, n$, where $\delta_{ij} = 0$ if $i \neq j$ and $\delta_{ij} = 1$ if $i = j$. Thus $\dim V = \dim V^*$.

Note: $S' = \{e^*_1, e^*_2, \dots, e^*_n\}$ is called the basis of the dual vector space V^* .

Proof: We know that for each $i = 1, 2, 3, \dots, n$, there exists a unique functional

$e^*_i \in V^*$ such that $e^*_i(e_j) = \delta_{ij}$ for $j = 1, 2, 3, \dots, n$, where $\delta_{ij} = 0$ if $i \neq j$ and $\delta_{ij} = 1$ if $i = j$. We want to show that $\{e^*_1, e^*_2, \dots, e^*_n\}$ is linearly independent;

$$\alpha_1 e^*_1 + \alpha_2 e^*_2 + \dots + \alpha_n e^*_n = 0, \text{ where } \alpha_i \in F$$

$$\Rightarrow (\alpha_1 e^*_1 + \alpha_2 e^*_2 + \dots + \alpha_n e^*_n)(e_i) = 0 \text{ for each } i = 1, 2, 3, \dots, n,$$

$$\Rightarrow (\alpha_1 e^*_1(e_i) + \alpha_2 e^*_2(e_i) + \dots + \alpha_n e^*_n(e_i)) = 0$$

$$\Rightarrow (\alpha_i e^*_i(e_i) = 0 \forall i \text{ [Since, } e^*_i(e_j) = 0 \text{ for } i \neq j \text{]})$$

$$\Rightarrow \alpha_i = 0 \forall i \text{ [Since, } e^*_i(e_i) = \delta_{ii} = 1 \text{]}$$

Hence the set $\{e^*_1, e^*_2, \dots, e^*_n\}$ is linearly independent.

Now we want to show that V^* is generated by $\{e^*_1, e^*_2, \dots, e^*_n\}$

Let $\phi \in V^*$ such that $\phi(e_i) = \beta_i$ for each $i = 1, 2, 3, \dots, n$, $\beta_i \in F$

Let $x = \alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_n x_n \in V$, $\alpha_i \in F$. Thus,

$$\begin{aligned}\phi(x) &= \phi(\alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_n x_n) \\ &= \alpha_1 \phi(x_1) + \alpha_2 \phi(x_2) + \dots + \alpha_n \phi(x_n)\end{aligned}$$

$$\text{Or, } \phi(x) = \alpha_1 \beta_1 + \alpha_2 \beta_2 + \dots + \alpha_n \beta_n \dots\dots\dots(1)$$

$$\text{Now, } e^*_i(x) = e^*_i(\alpha_1 e_1 + \alpha_2 e_2 + \dots + \alpha_n e_n) = \alpha_1 e^*_i(e_1) + \alpha_2 e^*_i(e_2) + \dots + \alpha_n e^*_i(e_n)$$

$$\text{Or, } e^*_i(x) = \alpha_i e^*_i(e_i) = \alpha_i \text{ for each } i = 1, 2, 3, \dots, n,$$

Hence, from equation (1) we have

$$\begin{aligned}\phi(x) &= \beta_1 e^*_1(x) + \beta_2 e^*_2(x) + \dots + \beta_n e^*_n(x) \\ &= \phi(x) = (\beta_1 e^*_1 + \beta_2 e^*_2 + \dots + \beta_n e^*_n)(x), \forall x \in V\end{aligned}$$

$$\text{Hence, } \phi = (\beta_1 e^*_1 + \beta_2 e^*_2 + \dots + \beta_n e^*_n)$$

Thus ϕ is a linear combination of $\{e^*_1, e^*_2, \dots, e^*_n\}$ which generate V^* .

Hence, $\{e^*_1, e^*_2, \dots, e^*_n\}$ is a basis of V^* which is called the dual of the basis

$S = \{e_1, e_2, \dots, e_n\}$ of V . Therefore, $\dim V = \dim V^*$.

Note: 1. Any vector $x \in V$ can be expressed as $x = e_1 e^*_1(x) + e_2 e^*_2(x) + \dots + e_n e^*_n(x)$.

2. Any linear functional $\phi \in V^*$ can be expressed as

$$\phi = \phi(e_1) e^*_1 + \phi(e_2) e^*_2 + \dots + \phi(e_n) e^*_n.$$

4.5 (b). Second Dual Space:

The set $\text{Hom}(V^*, F) = \{f | f: V^* \rightarrow F \text{ is a linear map}\}$, is called the second dual of V and is denoted by V^{**} .

Let $x \in V$. we consider a map $\hat{x}: V^* \rightarrow F$ is given by $\hat{x}(f) = f(x) \in F, \forall f \in V^*$

Claim: $\hat{x}$ is a linear functional on V^* .

$$\text{Let } f, g \in V^* \text{ and } \alpha, \beta \in F \text{ then } \hat{x}(\alpha f + \beta g) = (\alpha f + \beta g)(x) = \alpha f(x) + \beta g(x)$$

$$= \alpha \hat{x}(f) + \beta \hat{x}(g). \text{ Thus } \hat{x} \in V^{**}$$

Theorem: The set $\text{Hom}(V^*, F) = \{f | f: V^* \rightarrow F \text{ is a linear map}\}$, forms a vector space over F with respect to addition and scalar multiplication given by:

If $f \in V^*$ and $\hat{x}, \hat{y} \in V^{**}$, then

$$(\hat{x} + \hat{y})f = \hat{x}(f) + \hat{y}(f) = f(x) + f(y) = f(x+y) = \widehat{x+y}(f)$$

That is $(\hat{x} + \hat{y}) = \widehat{x+y}$ Also,

$$\alpha\hat{x}(f) = \alpha f(x) = f(\alpha x) = \widehat{\alpha x}(f)$$

That is $\alpha\hat{x} = \widehat{\alpha x}$, $\alpha \in F$.

Theorem: Let V be a finite dimensional vector space over a field F and V^{**} be its second dual then $V \cong V^{**}$.

Proof: we consider a map $T: V \rightarrow V^{**}$ defined by $T(x) = \hat{x}$ where $x \in V, \hat{x} \in \{e_i\}$

Let $x, y \in V$ and $\alpha, \beta \in F$, then

$$T(\alpha x + \beta y) = \widehat{\alpha x + \beta y} = \widehat{\alpha x} + \widehat{\beta y} = \alpha\hat{x} + \beta\hat{y} = \alpha T(x) + \beta T(y).$$

Hence, T is a homomorphism. Now if $x \in \ker T$, then $T(x) = \hat{x} = 0$ (zero linear functional on V^*). So, $\hat{x}(f) = 0, \forall f \in V^*$

Or, $f(x) = 0, \forall f \in V^*$. Hence, $x = 0$. Thus, $T(x) = 0 \Rightarrow x = 0$

So, $\ker(T) = \{0\}$. Hence, the map T is injective and therefore, it is surjective also, Since, V is finite dimensional. Thus T is an isomorphism and hence, $V \cong V^{**}$.

Note: If $\{e_i^*\}$ is a basis of V^* the dual basis to the basis $\{e_i\}$ of V then $\{e_i\}$ can be considered as a basis of V^{**} . It is the dual basis of the basis $\{e_i^*\}$ of V .

Example: Find the dual basis of the following basis of R^3 :

$$\{e_1 = (1, -1, 3), e_2 = (0, 1, -1), e_3 = (0, 3, -2)\}.$$

Solution: Suppose that required the dual basis be $\{e_1^*, e_2^*, e_3^*\}$. These are the linear functional on R^3 are given by

$$e_1^*(x, y, z) = a_1x + a_2y + a_3z$$

$$e_2^*(x, y, z) = b_1x + b_2y + b_3z$$

$e_3^*(x, y, z) = c_1x + c_2y + c_3z$, where $a_i, b_i, c_i \in R$ are to be determined which will make $\{e_1^*, e_2^*, e_3^*\}$ a basis of $(R^3)^*$

$$\text{Now, } 1 = e_1^*(e_1) = e_1^*(1, -1, 3) = a_1 - a_2 + 3a_3$$

$$0 = e_1^*(e_2) = e_1^*(0, 1, -1) = a_2 - a_3$$

$$0 = e_1^*(e_3) = e_1^*(0, 3, -2) = 3a_2 - 2a_3$$

Now, we solving these equations in a_1, a_2, a_3 , we get $a_1 = 1, a_2 = 0, a_3 = 0$

Thus, $e_1^*(x, y, z) = a_1x + a_2y + a_3z = 1.x + 0.y + 0.z = x$.

Similarly, we will get

$$e_2^*(x, y, z) = b_1x + b_2y + b_3z = 7x - 2y - 3z$$

$$\text{And } e_3^*(x, y, z) = c_1x + c_2y + c_3z = -2x + y + z$$

Example: Find the dual basis of the basis set $B = \{z_1, z_2\} = \{1 + i, 1 - i\}$ for the vector space $C(R)$.

Solution: Let $z = x + iy \in C$. if $\alpha, \beta \in R$ such that

$$\begin{aligned}x + iy &= \alpha(1 + i) + \beta(1 - i) \\&= (\alpha + \beta) + i(\alpha - \beta)\end{aligned}$$

Then solving these, we get

$$\alpha = \frac{1}{2}(x + y) \text{ and } \beta = \frac{1}{2}(x - y)$$

Let $B^* = \{f_1, f_2\}$ be the dual basis. then

$$\begin{aligned}f_1(z_1) &= 1, & f_1(z_2) &= 0 \\f_2(z_1) &= 0 \text{ and } f_2(z_2) &= 1 & (1) \\ \text{Thus } f_1(x + iy) &= \alpha f_1(z_1) + \beta f_1(z_2) \text{ , using (1)} \\&= \alpha(1) + \beta(0) \\&= \alpha = \frac{1}{2}(x + y) \text{ and } f_2(x + iy) = \alpha f_2(z_1) + \beta f_2(z_2) \text{ , using (1)} \\&= \alpha(0) + \beta(1) \\&= \beta = \frac{1}{2}(x - y)\end{aligned}$$

Hence the dual basis is $\{f_1, f_2\}$, where

$$f_1(x + iy) = \frac{1}{2}(x + y) \text{ and } f_2(x + iy) = \frac{1}{2}(x - y)$$

Example: Find the dual basis of the basis $\{(1,0), (1,1)\}$ of the vector space $R^2(R)$.

Solution: Let $(a, b) \in R^2$, if $\alpha, \beta \in R$ such that

$$\begin{aligned}(a, b) &= \alpha(1,0) + \beta(1,1) \\&= (\alpha + \beta, \beta)\end{aligned}$$

Then $\alpha + \beta = a$ and $\beta = b$ i. e, $\alpha = a - b$ and $\beta = b$

Let

Let $B^* = \{f_1, f_2\}$ be the dual basis. then

$$\begin{aligned}f_1(z_1) &= 1, & f_1(z_2) &= 0 \\f_2(z_1) &= 0 \text{ and } f_2(z_2) &= 1 & (1)\end{aligned}$$

$$\begin{aligned}
 \text{Thus } f_1(a, b) &= \alpha f_1(z_2) + \beta f_1(z_2) \text{ , using (1)} \\
 &= \alpha(1) + \beta(0) \\
 &= \alpha = (a - b) \text{ and } f_2(a, b) = \alpha f_2(z_2) + \beta f_2(z_2) \text{ , using (1)} \\
 &\quad \beta = b
 \end{aligned}$$

Hence the dual basis is $\{f_1, f_2\}$ where $f_1(a, b) = a - b$ and $f_2(a, b) = b$

Example: The vectors $\alpha_1 = (1, 1, 1)$, $\alpha_2 = (1, 1, -1)$ and $\alpha_3 = (1, -1, -1)$ form a basis of V_3 . If $\{f_1, f_2, f_3\}$ is the dual basis and if $\alpha = (0, 1, 0)$, find $f_1(\alpha)$, $f_2(\alpha)$, $f_3(\alpha)$
 Solution: As $\{f_1, f_2, f_3\}$ is a basis set of $V_3(C)$, for $\alpha \in V_3(C)$, we have

$$\alpha = a_1\alpha_1 + a_2\alpha_2 + a_3\alpha_3$$

Where $f_1(\alpha) = a_1$, $f_2(\alpha) = a_2$, $f_3(\alpha) = a_3$

$$\text{Now } \alpha = a_1\alpha_1 + a_2\alpha_2 + a_3\alpha_3 \quad (1)$$

$$\Rightarrow (0, 1, 0) = a_1(1, 1, 1) + a_2(1, 1, -1) + a_3(1, -1, -1)$$

$$\Rightarrow (0, 1, 0) = a_1 + a_2 + a_3, a_1 + a_2 - a_3, a_1 - a_2 - a_3$$

$$\Rightarrow a_1 + a_2 + a_3 = 0, a_1 + a_2 - a_3 = 1, a_1 - a_2 - a_3 = 0$$

$$\Rightarrow a_1 = 0, a_2 = \frac{1}{2}, a_3 = -\frac{1}{2}$$

Hence from (1), we get $f_1(\alpha) = 0$, $f_2(\alpha) = \frac{1}{2}$, $f_3(\alpha) = -\frac{1}{2}$

4.6. Transpose of a linear transformation: Let V, W be two vector spaces over the same field F . Let $T: V \rightarrow W$ be a linear transformation. Let $\phi: W \rightarrow F$ be a linear functional on W , then $\phi \circ T: V \rightarrow F$ will be a linear map. Since, composite of two linear maps is also linear. Thus $\phi \circ T$ is a linear functional on V . Thus for all $\phi \in W^*$ we get a unique element $\phi \circ T \in V^*$. Hence, the linear transformation $T: V \rightarrow W$ induces a map $T^t: W^* \rightarrow V^*$ given by $T^t(\phi) = \phi \circ T \in V^*$ where $\phi \in W^*$. The map $T^t: W^* \rightarrow V^*$ is also denoted by T^* which is linear. For if $\phi, \psi \in W^*$, $a, b \in F$, then

$$T^t(a\phi + b\psi) = (a\phi + b\psi) \circ T = a(\phi \circ T) + b(\psi \circ T) = aT^t(\phi) + bT^t(\psi)$$

The linear map map $T^t: W^* \rightarrow V^*$ defined as above is called the transpose of the linear map $T: V \rightarrow W$.

Theorem: Let $T: V \rightarrow W$ and $S: W \rightarrow Z$ be two linear transformations, then $(S \circ T)^t = T^t \circ S^t$.

Proof: Since, $T: V \rightarrow W$ and $S: W \rightarrow Z$ be two linear transformations then,

$(S \circ T): V \rightarrow Z$ be a linear transformation. So, $(S \circ T)^t: Z^* \rightarrow V^*$.

Again, $S^t: Z^* \rightarrow W^*$ and $T^t: W^* \rightarrow V^*$, so, $T^t \circ S^t: Z^* \rightarrow V^*$

From these we get $(S \circ T)^t$ and $T^t \circ S^t$ are both linear maps from $Z^* \rightarrow V^*$

Now we take, $z^* \in Z^*$, then $(S \circ T)^t(z^*) = z^* \circ (S \circ T) = (z^* \circ S) \circ T = S^t(z^*) \circ T$

$w^* \circ T$, where $w^* = S^t(z^*) \in W^*$.

$$= T^t(w^*) = T^t(S^t(z^*)) = (T^t \circ S^t)(z^*)$$

$$\text{Hence, } (S \circ T)^t = T^t \circ S^t.$$

Theorem: Let T and $S \in \text{Hom}(V, W)$ and $\alpha \in F$, then

$$(i). (T + S)^t = T^t + S^t$$

$$(ii). (\alpha T)^t = \alpha T^t.$$

Theorem: If V be a finite dimensional vector space then the linear transformation $T: V \rightarrow V$ is non-singular if and only if its transpose $T^t: V^* \rightarrow V^*$ is non-singular.

Proof: If $T: V \rightarrow V$ be a linear transformation and V be a finite dimensional vector space, then (a) $\text{Rank}(T^t) = \text{Rank}(T)$.

$$T \text{ is non-singular} \Leftrightarrow \text{Rank}(T) = \dim(V)$$

$$\Leftrightarrow \text{Rank}(T^t) = \dim(V),$$

$$\Leftrightarrow \text{Rank}(T^t) = \dim(V^*), \text{ (since } \dim(V) = \dim(V^*) \text{)}$$

$$\Leftrightarrow T^t: V^* \rightarrow V^* \text{ is non-singular.}$$

Example: Let ϕ be a linear functional on R^2 defined by $\phi(x, y) = x - 2y$.

Find the transpose of each of the linear transformation $T: R^2 \rightarrow R^2$ given by :

$$(a). T(x, y) = (x, 0),$$

$$(b). T(x, y) = (y, x + y),$$

$$(c). T(x, y) = (2x - 3y, 5x + 2y).$$

$$\text{Solution: (a). } (T^t(\phi))(x, y) = \phi(T(x, y)) = \phi(x, 0) = x.$$

$$(b). (T^t(\phi))(x, y) = \phi(T(x, y)) = \phi(y, x + y) = y - 2(x + y) = -2x - y$$

$$(c). (T^t(\phi))(x, y) = \phi(T(x, y)) = \phi(2x - 3y, 5x + 2y)$$

$$= 2x - 3y - 2(5x + 2y) = -8x - 7y.$$

4.7. Annihilators: Let V be a vector space over a field F , and S be a subset of V . The annihilator of S is denoted by S^0 and defined by

$$S^0 = \{\phi \in V^* | \phi(x) = 0, \forall x \in S\}$$

Remark: (1). S^0 is a subset of V^* .

(ii). Annihilator of $\{0\}$ is V^* .

(iii). V^0 is the zero subspace of V^* .

Theorem: : Let V be a vector space over a field F , and $\dim(V) = n$. Let W be a subspace of V then $\dim W + \dim W^0 = \dim V = n$.

Proof: Suppose $\dim W = r$ and let $S = \{x_1, x_2, \dots, x_r\}$ be a basis of W which can be extended to a basis of V say $B = \{x_1, x_2, \dots, x_r, x_{r+1}, \dots, x_n\}$ be a basis of V . Let $B' = \{x_1^*, x_2^*, \dots, x_r^*, x_{r+1}^*, \dots, x_n^*\}$ be its dual basis of V , that is ie is a basis of V^* .

Claim: To show that $D = \{x_{r+1}^*, \dots, x_n^*\}$ is a basis of W^0 .

Now we consider $x_i^*(x_j) = \delta_{ij} = 0$, [Since, $i \neq j$, as $i > r, j \leq r$] where $i = r+1, r+2, \dots, n$ and $j = 1, 2, \dots, r$.

Again, if $w = \sum_{j=1}^r a_j x_j \in W$, then for $i = r+1, r+2, \dots, n$, we have $x_i^*(w) = x_i^*(\sum_{j=1}^r a_j x_j) = \sum_{j=1}^r a_j x_i^*(x_j) = 0, \forall w \in W$. Thus $x_i^* \in W^0$ for $i = r+1, r+2, \dots, n$.

Thus the set $B' = \{x_1^*, x_2^*, \dots, x_r^*, x_{r+1}^*, \dots, x_n^*\} \subseteq W^0$ and the set is linearly independent, since it is subset of a basis of V^* .

Now if $\phi \in W^0$, then $\phi \in V^*$ [Since, $W^0 \subseteq V^*$]

So, ϕ can be expressed as $\phi = \sum_{j=1}^n \phi(x_j) x_j^*$

Since $\phi \in W^0$ so, $\phi(x_i) = 0$ for $i = 1, 2, \dots, r$.

Hence, $\phi = \phi(x_{r+1}) x_{r+1}^* + \phi(x_{r+2}) x_{r+2}^* + \dots + \phi(x_n) x_n^*$

Therefore, ϕ is generated by the set $D = \{x_{r+1}^*, \dots, x_n^*\}$ which is a basis of W^0

Thus $\dim(W^0) = n - r = n - \dim(W)$

Hence, $\dim W + \dim W^0 = \dim V = n$.

4.8. Rank of transpose of a linear transformation:

Theorem: If $T: V \rightarrow W$ be a linear transformation where V and W are finite dimensional, then $\text{rank}(T) = \text{rank}(T^t)$.

Proof: We want to firstly prove that $\ker(T^t) = (T(V))^0$ [annihilator of $T(V)$]

Now $T^t: W^* \rightarrow V^*, \phi \in \ker(T^t) \Leftrightarrow T^t(\phi) = 0 \Leftrightarrow \phi \circ T = 0$

$\Leftrightarrow \phi \circ T(x) = 0 \in F, \forall x \in V$

$$\Leftrightarrow \phi(T(x)) = 0 \in F, \forall T(x) \in T(V)$$

$$\Leftrightarrow \phi \circ T \in (T(V))^0$$

$$\text{So, } \ker(T^t) = (T(V))^0.$$

$$\text{Suppose } \dim(V) = n, \dim(W) = m \text{ and } \text{rank} T = \dim(T(V)) = r$$

We have, $\dim(T(V))^0 = \dim W - \dim(T(V))$ [since, $T(V)$ is a subspace of W].

$$\text{Hence, } \dim(T(V))^0 = \dim W - \dim(T(V)) = m - r$$

$$\text{So, } \dim(T(V))^0 = \dim(\ker T^t) = m - r$$

$$\text{So, Nullity}(T^t) = m - r$$

$$\text{Now, Nullity}(T^t) + \text{Rank}(T^t) = \dim W^*$$

$$\text{So, Rank}(T^t) = \dim W^* - \text{Nullity}(T^t) = m - (m - r) = r = \text{rank} T.$$

4.9. Summary :

After studying this unit we should be able to define $\text{Hom}(V, W)$ as a vector space, dual vector space, dual basis of a vector space. We also define Transpose of a linear transformation, Annihilator of a subset of a vector space and rank of transpose of a linear transformation.

4.10 Terminal Questions and Answers

1. If $A = \{(2,1), (3,1)\}$ be a basis of two dimensional vector space, find the dual basis of A .
2. Let $\{e_1 = (1,0,0), e_2 = (0,1,0), e_3 = (0,0,1)\}$ be the natural basis of $R^3(R)$. Determine the dual basis.
3. If $B = \{(1,1,1), (1,1,-1), (1,-1,1)\}$ is a basis of $V_3(R)$, then find the dual basis of B .
4. If $B = \{(1,-2,3), (1,-1,1), (2,-4,7)\}$ is a basis of $V_3(R)$, then find the dual basis of B .
5. Define a non-zero linear function f on R^3 such that if $x_1 = (1,1,1), x_2 = (1,1,-1)$ then $[x_1, f] = [x_2, f] = 0$



**Uttar Pradesh Rajarshi Tondon
Open University**

Bachelor of Science

UGMM - 107

Linear Algebra

BLOCK

2

Unit - 5	Matrices	107
Unit - 6	Rank of a Matrix	129
Unit - 7	Determinants	150
Unit - 8	Eigen vector and eigen space	164

Course Design Committee

Prof. Ashutosh Gupta, School of Computer and Information Science, UPRTOU, Prayagraj	Chairman
Prof. Sudhir Srivastav Dept. of Mathematics, DDU Gorakhpur University, Gorakhpur	Member
Prof. P. K. Singh Dept. of Mathematics, Allahabad University, Prayagraj	Member
Prof. Mona Khare Dept. of Mathematics, Allahabad University, Prayagraj	Member
Dr. A. K. Pandey Associate Professor, ECC Prayagraj	Member
Dr. Vikas Singh Academic Consultant, UPRTOU, Prayagraj	Member
Dr. S. S. Tripathi Academic Consultant, UPRTOU, Prayagraj	Member

Course Preparation Committee

Dr. Kamran Alam Khan Assistant Professor (Dept. of Mathematics), Ram Lubhai Sahani Govt. Mahila Degree College, Pilibhit	Author (Unit – 1 - 3)
Dr. Ashok Kumar Pandey Associate Professor, E.C.C, Prayagraj	Author (Unit 4 - 9)
Dr. S. S. Tripathi Academic Consultant School of Science, UPRTOU, Prayagraj	Editor (Unit 1 - 9)
Prof. Ashutosh Gupta Director, School of Computer and Information Science, UPRTOU, Prayagraj	

© UPRTOU , Prayagraj - 2021**© UGMM - 107 Under Algebra****ISBN :****Reprint – 2021 – 100 Copy**

All Rights are reserved. No Part of this work may reproduced in any form, by mimeograph or any other means, without permission in writing from the Uttar Pradesh Rajarshi Tandon Open University.

Printed and Published by Pro. (Dr.) Arun Kumar Gupta Registrar, Uttar Pradesh Rajarshi Tandon Open University, 2021

Printed By. – M/s K.C.Printing & Allied Works, Panchwati, Mathura -281003.

Unit 05

Matrices

Structure

- 5.1. Introduction
- 5.2. Objectives
- 5.3. Matrix
- 5.4. Matrix representation of a linear transformation
- 5.5. Equality of matrices
- 5.6. Algebra of Matrix
- 5.7. Multiplication of matrices
- 5.8. vector space of all $m \times n$ of matrices over a field
- 5.9. Ring of n – square matrices
- 5.10. Invertible matrix (non- singular matrix)
- 5.11. Transpose of a matrix
- 5.12. Equivalent matrices
- 5.13. Similar matrices
- 5.14. Orthogonal matrices
- 5.15. Summary
- 5.16 Terminal Questions

5.1. INTRODUCTION

This unit introduces matrices over number sets like sets of integers, rationals, reals and complex numbers. We define matrix representation of a linear transformation, equality of matrices, addition, subtraction and multiplication of matrices, vector space of all $m \times n$ of matrices over a field, ring of n – square matrices. We also define invertible matrix (non- singular matrix), transpose of a matrix, equivalent matrices, similar matrices and orthogonal matrices.

5.2. Objectives

After studying this unit we should be able to:

- Define matrix representation of a linear transformation;
- Equality of matrices;
- Define sum, difference and product of matrices;
- Vector space of all $m \times n$ of matrices over a field;
- Ring of n – square matrices;
- Invertible matrix (non- singular matrix);
- Define Transpose of Matrices;
- Equivalent matrices;
- Similar matrices;
- orthogonal matrices;

5.3. Matrix

In this block we shall deal with matrices, their algebra i.e. various ways of combining and altering them to produce other matrices, and their role in the solution of systems of linear equations, which will also involve the concept of determinant of so called square matrices. We shall also touch upon Boolean matrices and some operations and properties associated with them.

Definition: For positive integers m and n , an $(m \times n)$ matrix A stands for a rectangular array,

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & & & \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \quad (1)$$

of mn entities a_{ij} , where $1 \leq i \leq m$, $1 \leq j \leq n$. Note that the matrix does not represent a single value which could possibly be evaluated by some process.

The “entries” a_{ij} may belong to any definite set S whose members can be added, subtracted and multiplied to obtain member of S . For example S can be $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ which are standard notations for the number sets of all integer, rational numbers, real numbers and complex numbers, respectively. S can also be the set of all polynomials in one or more variables with the coefficients coming from any of the above mentioned number sets.

Example:

$$(1). \begin{bmatrix} -1 & 4 & 7 \\ 2 & 5 & -6 \end{bmatrix} \text{ is a } 2 \times 3 \text{ matrix 'over' } \mathbb{Z}$$

$$(2). \begin{bmatrix} 0 & -2 \\ -3 & 10 \\ 6 & 6 \end{bmatrix} \text{ is a } 3 \times 2 \text{ matrix over } \mathbb{Z}$$

$$(3). \begin{bmatrix} 2/3 & 0 \\ 1 & -5/6 \end{bmatrix} \text{ is a } 2 \times 2 \text{ matrix over } \mathbb{Q}$$

$$(4). [0, \frac{3}{4}, \pi, 2, -1] \text{ is a } 1 \times 5 \text{ matrix over } \mathbb{R}$$

$$(5). \begin{bmatrix} 4+2i & -6+\pi i \\ 0 & 3-\sqrt{5}i \end{bmatrix} \text{ is a } 2 \times 2 \text{ matrix over } \mathbb{C}$$

$$(6). \begin{bmatrix} x \\ 5-x^3 \\ 1+2x+6x^4 \end{bmatrix} \text{ is a } 3 \times 1 \text{ matrix with polynomial entries.}$$

The matrix of (1) is denoted by $A = [a_{ij}]$ where $1 \leq i \leq m$, $1 \leq j \leq n$.

If $m = n$ the matrix is said to be a **square matrix**. The above matrix A has m rows, (which are n -component vectors)

$R_i = [a_{i1}, a_{i2}, \dots, a_{in}]$ for $i = 1, 2, \dots, m$. and n columns, (which are m -component vectors)

$$C_j = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix} \text{ for } j = 1, 2, \dots, n.$$

5.4. Matrix representation of a linear transformation:

Let V, U be two vector spaces over the same field F of dimensions n and m respectively. Let $S = \{x_1, x_2, \dots, x_n\}$ and $S' = \{y_1, y_2, \dots, y_m\}$ be ordered bases of V and U respectively. Let $T: V \rightarrow U$ be a linear transformation, then $T(x_j) \in U$ for $j = 1, 2, 3, \dots, n$ and hence can be expressed uniquely as a linear combinations of elements of the basis $\{y_i\}$ of U as follows:

$$T(x_j) = a_{1j}y_1 + a_{2j}y_2 + a_{3j}y_3 + \dots + a_{mj}y_m$$

$$\text{Or, } T(x_j) = \sum_{i=1}^m a_{ij}y_i \text{ for } j = 1, 2, 3, \dots, n$$

Where $a_{ij} \in F$ are unique, $i = 1, 2, 3, \dots, m$ and $j = 1, 2, 3, \dots, n$

Thus the linear transformation $T: V \rightarrow U$ determines an unique ordered mn scalars a_{ij} which may be arranged in the matrix form as given

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

This matrix $[a_{ij}]$ is of order $m \times n$ and is called the matrix of the linear map $T: V \rightarrow U$ relative to the ordered bases $\{x_j\}$ of V and $\{y_i\}$ of U .

Conversely given an matrix $[a_{ij}]$ is of order $m \times n$ over a field F and two finite dimensional vector spaces V and U over F of dimensions n and m and bases $S = \{x_1, x_2, \dots, x_n\}$ and $S' = \{y_1, y_2, \dots, y_m\}$ respectively, we can find a unique linear map $T: V \rightarrow U$ whose matrix relative to the ordered bases $\{x_j\}$ of V and $\{y_i\}$ of U will be $[a_{ij}]$ as follows:

The linear map $T: V \rightarrow U$ defined by $T(x_j) = \sum_{i=1}^m a_{ij}y_i$ for $j = 1, 2, 3, \dots, n$

It can be extended over V by linearity. T is called the linear map associated with the matrix $[a_{ij}]$.

Remark: 1. The coordinates of $T(x_j)$ are the elements of the j^{th} column of the matrix of the linear transformation T . The matrix of a linear map and the linear map associated with a matrix depend on the choice and order of the bases of the vector spaces. If we change any of the bases or their orders, we find the different matrices for the same linear map.

2. For every $T \in \text{Hom}(V, U)$, there exists a unique $m \times n$ matrix over F (relative to fixed ordered bases of V and U). Also, conversely for every $m \times n$ matrix $[a_{ij}]$ over F , there exists a unique linear map $T: V \rightarrow U$ whose matrix relative to fixed ordered bases $\{x_j\}$ of V and $\{y_i\}$ of U will be $[a_{ij}]$.

Example: Let $0: V \rightarrow U$ be the zero linear transformation, then the matrix of this zero linear map relative to any ordered bases of V and U is

$$\begin{bmatrix} 1 & \dots & \dots & \dots & 0 \\ 0 & 1 & \dots & \dots & 0 \\ 0 & \dots & \dots & \dots & 1 \end{bmatrix}$$

In which each entry is $0 \in F$. It is called Null matrix or zero matrix.

Example: Let $I: V \rightarrow V$ be the identity linear transformation on an n - dimensional vector space V , then the matrix of this identity linear map relative to any ordered basis $S = \{x_1, x_2, \dots, x_n\}$ of V is

$$\begin{bmatrix} 1 & \dots & \dots & 0 \\ 0 & \dots & \dots & 0 \\ \vdots & & & \\ 0 & \dots & \dots & 1 \end{bmatrix}$$

In which $a_{ij} = 0$, if $i \neq j$ and $= 1$ if $i = j$ is $0 \in F$. It is called identity matrix this matrix is denoted by I_n . I_n is of ordered n and is independent of the choice or ordered of the basis $\{x_j\}$.

Example: Let $T: V \rightarrow V$ be the constant linear transformation on an n - dimensional vector space V defined by $T(x) = cx, \forall x \in$

$V, c(\neq 0) \in F$ then the matrix of this constant linear map relative to any ordered basis $S = \{x_1, x_2, \dots, x_n\}$ of V is

$$\begin{bmatrix} c & \dots & \dots & \dots & 0 \\ 0 & c & \dots & \dots & 0 \\ 0 & \dots & \dots & \dots & 1 \end{bmatrix}$$

In which $a_{ij} = 0$, if $i \neq j$ and $= c$ if $i = j$ is $c \in F$. It is called constant (scalar) matrix, this matrix is denoted by cI_n . cI_n is of ordered n and is independent of the choice or ordered of the basis $\{x_j\}$.

Example: Let $S = \{x_1, x_2, \dots, x_n\}$ be a basis of V . Let $T: V \rightarrow V$ be the diagonal linear transformation on an n -dimensional vector space V defined by $T(x_j) = \lambda_j x_j, \forall x_j \in S \subseteq V, \lambda_j (\neq 0) \in F$ then the matrix of this diagonal linear map relative to any ordered basis $S = \{x_1, x_2, \dots, x_n\}$ of V is

$$\begin{bmatrix} \lambda_1 & \dots & \dots & \dots & 0 \\ 0 & \lambda_2 & \dots & \dots & 0 \\ 0 & \dots & \dots & \dots & 1 \end{bmatrix}$$

In which $a_{ij} = 0$, if $i \neq j$ and $= \lambda_j$ if $i = j$ is $\lambda_j \in F$. It is called diagonal matrix, this matrix is denoted by $\lambda_n I_n$. $\lambda_n I_n$ is of ordered n and is independent of the choice or ordered of the basis $\{x_j\}$.

Check your progress

1. Prove that the matrix of D where $D \equiv \frac{d}{dx}$

(A). relative to the basis $S = \{x^2, x, 1\}$ is $\begin{bmatrix} 0 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

(B). relative to the basis $S = \{1, x, x^2\}$ is $\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}$

2. (a). Let T be the linear map on R^3 defined by $T(x, y, z) = (x, y, 0)$. Find the matrix of T relative to the usual basis of R^3 .

(b). Find the matrix of T relative to the basis $S = \{(1, 0, 0), (0, 1, 0), (1, 1, 0)\}$ of R^3 .

3. Let T be the linear map on R^3 defined by

$$T(x, y, z) = (a_1x + a_2y + a_3z, b_1x + b_2y + b_3z, c_1x + c_2y + c_3z).$$

Find the matrix of T relative to the usual basis of R^3 .

5.5. Equality of matrices:

Let V, U be two vector spaces over the same field F of dimensions n and m respectively. Let $S = \{x_1, x_2, \dots, x_n\}$ and $S' = \{y_1, y_2, \dots, y_m\}$ be ordered bases of V and U respectively. Let $T: V \rightarrow U$ and $S: V \rightarrow U$ be linear transformations whose matrices with respect to above bases be $[a_{ij}]$ and $[b_{ij}]$ respectively. Then

$$T = S \Leftrightarrow [a_{ij}] = [b_{ij}]$$

$$\text{Since, } T = S \Leftrightarrow T(x_j) = S(x_j) \Leftrightarrow \sum_{i=1}^m a_{ij}y_i = \sum_{i=1}^m b_{ij}y_i$$

$$\Leftrightarrow a_{ij} = b_{ij}, \forall i \text{ and } \forall j \Leftrightarrow [a_{ij}] = [b_{ij}]$$

That is If $A = [a_{ij}]$ and $B = [b_{ij}]$ be two $m \times n$ matrices over a field F then we can say that these two matrices are equal if $[a_{ij}] = [b_{ij}]$

$$\Leftrightarrow a_{ij} = b_{ij}, \forall i \text{ and } \forall j$$

5.6. Algebra of Matrix:

We define addition of two matrices of same size (order) $m \times n$ by entry wise sum:

$$\begin{bmatrix} a_{11} & \dots & a_{1n} \\ a_{m1} & \dots & a_{mn} \end{bmatrix} + \begin{bmatrix} b_{11} & \dots & b_{1n} \\ b_{m1} & \dots & b_{mn} \end{bmatrix} = \begin{bmatrix} a_{11} + b_{11} & \dots & a_{1n} + b_{1n} \\ a_{m1} + b_{m1} & \dots & a_{mn} + b_{mn} \end{bmatrix}$$

$$\text{In short if } A = [a_{ij}], B = [b_{ij}], 1 \leq i \leq m, 1 \leq j \leq n$$

$$\text{then } A + B = [a_{ij} + b_{ij}]$$

$$\text{Similarly } A - B = [a_{ij} - b_{ij}], 1 \leq i \leq m, 1 \leq j \leq m.$$

We may also multiply any matrix A over S by any member α of S (scalars)

$$\alpha \begin{bmatrix} a_{11} & \dots & a_{1n} \\ a_{m1} & \dots & a_{mn} \end{bmatrix} = \begin{bmatrix} \alpha a_{11} & \dots & \alpha a_{1n} \\ \alpha a_{m1} & \dots & \alpha a_{mn} \end{bmatrix}$$

The matrix O stands for a matrix of any size with each entry equal to zero, for matrices with entries from any of the above mentioned number sets or sets of polynomials. If $A = [a_{ij}]$ we denote $(-1)A = [-a_{ij}]$ by $-A$.

It may be easily verified by using the properties of addition, subtraction etc in the number sets, that for any of the matrices A, B, C of same order and for any $\alpha, \beta \in S$

1. $A + (B + C) = (A + B) + C$
2. $A + B = B + A$
3. $A + O = O + A = A$
4. $A + (-A) = O = (-A) + A$
5. $\alpha(A + B) = \alpha A + \alpha B$
6. $(\alpha + \beta)A = \alpha A + \beta A$
7. $(\alpha\beta)A = \alpha(\beta A)$
8. $1.A = A$

As expected, we denote the matrix $A + (-B)$ by $(A - B)$

Example:

$$1. \begin{bmatrix} 2 & \frac{3}{4} & -1 \\ \sqrt{3} & 2 & \pi \end{bmatrix} + \begin{bmatrix} -3 & \frac{1}{4} & -4 \\ -1 & -\pi & \sqrt{2} \end{bmatrix} = \begin{bmatrix} -1 & 1 & -5 \\ \sqrt{3}-1 & 2-\pi & \pi+\sqrt{2} \end{bmatrix}$$

as matrices over $\mathbb{R}$. (set of all reals).

$$2. \frac{3}{7} \begin{bmatrix} 5 & 7 \\ 2 & -5 \\ 3 & 0 \\ 2 & 0 \end{bmatrix} = \begin{bmatrix} \frac{15}{7} & 3 \\ \frac{2}{7} & -\frac{15}{7} \\ \frac{6}{7} & 0 \\ \frac{2}{7} & 0 \end{bmatrix} \text{ as matrices over } \mathbb{Q}.$$

5.7. Multiplication of matrices

Multiplication of matrices is not defined through multiplication of corresponding entries nor it is defined in general for matrices of same order.

Suppose that $A = [a_{ik}]$; $1 \leq i \leq m$, $1 \leq k \leq n$. And $B = [b_{kj}]$; $1 \leq k \leq n$, $1 \leq j \leq p$

be matrices of order $m \times n$ and $n \times p$ respectively. We define $AB = [c_{ij}]$ of order $m \times p$, where $c_{ij} = \sum_{k=1}^n a_{ik} b_{kj} = a_{i1}b_{1j} + a_{i2}b_{2j} + \dots + a_{in}b_{nj}$

That is the typical $(i, j)^{th}$ element of the product of AB is the sum of the products of elements of i^{th} row of A with the corresponding elements of the j^{th} column of B .

Note:1. For this multiplication the number of elements of i^{th} row of A = the number of elements of j^{th} column of B . Hence, the product AB can be found only when number of columns of A = number of rows of B .

2. In general matrix of $ST \neq$ matrix of TS .

3. matrix of $ST = (\text{matrix of } S) \cdot (\text{matrix of } T) \neq (\text{matrix of } T) \cdot (\text{matrix of } S) = \text{matrix of } TS . e. i. AB \neq BA$.

Hence, matrix multiplication is not commutative in general.

Let us denote the rows of A by $\underline{R}_1, \underline{R}_2, \dots, \underline{R}_m$ which are clearly n component vectors e.g.

$\underline{R}_i = (a_{i1}, a_{i2}, \dots, a_{in})$ for $i = 1, 2, \dots, m$. Similarly we denote the rows of B by $\underline{R}'_1, \dots, \underline{R}'_r$, which are r component vectors :

$\underline{R}'_k = (b_{k1}, \dots, b_{kn})$ ($k=1,2,\dots,n$). We shall denote the column of A and B by

$\underline{C}_k$ ($k = 1,2,\dots,n$) and $\underline{C}'_j$ ($j = 1, 2, \dots, r$) i.e.

$$\underline{C}_k = \begin{bmatrix} a_{1k} \\ \vdots \\ a_{mk} \end{bmatrix} \text{ and } \underline{C}'_j = \begin{bmatrix} b_{1j} \\ \vdots \\ b_{rj} \end{bmatrix}$$

which are m and n component vectors respectively. Recall that the scalar or dot product of two n component vectors

$\underline{x} = (x_1, \dots, x_n)$ and $\underline{y} = (y_1, \dots, y_n)$ is defined as the scalar

$$\underline{x} \cdot \underline{y} = x_1 y_1 + x_2 y_2 + \dots + x_n y_n$$

$\|\underline{x}\|$ of $\underline{x}$ by $\|\underline{x}\| = \sqrt{\underline{x} \cdot \underline{x}} = \sqrt{[x_1^2 + x_2^2 + \dots + x_n^2]}$ which is related to the length

We define the product of the above matrices A and B of orders $m \times n$ and $n \times r$ respectively, as the $m \times r$ matrix AB given by.

$$AB = \begin{bmatrix} \underline{R}_1 \cdot \underline{C}'_1 & \underline{R}_1 \cdot \underline{C}'_2 & \dots & \underline{R}_1 \cdot \underline{C}'_r \\ \underline{R}_2 \cdot \underline{C}'_1 & \underline{R}_2 \cdot \underline{C}'_2 & \dots & \underline{R}_2 \cdot \underline{C}'_r \\ \vdots & \vdots & \ddots & \vdots \\ \underline{R}_m \cdot \underline{C}'_1 & \underline{R}_m \cdot \underline{C}'_2 & \dots & \underline{R}_m \cdot \underline{C}'_r \end{bmatrix}$$

In other words if $AB = D = [d_{ij}]$, then

$$d_{ij} = R_i \cdot C'_j = a_{i1}b_{1j} + a_{i2}b_{2j} + \dots + a_{in}b_{nj} \text{ for } 1 \leq i \leq m, 1 \leq j \leq r$$

(Thumb rule: put the i^{th} row of A on the j^{th} column of B, multiply term by term and add to obtain d_{ij})

Example 1: Let $A = \begin{bmatrix} 1 & 0.5 & -2 & 3.2 \\ 0 & -4 & 6 & \sqrt{2} \end{bmatrix}$ and $B = \begin{bmatrix} 0 & -1 & \sqrt{3} \\ 6 & 0 & -5 \\ 1 & 4 & 0 \\ 2 & -2 & 3 \end{bmatrix}$

which are 2×4 and 4×3 matrices respectively. Thus the product AB will be the 2×3 matrix

$$\begin{aligned} & \begin{bmatrix} 1 \times 0 + .5 \times 6 + (-2) \cdot 1 + 3.2 \times 2, 1(-1) + .5 \times 0 + (-2)4 + 3.2(-2), 1 \times \sqrt{3} + .5 \times (-5) + (-2)0 + 3.2 \times \\ 0 \times 0 + (-4)6 + 6 \times 1 + \sqrt{2}(2), 0(-1) + (-4)0 + 6 \times 4 + \sqrt{2}(-2), 0(\sqrt{3}) + (-4)(-5) + 6(0) + \sqrt{2}(3) \end{bmatrix} \\ &= \begin{bmatrix} 7.4, & -15.4 & 7.1 + \sqrt{3} \\ -18 + 2\sqrt{2} & 24 - 2\sqrt{2} & 20 + 3\sqrt{2} \end{bmatrix} \end{aligned}$$

The above rule of matrix multiplication may seem weird and complicated but actually it has its roots in the composition of linear changes of variables (e.g. in the change of coordinates of a point when the coordinate axis in a plane are rotated about a line through the origin and perpendicular to the plane). Consider the successive change of variables

$$x'_1 = a_{11}x_1 + a_{12}x_2$$

$$x'_2 = a_{21}x_1 + a_{22}x_2$$

$$= \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

followed by $\begin{aligned} x''_1 &= b_{11}x'_1 + b_{12}x'_2 \\ x''_2 &= b_{21}x'_1 + b_{22}x'_2 \end{aligned}$

$$= \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} x'_1 \\ x'_2 \end{bmatrix}$$

then substituting for x'_1, x'_2 in second pair of equations from the first, we may express x''_1 and x''_2 in term of x_1 and x_2 :

$$x_1' = b_{11} (a_{11}x_1 + a_{12}x_2) + b_{12} (a_{21}x_1 + a_{22}x_2)$$

$$x_2' = b_{21} (a_{11}x_1 + a_{12}x_2) + b_{22} (a_{21}x_1 + a_{22}x_2)$$

$$\text{or } x_1' = (b_{11} a_{11} + b_{12} a_{21}) x_1 + (b_{11} a_{12} + b_{12} a_{22}) x_2$$

$$x_2' = (b_{21} a_{11} + b_{22} a_{21}) x_1 + (b_{21} a_{12} + b_{22} a_{22}) x_2$$

$$\therefore \begin{bmatrix} x_1' \\ x_2' \end{bmatrix} = \begin{bmatrix} b_{11}a_{11} + b_{12}a_{21} & b_{11}a_{12} + b_{12}a_{22} \\ b_{21}a_{11} + b_{22}a_{21} & b_{21}a_{12} + b_{22}a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$= \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Note that everywhere in the above calculations, the products of matrices or of matrices and column vectors (which are nothing but matrices with a single column) are according to the rule for multiplications of matrices defined earlier. So now it must be clear why matrix multiplication has been defined in the above manner.

Also from the above discussion, the following facts are evident

1. If for matrices A and B, the product AB is defined, the reverse product BA need not be defined. For, a product AB to be defined, the number of columns of A must be equal to the number of rows of B. In the example, considered earlier A was a 2×4 matrix and B was a 4×3 matrix. So AB was defined but as $3 \neq 2$, BA will not be defined.
2. In fact both AB and BA are defined only when one is say of order $m \times n$, and the other is of order $n \times m$. Even then the two matrices AB and BA will be of order $m \times m$ and $n \times n$ respectively and so they cannot be same if $m \neq n$.
3. Even if $m = n$, so that A, B, AB and BA are all of order $n \times n$, these products AB and BA are not equal in general. E.g.

$$\text{If } A = \begin{bmatrix} 2 & 0 \\ 3 & 1 \end{bmatrix} \text{ and } B = \begin{bmatrix} -1 & 2 \\ 3 & 0 \end{bmatrix} \text{ then}$$

$$AB = \begin{bmatrix} -2 & 4 \\ 0 & 6 \end{bmatrix} \neq \begin{bmatrix} 4 & 2 \\ 6 & 0 \end{bmatrix} = BA.$$

4. Since each of the number sets $\mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C}$ are closed under the arithmetic operations of addition, subtraction and multiplication, and the processes of addition, subtraction and multiplication of matrices involve the above operations only on their entries, sum, difference and

product of matrices over any of the above number sets is again a matrix over the same number set.

5. It may be readily verified that for matrices A, B, C:
 - (i) $(A B) C = A (BC)$ whenever either side is defined.
 - (ii) $A (B+C) = AB + AC$ whenever either side is defined.
 - (iii) $(A+B)C = AC + BC$ same way.
 - (iv) $(\alpha A) B = \alpha (AB) = A (\alpha B)$. for matrices over any of the above number sets S for every $\alpha \in S$ and whenever either of the above sides is defined.
6. For each positive integer m the $m \times m$ matrix $I_m = [\delta_{ij}]$ $1 \leq i, j \leq m$, where $\delta_{ij} = 1$ or 0 according as $i = j$ or not is called the identity matrix of size m. It satisfies $I_m A = A = A I_n$ for every $m \times n$ matrix A.
7. In order that $A+B$, AB and BA are all defined, clearly A and B should be square matrices of the same size.
8. Unlike as in the number systems, the products of two nonzero matrices (even when they both are square matrices of same size) can be a zero matrix 0. (each of whose entries are 0) e.g. If $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \neq 0$; $B =$

$$\begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix} \neq 0 \text{ but}$$

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

Example: $A = \begin{bmatrix} 1 & -2 & 5 \\ 0 & 3 & -7 \end{bmatrix}$

$$I_2 A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -2 & 5 \\ 0 & 3 & -7 \end{bmatrix} = \begin{bmatrix} 1 \times 1 + 0 \times 0 & 1 \times (-2) + 0 \times 3 & 1 \times 5 + 0 \times (-7) \\ 0 \times 1 + 1 \times 0 & 0 \times (-2) + 1 \times 3 & 0 \times 5 + 1 \times (-7) \end{bmatrix}$$

$$= \begin{bmatrix} 1 & -2 & 5 \\ 0 & 3 & -7 \end{bmatrix} = A$$

$$\text{And } A I_3 = \begin{bmatrix} 1 & -2 & 5 \\ 0 & 3 & -7 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 \times 1 + (-2) \times 0 + 5 \times 0 & 1 \times 0 + (-2) \times 1 + 5 \times 0 & 1 \times 0 + (-2) \times 0 + 5 \times 1 \\ 0 \times 1 + 3 \times 0 + (-7) \times 0 & 0 \times 0 + 3 \times 1 + (-7) \times 0 & 0 \times 0 + 3 \times 0 + (-7) \times 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -2 & 5 \\ 0 & 3 & -7 \end{bmatrix} = A$$

5.8. vector space of all $m \times n$ of matrices over a field: The set M_{mn} of all $m \times n$ matrices over a field F is a vector space over F with respect to addition and scalar multiplication of matrices as given below:

Let $A = [a_{ij}]$, $B = [b_{ij}]$ and $C = [c_{ij}]$ be $m \times n$ matrices. Then

$$\begin{aligned}(A + B) + C &= ([a_{ij}] + [b_{ij}]) + [c_{ij}] = [(a_{ij} + b_{ij}) + c_{ij}] \\ &= [a_{ij} + (b_{ij} + c_{ij})] = [a_{ij}] + ([b_{ij}] + [c_{ij}]) = A + (B + C).\end{aligned}$$

Hence, matrix addition is associative.

$$\text{Also, } A + B = [a_{ij}] + [b_{ij}] = [a_{ij} + b_{ij}] = [b_{ij} + a_{ij}] = [b_{ij}] + [a_{ij}] = B + A.$$

Hence, matrix addition is commutative.

$$\text{There is a unique } m \times n \text{ matrix } 0_{mn} \in M_{mn} \text{ such that } A + 0 = [a_{ij}] + 0_{mn} = [a_{ij} + 0] = [a_{ij}] = A.$$

Hence, null matrix 0_{mn} is the identity for the matrix addition.

$$\text{There is a unique } m \times n \text{ matrix } -A = -[a_{ij}] \in M_{mn} \text{ such that } A + -A = [a_{ij}] + -[a_{ij}] = [a_{ij} + -a_{ij}] = [0] = 0_{mn}.$$

Hence, the matrix $-A = -[a_{ij}] = [-a_{ij}]$ is the additive inverse of the matrix $A = [a_{ij}]$.

Thus the pair $(M_{mn}, +)$ is an abelian group.

Now, we define scalar multiplication as follows:

Let $\alpha \in F$, $A = [a_{ij}] \in M_{mn}$, then $\alpha A = \alpha[a_{ij}] = [\alpha a_{ij}]$ which satisfied the following conditions.

$$\begin{aligned}\text{(a). } \alpha(A + B) &= \alpha([a_{ij}] + [b_{ij}]) = \alpha[a_{ij} + b_{ij}] = [\alpha(a_{ij} + b_{ij})] \\ &= [\alpha a_{ij} + \alpha b_{ij}] = [\alpha a_{ij}] + [\alpha b_{ij}] = \alpha[a_{ij}] + \alpha[b_{ij}] = \alpha A + \alpha B.\end{aligned}$$

$$\begin{aligned}\text{(b). } (\alpha + \beta)A &= (\alpha + \beta)[a_{ij}] = [(\alpha + \beta)a_{ij}] = [\alpha a_{ij} + \beta a_{ij}] = [\alpha a_{ij}] + [\beta a_{ij}] \\ &= \alpha[a_{ij}] + \beta[a_{ij}] = \alpha A + \beta A.\end{aligned}$$

$$\begin{aligned}\text{(c). } \alpha(\beta A) &= \alpha(\beta[a_{ij}]) = \alpha([\beta a_{ij}]) = [\alpha(\beta a_{ij})] \\ &= [\alpha \beta a_{ij}] = (\alpha \beta)[a_{ij}] = (\alpha \beta)A.\end{aligned}$$

$$(d). 1.A = 1.[a_{ij}] = [1.a_{ij}] = [a_{ij}] = A.$$

Thus the set of all $m \times n$ matrices over a field F is a vector space over F with respect to addition and scalar multiplication of matrices.

Theorem: If V and W are vector spaces over the field F of dimensions n and m respectively, then $\text{Hom}(V, W) \cong M_{mn}$ (the set of all $m \times n$ matrices over a field F) and hence, dimension of the vector space M_{mn} is mn .

Hint: We define a map $\phi: \text{Hom}(V, W) \rightarrow M_{mn}$ as if $T \in \text{Hom}(V, W)$, then $\phi(T) = \text{matrix of } (T) = [a_{ij}]$. clearly this map is a bijective homomorphism. Hence, $\text{Hom}(V, W) \cong M_{mn}$.

Note: The set of all symmetric matrices of order $n \times n$ is also a vector space which is a subspace of $M_{nn} = M_n$.

5.9. Ring of n – square matrices: The set M_n of all $n \times n$ matrices over a field F is a ring respect to addition and multiplication of matrices as given below:

We already know that the pair $(M_n, +)$ is an abelian group where $+$ is the addition of the n – square matrices.

We define matrix multiplication as follows: let A, B be two n – square matrices, then AB is also an n – matrix, hence the set M_n is closed under matrix multiplication.

Let V_1, V_2, V_3, V_4 be vector spaces over the field F of dimension n .

Let $T_1: V_1 \rightarrow V_2, T_2: V_2 \rightarrow V_3$ and $T_3: V_3 \rightarrow V_4$ be linear transformation, such that matrices of T_3, T_2, T_1 with respect to suitable ordered bases be A, B, C respectively each of order $n \times n$. Then

$$T_3(T_2.T_1): V_1 \rightarrow V_4 \text{ and } (T_3T_2) T_1: V_1 \rightarrow V_4$$

$$\text{Hence, } T_3(T_2.T_1) = (T_3T_2) T_1.$$

$$\text{So, } \text{matrix}(T_3(T_2.T_1)) = \text{matrix}(T_3T_2) T_1)$$

$$= \text{matrix of } T_3, \text{matrix of } (T_2.T_1) = \text{matrix}(T_3T_2). \text{matrix of } T_1$$

$$= A.(B.C) = (A.B).C$$

Hence, multiplication of matrices is associative.

Let $T_1: V_1 \rightarrow V_2, T_2: V_1 \rightarrow V_2$ and $T_3: V_2 \rightarrow V_3$ be linear transformations and each vector spaces over the field F of dimension n . Let matrices of T_1, T_2, T_3

with respect to suitable ordered bases be A, B, C respectively each of order $n \times n$. Then

$$\text{We have, } T_3(T_1 + T_2) = T_3T_1 + T_3T_2.$$

Since, multiplication of linear transformations is distributive over addition. Hence,

$$\begin{aligned} \text{matrix of } T_3(T_1 + T_2) &= \text{matrix of } (T_3T_1 + T_3T_2) \\ &= \text{matrix of } T_3(T_1 + T_2) = \text{matrix of } (T_3T_1) + \text{matrix of } (T_3T_2) \\ &= \\ \text{matrix of } T_3, \text{matrix of } (T_1 + T_2) &= \text{matrix of } (T_3T_1) + \text{matrix of } (T_3T_2) \end{aligned}$$

$$\text{Or, } C.(A + B) = C.A + C.B$$

Similarly, we can prove that $(A + B).C = A.C + B.c$

Thus, multiplication of matrices is distributive over addition of matrices.

Hence, the set M_n satisfies all the conditions of ring, so $(M_n, +, \cdot)$ is a ring.

The ring has multiplicative identity viz the identity matrix I_n .

In general multiplication of the matrices is not commutative, so the ring is not commutative in general.

Note: If A, B are n - square matrices and $\alpha \in F$ then, $(\alpha A)B = \alpha(AB)$

Check your progress

1. Let V be a vector space of all 2×2 symmetric real matrices. Show that $\dim V = 3$. Also, find the usual basis of V .
2. Let V be a vector space of all $n \times n$ matrices over F . Let M be a matrix in V and let $T: V \rightarrow V$ be a map defined by $T(A) = AM \pm MA, \forall A \in V$. Prove that T is linear and find $\text{Ker} T$ in the second case.
3. Let M_n be the set of all invertible n - square matrices over a field F . Show that M_n is a group under matrix multiplication. Is this group abelian?

5.10. Invertible matrix (non- singular matrix):

An $n \times n$ (square) matrix A is said to be Invertible matrix (non- singular matrix) if there is a square matrix B of the same size $n \times n$ (square), such that

$$AB = BA = I_n \text{ i.e. if } A = [a_{ij}] \text{ and}$$

$$B = [b_{ij}] \text{ for } 1 \leq i, j, \leq n.$$

then $\sum_{k=1}^n a_{ik} b_{kj} = \delta_{ij} = \sum_{k=1}^n b_{kj} a_{ik} \forall 1 \leq i, j \leq n.$

where $\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$ is the $(i, j)^{\text{th}}$ element of the identity matrix I_n .

Example: let $A = \begin{bmatrix} 1 & 1 \\ 2 & 3 \end{bmatrix}$ and

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

$$\text{Then } AB = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2 = BA.$$

$\therefore A$ and B are both invertible.

Remark: If $AB = I_n = BA$ for $n \times n$ matrices, we say that A and B are inverse matrix of each other and write $B = A^{-1}$ and $A = B^{-1}$.

Note: that if the inverse of an invertible matrix A exists then it is unique, for if both B_1 and B_2 are inverse matrices of the same matrix A then,

$$\begin{aligned} B_1 &= I_n B_1 = (B_2 A) B_1 \quad (\because B_2 A = I_n) \\ &= B_2 (AB_1) \quad (\text{by associativity of matrix product}), = B_2 I_n = B_2. \end{aligned}$$

If A and B are invertible n - square matrices, then AB is also, invertible and $(AB)^{-1} = B^{-1}A^{-1}$ (reversal rule).

$$\text{Since, } B^{-1}A^{-1}(AB) = B^{-1}(A^{-1}A)B = B^{-1}(I_n B) = B^{-1}B = I_n.$$

$$\text{Similarly, } (AB)B^{-1}A^{-1} = I_n.$$

Hence, AB is invertible and $(AB)^{-1} = B^{-1}A^{-1}$.

In general this result can be generalized as if $A_1, A_2, \dots, A_n$ are n - square invertible matrices, then $(A_1 A_2 \dots A_n)^{-1} = A_n^{-1} A_{n-1}^{-1} \dots A_2^{-1} A_1^{-1}$

Remark: It is clear that all square matrices are not invertible.

Suppose that $A = \begin{bmatrix} 3 & 6 \\ 2 & 4 \end{bmatrix}$ is invertible and

$B = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is the (in view of the uniqueness of inverse seen above) inverse of A then we have

$$AB = \begin{bmatrix} 3a+6c & 3b+6d \\ 2a+4c & 2b+4d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \text{ which means,}$$

$$3a + 6c = 1, \text{-----(1)}$$

$$3b + 6d = 0, \text{-----(2)}$$

$$2a + 4c = 0 \quad \text{and} \text{-----(3)}$$

$$2b + 4d = 1. \text{-----(4)}$$

Multiplying the equation (1) by 2 and equation (3) by 3 and subtracting, gives $2 = 0$; clearly it is not possible. Therefore the above matrix A is not invertible.

Note: We shall have several methods to detect whether a given square matrix is invertible and in case it is so, then for finding their inverse matrices. At this point we only make this remark that invertible matrices over the number sets $\mathbb{Q}$, $\mathbb{R}$, or $\mathbb{C}$ have their inverse matrices with entries in these sets only respectively but an inverse of an invertible matrix over $\mathbb{Z}$ is in general a matrix over $\mathbb{Q}$. The reason for this will be clear later on.

5.11. Transpose of a matrix:

If the matrix $A = [a_{ij}]$ of order $m \times n$, then the matrix $B = [b_{ji}]$ of order $n \times m$ is called the transpose of A if $a_{ij} = b_{ji}$ for $\forall i$ and $\forall j$ i.e. $(i, j)^{\text{th}}$ element of $A = (j, i)^{\text{th}}$ element of B . Hence, rows of A are the columns of B and vice-versa. Transpose of A is written as A^T .

Given any $m \times n$ matrix A , the matrix A^T obtained by writing the rows as columns and vice-versa gives an $n \times m$ matrix called the transpose of the matrix A . That is,

$$\begin{bmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix}^T = \begin{bmatrix} a_{11} & \dots & a_{m1} \\ a_{12} & \dots & a_{n2} \\ \vdots & & \vdots \\ a_{1n} & \dots & a_{mn} \end{bmatrix}$$

Example:

$$\begin{bmatrix} 1 & 0 & -6 \\ 2 & 4 & 5 \\ 0 & -3 & 3 \\ 4 & 0 & -2 \end{bmatrix}^T = \begin{bmatrix} 1 & 2 & 0 & 4 \\ 0 & 4 & -3 & 0 \\ -6 & 5 & 3 & -2 \end{bmatrix}$$

Note:1. A matrix A is said to be a square matrix if it is of order $n \times n$ for some positive integer n i.e. it has the same number of rows and columns.

2. A square matrix A is symmetric if $A = A^T$ and skew symmetric if $A = -A^T$. clearly, it is not necessary for a square matrix to be of either of these types.

Thus $A = [a_{ij}]$ is symmetric iff $a_{ij} = a_{ji} \forall 1 \leq i, j \leq n$ and is skew symmetric iff $a_{ij} = -a_{ji} \forall 1 \leq i, j \leq n$

Example:

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & -6 & -5 \\ 3 & -5 & 0 \end{bmatrix} \text{ is symmetric whereas}$$

$$B = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \text{ is skew symmetric.}$$

$$C = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \text{ is not symmetric nor skew symmetric.}$$

Theorem: Prove that the product AB of two symmetric matrices A and B is symmetric if and only if $AB = BA$.

Proof: Let $A = [a_{ij}]$ and $B = [b_{ij}]$ be two n – square symmetric matrices. Then $a_{ij} = a_{ji}$ and $b_{ij} = b_{ji} \forall 1 \leq i, j \leq n$.

Hence, $A = A^T$ and $B = B^T$.

$$\text{Now } (AB)^T = B^T A^T = BA$$

If A and B are symmetric matrices, then $(AB)^T = BA$.

Now, if AB is symmetric, then $(AB)^T = AB$

Hence, $AB = BA$.

Conversely, if $AB = BA$ and $(AB)^T = BA = AB$

Hence, $(AB)^T = AB$. So, AB is symmetric.

Check your progress

1. If $A = \begin{bmatrix} 1 & 2 & -3 \\ 4 & 0 & 2 \\ 7 & -1 & 0 \end{bmatrix}$ and

$$B = \begin{bmatrix} 2 & 6 & 0 \\ -3 & 5 & -2 \\ -1 & 7 & 4 \end{bmatrix}$$

Find (1). $A + B, A - B, AB$ and BA .

(2). (i). $(A + B)^T = A^T + B^T$; (ii). $(AB)^T = B^T A^T$

(3). $(AB)^{-1} = B^{-1}A^{-1}$; $(\alpha A)^T = \alpha A^T$, where $\alpha \in F$

2. Construct two nonzero matrices A and B of order 3×3 for which $AB \neq BA$.

3. Construct two nonzero 3×3 matrices whose product BA is a zero matrix.

4. Prove that if A is an invertible matrix then for every nonzero square matrix B of same order as A , the product $AB \neq 0$ (zero matrix).

5. Show that an n - square matrix $A = [a_{ij}]$ is non-singular if and only if A is invertible.

6. Product of two non-singular matrices is non-singular.

7. Product of two matrices one of which is singular is singular.

5.12. Equivalent matrices: Two $m \times n$ matrices (A) and $(\bar{A})$ are said to be equivalent if they are the matrices of the same linear transformation $T: V \rightarrow U$, relative to different pairs of bases and we write $(A) \cong (\bar{A})$.

Note: We see that if (A) and $(\bar{A})$ are equivalent, then there exists non-singular matrices (γ) and (μ) such that $(\bar{A}) = (\gamma)(A)(\mu)$.

The converse of this statement is also true.

That is two $m \times n$ matrices (A) and $(\bar{A})$ are equivalent if and only if there exists non-singular matrices (γ) and (μ) of order m and n respectively such that $(\bar{A}) = (\gamma)(A)(\mu)$.

5.13. Similar matrices: Two n – square matrices are called similar if they are the matrices of the same linear transformation T on a vector space V of dimension n relative to different bases.

Note: Two $n \times n$ matrices (A) and $(\bar{A})$ are similar if and only if there exists a non- singular matrix (μ) of order n such that $(\bar{A}) = (\mu)(A)(\mu)^{-1}$.

Theorem: A relation R is defined in the set all $m \times n$ matrices as follows:

$(A)R(\bar{A})$ if $(A) \cong (\bar{A})$, then R is an equivalence relation where (A) and $(\bar{A})$ are $m \times n$ matrices.

Proof: (I). $(A)R(A)$ since $(A) = I(A)I$ where I is non- singular matrix. So, R is a reflexive relation.

(II). $(A)R(\bar{A}) \Rightarrow (\bar{A}) = (\gamma)(A)(\mu)$ where non- singular matrices (γ) and (μ) of order m and n respectively.

$$\Rightarrow (A) = (\gamma)^{-1}(\bar{A})(\mu)^{-1} \Rightarrow (\bar{A})R(A)$$

Hence, R is symmetric.

(III). $(A)R(B) \Rightarrow (B) = (\gamma_1)(A)(\mu_1)$ where non- singular matrices (γ_1) and (μ_1) of order m and n respectively and $(B)R(C) \Rightarrow (C) = (\gamma_2)(B)(\mu_2)$ where non- singular matrices (γ_2) and (μ_2) of order m and n respectively.

$$(C) = (\gamma_2)((\gamma_1)(A)(\mu_1))(\mu_2) \Rightarrow (C) = ((\gamma_2)(\gamma_1))(A)((\mu_1)(\mu_2))$$

$\Rightarrow (C) = (\gamma_3)(A)(\mu_3)$ where, $(\gamma_3) = ((\gamma_2)(\gamma_1))$ and $(\mu_3) = ((\mu_2)(\mu_1))$ are non- singular matrices. $\Rightarrow (A)R(C)$.

So, R is transitive . Hence, R is an equivalence relation.

Check your progress

1. Consider the bases of R^3 : $\{e_1 = (1,0,0), e_2 = (0,1,0), e_3 = (0,0,1)\}$ and $\{f_1 = (1,1,1), f_2 = (1,1,0), f_3 = (1,0,0)\}$, find the transition matrix (A) from $\{e_i\}$ to $\{f_i\}$ and the transition matrix (B) from $\{f_i\}$ to $\{e_i\}$ and verify that $(B) = (A)^{-1}$.

2. Let $S = \{x_1, x_2, \dots, x_n\}$ be a basis of a vector space V over a field F . We defined $y, y_2, \dots, y_n$ as follows:

$y_j = \sum_{i=1}^n a_{ij} x_i = a_{1j}x_1 + a_{2j}x_2 + \dots + a_{nj}x_n$ where, $j = 1, 2, 3, \dots, n$ and $a_{ij} \in F$ then prove that $\{y, y_2, \dots, y_n\}$ is a basis of V if and only if the matrix $[a_{ij}]$ is invertible.

3. Prove that (i). $tr(AB) = tr(BA)$ and (ii). If A is similar to B then $tr(A) = tr(B)$ where the trace of an n – square matrix $A = [a_{ij}]$ is written as $tr(A)$ and defined as $tr(A) = \sum_{i=1}^n a_{ii} = a_{11} + a_{22} + \dots + a_{nn}$.

5.14. Orthogonal matrices: An n – square matrix A is said to be orthogonal if $A^T A = I$.

Example: 1. $\begin{bmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{bmatrix}$ 2. $\begin{bmatrix} \frac{2}{3} & \frac{2}{3} & \frac{1}{3} \\ -\frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \end{bmatrix}$ are orthogonal matrices.

Corollary: If An n – square matrix A is orthogonal then A is non- singular and (i). $A^{-1} = A^T$ (ii). $|A| = \pm 1$.

Proof: Since A is orthogonal so, $A^T A = I \Rightarrow A(A^T A) = A.I = A$

$$\Rightarrow (AA^T)A = A \Rightarrow AA^T = I$$

Thus, $A^T A = I = AA^T$. So, A is non- singular and $A^T = A^{-1}$

(ii). $|A^T A| = |I| = 1$, or, $|A^T A| = 1$

Or, $|AA| = 1$, Since, $|A^T| = |A|$

Or, $|A|^2 = 1$, so, $|A| = \pm 1$

Check your progress

1. Prove that product of two orthogonal matrices is orthogonal.
2. Prove that inverse of an orthogonal matrix is orthogonal.
3. Prove that transpose of an orthogonal matrix is orthogonal.
4. Show that the set of orthogonal matrices forms a group with respect to multiplication of matrices.

5.15. Summary :

After studying this unit we will be able to introduces matrices over number sets like sets of integer, rational, real and complex numbers. We define matrix representation of a linear transformation, equality of matrices, addition, subtraction and multiplication of matrices, vector space of all $m \times n$ of matrices over a field, ring of n – square matrices. We also define invertible matrix (non- singular matrix), transpose of a matrix, equivalent matrices, similar matrices and orthogonal matrices.

5.16 Terminal Questions

Q.1- Verify whether the following matrix is orthogonal or not. Also find A^{-1} .

$$A = \begin{bmatrix} \cos \theta & 0 & \sin \theta \\ 0 & 1 & 0 \\ -\sin \theta & 0 & \cos \theta \end{bmatrix}.$$

Hint: for orthogonal $AA^T = I$ & $A^{-1} = A^T$.

Q.2- If $A = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} & a \\ \frac{2}{3} & \frac{1}{3} & b \\ \frac{2}{3} & -\frac{2}{3} & c \end{bmatrix}$ is orthogonal then find a, b, c.

Ans. $\left(a = \frac{2}{3}, b = -\frac{2}{3}, c = \frac{1}{3}\right)$

Q.3- Show that $A + A^T$ is symmetric and $A - A^T$

is skew symmetric and hence show that matrix A is the sum of a symmetric matrix & a skew symmetric.

Q.4- Find the inverse of the matrix

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

UNIT-06

Rank of a Matrix

Structure

- 6.1. Introduction
- 6.2. Objectives
- 6.3. Rank of a Matrix
- 6.4. Row rank and Column rank of a matrix
- 6.5. Elementary matrices
- 6.6. Elementary Row and column operations of a matrix
- 6.7. Normal form of a matrix
- 6.8. Echelon form of matrix
- 6.9. Inverse of a non- singular matrix
- 6.10. Determinantal rank of a matrix
- 6.11. Non homogeneous and homogeneous linear equations
- 6.12. Summary
- 6.13 Terminal Questions

6.1. INTRODUCTION

In this unit we introduce linear systems and two methods for solving it. In this unit we discuss a method for solving a particular type of linear system. You can find a detailed study in our course on linear algebra. In this unit we consider elementary row operations and Gaussian elimination, also.

In this unit we also discuss a method which uses determinants for solving certain systems of linear equations. This method is due to the eighteenth century mathematician Cramer. It only applies to some of those linear systems in which the number of variables equals the number of equations.

6.2. Objectives

After studying this unit you should be able to:

- To find the Rank of a Matrix
- To find Row rank and Column rank of a matrix
- Elementary matrices
- Elementary Row and column operations of a matrix
- To find Normal form of a matrix
- To find Echelon form of matrix
- To find Inverse of a non- singular matrix
- To find Determinantal rank of a matrix
- Identify the linear systems which can be solved by using Cramer's rule, and apply the rule to solve them. Also to find Gaussian Elimination to solve systems of linear equations

Now let us look at a convenient way of representing a linear system.

6.3. Rank of a matrix: Let, A be a $m \times n$ matrix whose elements are members of the field F . Let k be any natural number such that $k \leq \min(m, n)$, where $\min(m, n)$ denotes the smaller of m and n .

Definition: The natural number r is called the rank of a $m \times n$ matrix A if

- (i). There is at least one non- zero minor of order r .
- (ii). Every minor of order $(r + 1)$ is zero.

In fact every minor of order greater than r will be zero. That is rank of a matrix A is the largest order of a non- zero minor of the matrix

Note: 1. There exists a non- zero minor of order r then the rank of A is greater than or equal to r .

2. All minors of order $(r + 1)$ are zero, implies the rank of A is less than or equal to r .

3. The rank of a matrix A is greater than or equal to 0, rank of zero matrix is zero. In particular the rank of every n – rowed non- singular matrix A is n . The rank of a matrix A is denoted by the symbol $\rho(A)$.

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

Since, $\det. (A) = |A| = 2 \neq 0$, so the rank of A is $\rho(A) = 3$

Example: $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ -3 & -6 & -9 \end{bmatrix}$

Since $|A| = 0$ and each of the second order minor is also 0. Also, the matrix is non- zero. so the rank of A is $\rho(A) = 1$

Example: $A = \begin{bmatrix} 2 & 3 & 4 \\ 3 & 1 & 2 \\ -1 & 2 & 2 \end{bmatrix}$

Since, $|A| = 0$ and there is a non- zero minor of order 2, viz $\begin{bmatrix} 2 & 3 \\ 3 & 1 \end{bmatrix}$ such that its determinant is $-7 \neq 0$, so the rank of A is $\rho(A) = 2$

Note: The rank of a transpose of a matrix A is the same as the rank of the original matrix A .

Check your progress

Determine the rank of the following matrices:

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

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

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

7. $A = \begin{bmatrix} 3 & -1 & 2 \\ -6 & 2 & 4 \\ -3 & 1 & 2 \end{bmatrix}$

$$\begin{aligned}
 8. A &= \begin{bmatrix} 0 & i & -i \\ -i & 0 & i \\ i & -i & 0 \end{bmatrix} \\
 9. A &= \begin{bmatrix} 1 & 1 & -1 \\ 2 & -1 & 4 \\ 3 & -2 & 3 \end{bmatrix} \\
 10. A &= \begin{bmatrix} -1 & -2 & -1 \\ 6 & 12 & 6 \\ 5 & 10 & 5 \end{bmatrix} \\
 11. A &= \begin{bmatrix} 1 & 1 & 1 \\ b+c & c+a & a+b \\ bc & ca & ab \end{bmatrix}
 \end{aligned}$$

6.4. Row rank and Column rank of a matrix: Let $A = [a_{ij}]$ be a $m \times n$ matrix over a field F . Then the m rows $A_i = (a_{i1}, a_{i2}, a_{i3}, \dots, a_{in})$, where $i = 1, 2, 3, \dots, m$ are elements of F^n , and the generated subspace of F^n is called the row space of the matrix A .

Similarly the n columns $A^j = \begin{bmatrix} x_{1j} \\ \vdots \\ x_{mj} \end{bmatrix}$ where $j = 1, 2, 3, \dots, n$ are elements of F^m ,

and the generated subspace of F^m is called the column space of the matrix A

The dimension of the row space is called the **row rank** of the matrix A , which is the maximum number of linearly independent rows of A .

The dimension of the column space is called the **column rank** of the matrix A , which is the maximum number of linearly independent columns of A .

Theorem: Let $T: V \rightarrow U$ be a linear transformation. Then $\text{Rank}(T)$ is equal to the column rank of the matrix of T .

Theorem: The row rank and the column rank of any matrix A are equal and this equal to the determinantal rank of A .

Note: $\text{rank } T = \text{rank } T^t$; $\text{rank } A = \text{rank } A^t$.

Theorem: The rank of the product of two matrices can not exceed the rank of either matrix.

Proof: Let A be an $m \times n$ and B be an $n \times p$ matrix

Let $T_1: V^p \rightarrow U^n$ and $T_2: U^n \rightarrow Z^m$ be linear transformations such that matrices of T_1 and T_2 relative to the suitable bases be B and A respectively.

The matrix of $(T_2T_1) = AB$.

$rank(T_2T_1) = rank(AB); rank(T_2) = rankA$ and $rank(T_1) = rankB$

Since, we know that $rank(T_2T_1)$ can not exceed $rank(T_2)$ or $rank(T_1)$.

Hence, $rank(AB)$ can not exceed $rank(A)$ or $rank(B)$.

Thus $rank(AB) \leq \text{Min}(rankA, rankB)$

Remark: 1. The rank of an n –square matrix A is n if and only if it is non-singular.

2. If the n – square matix A is non- singular and B is any n – square matrix then $rank(AB) = rankB = rank(BA)$

3. If A is non- singular then the linear map T associated with A is also non-singular.

4. The ranks of equivalent matrices are equal.

5. The rank of a zero matrix of any order is zero.

Example: Let A and B be two matrices where $A + B$ and AB both are defined. Let $rankA = m, rankB = n$ where $m < n$ then find the respective upper limits for $rank(A + B)$ and $rank(AB)$.

Proof: Since, $rank(AB) \leq \text{Min}(rankA, rankB) = m$

So, upper limit of $rank(AB) = m$.

Let $T_1: V \rightarrow W$ and $T_2: V \rightarrow W$ be linear transformations such that matrices of T_1 and T_2 relative to the suitable bases be A and B respectively.

$rank(T_2) = rankB = n$ and $rank(T_1) = rankA = m$

Matrix of $(T_2+T_1) = A + B$, so, $rank(T_2+T_1) = rank(A + B)$

Now, $T_1(V) = W_1 \subseteq W, T_2(V) = W_2 \subseteq W$

We know that $\dim(W_1 + W_2) = \dim(W_1) + \dim(W_2) - \dim(W_1 \cap W_2)$

Hence, $\dim(T_2+T_1)V \leq \dim T_1(V) + \dim T_2(V)$

Or, $rank(T_2+T_1) \leq rank(T_1) + rank(T_2)$

Hence, $rank(A + B) \leq rank(A) + rank(B)$

Thus, $rank(A + B) \leq m + n$.

So, upper limit of $rank(A + B) = m + n$.

Remark: If an $m \times n$ matrix $A = [a_{ij}]$ over a field F has rank r ,

$$\text{Then } A = [a_{ij}] \cong \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$$

Where sub- matrix I_r is the identity matrix of order $r \times r$ and other sub-matrices are null matrices.

6.5. Elementary matrices: The matrix obtained when any one of the three elementary operations (row or column) is applied on the identity matrix, is called an elementary matrix. Thus we have three types of elementary matrices corresponding to the three elementary operations on a matrix.

Example: Determine the elementary matrices corresponding to the following

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

Solution: (i). $R_1 \leftrightarrow R_2 \sim E_1 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

(ii). $R_3 \rightarrow -7R_3 \sim E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -7 \end{bmatrix}$

(iii). $R_2 \rightarrow R_2 - 3R_1 \sim E_1 = \begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

Remark:1. Each elementary row (respectively column) operation on an $m \times n$ matrix A can be effected by pre- multiplying (respectively post multiplying) A by the corresponding elementary matrix.

2. The elementary matrices are non- singular.

6.6. Elementary Row and column operations of a matrix:

Let $A = [a_{ij}]$ be any $m \times n$ matrix over any of the number sets $\mathbb{Q}$, $\mathbb{R}$ or $\mathbb{C}$. The following transformation applied on the matrix A are called elementary row operations:

1. Interchange of i th and j th rows of A . We denote it by R_{ij}

$$\begin{array}{c} i \\ j \end{array} \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{i1} & \dots & a_{in} \\ \vdots & & \vdots \\ a_{j1} & \dots & a_{jn} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} \xrightarrow{R_{ij}} \begin{array}{c} i \\ j \end{array} \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{j1} & \dots & a_{jn} \\ \vdots & & \vdots \\ a_{i1} & \dots & a_{in} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix}$$

2. Multiplying any row (say i^{th}) by a non zero scalar α i.e. $0 \neq \alpha \in S (= \mathbb{Q}, \mathbb{R} \text{ or } \mathbb{C})$ we denote this by αR_i .

$$\begin{array}{c} i \\ j \end{array} \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{i1} & \dots & a_{in} \\ \vdots & & \vdots \\ a_{j1} & \dots & a_{jn} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} \xrightarrow{\alpha R_i} \begin{array}{c} i \\ j \end{array} \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ \alpha a_{i1} & \dots & \alpha a_{in} \\ \vdots & & \vdots \\ a_{j1} & \dots & a_{jn} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix}$$

3. Adding β times j^{th} row to the i^{th} row : We denote it by $R_i + \beta R_j$ (where $\beta \in S$)

$$\begin{array}{c} i \\ j \end{array} \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{i1} & \dots & a_{in} \\ \vdots & & \vdots \\ a_{j1} & \dots & a_{jn} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} \xrightarrow{R_i + \beta R_j} \begin{array}{c} i \\ j \end{array} \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{i1} + \beta a_{j1} & \dots & a_{in} + \beta a_{jn} \\ \vdots & & \vdots \\ a_{j1} & \dots & a_{jn} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix}$$

It should be noticed that each of the elementary row operations is reversible and the reverse operation is also an elementary row operation of the same type. Thus reverse of R_{ij} is $R_j = R_{ij}$ itself. Reverse of αR_i is $(1/\alpha) R_i$

$\because \alpha \neq 0 \Rightarrow \frac{1}{\alpha} \in S \text{ \& } \frac{1}{\alpha} \neq 0$. Reverse of $R_i + \beta R_j$ is $R_i + (-\beta) R_j$.

Also note that the II and the III type of operations only affects the i th row and I type operation affects the i th and j th rows only while all the other rows remain unchanged.

Example 1:
$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \xrightarrow{R_{13}} \begin{bmatrix} 7 & 8 & 9 \\ 4 & 5 & 6 \\ 1 & 2 & 3 \end{bmatrix}$$

$$2. \quad \begin{bmatrix} 1 & 4 & 5 \\ -1 & 6 & 4 \\ 2 & -3 & 6 \end{bmatrix} \xrightarrow{\frac{1}{5}R_2} \begin{bmatrix} 1 & 4 & 5 \\ -\frac{1}{5} & \frac{6}{5} & \frac{4}{5} \\ 2 & -3 & 6 \end{bmatrix}$$

$$3. \quad \begin{bmatrix} 2 & 1 & -3 \\ 4 & -3 & 6 \\ 7 & 2 & 5 \end{bmatrix} \xrightarrow{R_2 + (-2)R_1} \begin{bmatrix} 2 & 1 & -3 \\ 0 & -5 & 12 \\ 7 & 2 & 5 \end{bmatrix}$$

The reverse operations of these are respectively R_{13} , $5R_2$ and R_2+2R_1 and upon application on the matrices on the right, bring them back to the original ones.

In a similar manner we can define elementary column operations but they are not so useful, as the row operations are mainly used for the purpose of solving arbitrary systems of linear algebraic equations.

Although elementary row operations can be applied on any matrix A (not necessarily a square matrix only) but when applied on square matrices, the first type multiplies $\det(A)$ by -1 , the second type multiplies it by $\alpha \neq 0$ and the third type does not change the value of the determinant. Since invertible matrices are those square matrices which have nonzero determinant, it is now clear that elementary operations keep invertible matrices invertible and keep non invertible matrices noninvertible.

To obtain a non- singular matrix B of order $m \times n$ from an $m \times n$ matrix A by a finite number of elementary row and column operations on A then $B \cong A$.

Proof: Let $E_1, E_2, E_3, \dots, E_r$ be the elementary matrices corresponding to the row operations on A and $E'_1, E'_2, E'_3, \dots, E'_s$ be the elementary matrices corresponding to the column operations on A .

Hence, we have $B = (E_r \dots E_3 E_2 E_1) A (E'_1 E'_2 E'_3 \dots E'_s) = SAT$

Where, $S = E_r \dots E_3 E_2 E_1$ and $T = E'_1 E'_2 E'_3 \dots E'_s$ and so S and T are non- singular matrices being products of elementary matrices which are non-singular.

Now, $\text{rank} B = \text{rank}(SAT) = \text{rank} A$

Hence, $B \cong A$.

6.7. Normal form of a matrix: The matrix $A = [a_{ij}] \cong \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$ Where sub- matrix I_r is the identity matrix of order $r \times r$ and other sub- matrices are null matrices. The matrix $\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$ to be equivalent to the matrix $[a_{ij}]$, is called the normal form of the matrix.

Note:1. Two matrices $A = [a_{ij}]$ and $B = [b_{ij}]$ having the same rank r are equivalent, since each of them is equivalent to the matrix $\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$.

2. To find the rank of a matrix A , it can be reduced to its normal form $\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$ by applying elementary row and column operations. Then the rank of $A = r$ (order of the identity sub-matrix I_r in the normal form.)

Example: Reduce the matrix in the normal form and hence find the rank of

$$\text{the matrix } A = \begin{bmatrix} 1 & 2 & -1 & 2 & 1 \\ 2 & 4 & 1 & -2 & 3 \\ 3 & 6 & 2 & -6 & 5 \end{bmatrix}$$

Solution: We apply row operations $R_2 \rightarrow R_2 - 2R_1, R_3 \rightarrow R_3 - 3R_1$, we get

$$A \sim \begin{bmatrix} 1 & 2 & -1 & 2 & 1 \\ 0 & 0 & 3 & -6 & 1 \\ 0 & 0 & 5 & -12 & 2 \end{bmatrix}$$

By applying column operations $C_2 \rightarrow C_2 - 2C_1, C_3 \rightarrow C_3 + C_1, C_4 \rightarrow C_4 - 2C_1$

$$C_5 \rightarrow C_5 - C_1 \sim \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & -6 & 1 \\ 0 & 0 & 5 & -12 & 2 \end{bmatrix}$$

$$\text{By applying } C_2 \leftrightarrow C_5 \sim \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 3 & -6 & 0 \\ 0 & 2 & 5 & -12 & 0 \end{bmatrix}$$

$$\text{By applying } R_3 \rightarrow R_3 - 2R_2 \sim \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 3 & -6 & 0 \\ 0 & 0 & -1 & 0 & 0 \end{bmatrix}$$

$$\text{By applying } C_3 \rightarrow C_3 - 3C_2, C_4 \rightarrow C_4 + 6C_2 \sim \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \end{bmatrix}$$

Now applying $R_3 \rightarrow (-1)R_3 \sim \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} I_3 & 0 \\ 0 & 0 \end{bmatrix}$

Hence rank of the matrix A is 3

Note: To reduce A into a normal form, we apply row and column operations on A so as to make the diagonal elements of A is equal to I as far as possible after which the diagonal elements if exists are all equal to zeros, and all non-diagonal elements are made zeros.

6.8. Echelon form of a matrix: A matrix is said to in Echelon form if:

- (i). All the non- zero rows if any precede the zero rows.
- (ii). The number of zeros preceding the first non- zero entry of a row, increases row by row until only zero rows remain.
- (iii). The first non- zero entry in arrow is 1.

The matrix is said to be in reduced Echelon form if further the first non- zero entry 1 of the i^{th} row in the j^{th} column, then all other elements if the j^{th} column are zeros.

Theorem: The rank of a matrix in Echelon form is equal to the number of non- zero rows of the matrix.

Example: $\begin{bmatrix} 1 & 2 & -1 & 2 & 1 \\ 0 & 0 & 3 & -6 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$ is in the Echelon form.

Check your progress

Reduce the following matrices to their normal forms and hence find their ranks

1. $\sim \begin{bmatrix} 2 & 3 & -1 & -1 & 1 \\ 1 & -1 & -2 & -4 & 1 \\ 3 & 1 & 3 & -2 & 2 \end{bmatrix}$

2. $\begin{bmatrix} 2 & 6 & -2 \\ 5 & 1 & -4 \\ -2 & -4 & -3 \end{bmatrix}$

Compute the ranks of the following matrices:

1. $\begin{bmatrix} 2 & 1 & 4 \\ 3 & 6 & 2 \\ 0 & 0 & 1 \end{bmatrix}$

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

3. $\begin{bmatrix} -1 & -2 & -1 \\ 6 & 12 & 6 \\ 5 & 10 & 5 \end{bmatrix}$

6.9. Inverse of a non- singular matrix: If a $n -$ square matrix A is reduced to the identity matrix I_n by applying a finite number of elementary row operations, then the matrix A is non- singular or invertible. Then A^{-1} can be obtained by applying the same sequence of elementary row operations on I_n which were applied on A to reduce it to I_n .

Proof: Since, A is reduced to I_n by applying a finite number of elementary row operations applied on A , So, $A \cong I_n$.

Hence, the rank $A = \text{rank} I_n = n$.

Thus the rank of $n -$ square matrix A is n .

So, A must be non- singular and so invertible. Hence, A^{-1} exists.

Let $E_1, E_2, E_3, \dots, E_r$ be the elementary matrices, corresponding to the elementary row operations applied on A to reduce it to I_n .

Then $(E_1, E_2, E_3, \dots, E_r)A = I_n$.

Therefore, $(E_1, E_2, E_3, \dots, E_r)AA^{-1} = I_n A^{-1} = A^{-1}$

$$(E_1, E_2, E_3, \dots, E_r)I_n = A^{-1}$$

Remark: This theorem gives us the to find the inverse of a non- singular matrix.. Adjoin the matrix A with I_n by applying elementary row operations on the adjoined matrix to reduce its sub- matrix A to I_n , then the reduced form of its sub- matrix I_n will be the inverse of A .

Example: Compute the inverse of the matrix $\begin{bmatrix} 2 & 1 & 2 \\ 1 & 0 & 5 \\ 2 & 5 & 11 \end{bmatrix}$

Solution: adjoining A with I_3 , we get,

$$\begin{aligned}
B &= \begin{bmatrix} 2 & 1 & 2 \\ 1 & 0 & 5 \\ 2 & 5 & 11 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\
&\sim \begin{bmatrix} 1 & 0 & 5 \\ 2 & 1 & 2 \\ 2 & 5 & 11 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{ by } R_1 \leftrightarrow R_2 \\
&\sim \begin{bmatrix} 1 & 0 & 5 \\ 0 & 1 & -8 \\ 0 & 5 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 1 & -2 & 0 \\ 0 & -2 & 1 \end{bmatrix} \text{ by } R_2 \rightarrow R_2 - 2R_1, R_3 \\
&\quad \rightarrow R_3 - 2R_2 \\
&\sim \begin{bmatrix} 1 & 0 & 5 \\ 0 & 1 & -8 \\ 0 & 0 & 41 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 1 & -2 & 0 \\ -5 & 8 & 1 \end{bmatrix} \text{ by } R_3 \rightarrow R_3 - 5R_2 \\
&\sim \begin{bmatrix} 1 & 0 & 5 \\ 0 & 1 & -8 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 1 & -2 & 0 \\ -\frac{5}{41} & \frac{8}{41} & \frac{1}{41} \end{bmatrix} \text{ by } R_3 \rightarrow \frac{1}{41}R_3 \\
&\sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{25}{41} & \frac{1}{41} & \frac{-5}{41} \\ \frac{1}{41} & -\frac{18}{41} & \frac{8}{41} \\ -\frac{5}{41} & \frac{8}{41} & \frac{1}{41} \end{bmatrix} \text{ by } R_1 \rightarrow R_1 - 5R_3, R_2 \rightarrow \\
&\quad R_2 + 8R_3
\end{aligned}$$

$$\text{Hence, } A^{-1} = \begin{bmatrix} \frac{25}{41} & \frac{1}{41} & \frac{-5}{41} \\ \frac{1}{41} & -\frac{18}{41} & \frac{8}{41} \\ -\frac{5}{41} & \frac{8}{41} & \frac{1}{41} \end{bmatrix} = \frac{1}{41} \begin{bmatrix} 25 & 1 & -5 \\ 1 & -18 & 8 \\ -5 & 8 & 1 \end{bmatrix}$$

$$\text{and } I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \text{ We can verified that } AA^{-1} = I_3$$

Check your progress

Compute the inverse of the following matrices:

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

$$5. \begin{bmatrix} 1 & 2 & -2 \\ -1 & 3 & 0 \\ 0 & -2 & 2 \end{bmatrix}$$

$$6. \begin{bmatrix} 2 & 4 & 3 \\ 0 & 1 & 1 \\ 2 & 2 & -1 \end{bmatrix}$$

$$7. \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 0 \\ 0 & 1 & 2 \end{bmatrix}$$

$$8. \begin{bmatrix} i & -1 & 2i \\ 2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$$

6.10. Determinantal rank of a matrix: Let $A = [a_{ij}]$ be a matrix of order $m \times n$. The determinantal rank of a square sub-matrix of A of order r obtained by deleting a certain number of rows and columns, is called a minor of the matrix A of order r .

The matrix A is said to have determinantal rank ρ if

- (i). Every minor of A of order $(\rho + 1)$ has the value 0
- (ii). There exists a minor of A of order ρ whose value is not zero.

Example: The determinantal rank of the matrix $A = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 4 & 5 \\ 4 & 5 & 6 \end{bmatrix}$

We have seen that minors of order 2 of A are $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$, $\begin{bmatrix} 1 & 2 \\ 4 & 5 \end{bmatrix}$, $\begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix}$ etc

Only minor of A of order 3 is $|A| = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 4 & 5 \\ 4 & 5 & 6 \end{bmatrix}$

We can verify that $|A| = 0$ and there exists a minor of order 2, $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = -2 \neq 0$. Hence, the determinantal rank of A is 2.

Theorem: Let $A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ a_{r1} & \dots & a_{rn} \end{bmatrix}$ be an $r \times n$ matrix over a field F .

Then the r row vectors $A_i = (a_{i1}, a_{i2}, \dots, a_{in})$ where $i = 1, 2, 3, \dots, r$ are linearly independent if and only if determinantal rank ρ of A is equal to r .

Theorem: Let $A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ a_{m1} & \dots & a_{mn} \end{bmatrix}$ be an $m \times n$ matrix over a field F .

Then its row rank will be equal to its determinantal rank.

6.11. Non homogeneous and homogeneous linear equations:

We now indicate an application of determinants for solving a system of n linear algebraic equations in n unknowns, in case the coefficient matrix is invertible i.e. when the associated determinant has a nonzero value.

Suppose the given system is

$$a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1$$

$$a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2$$

.....

.....

$$a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n = b_n$$

where $a_{11}, \dots, a_{nn}$ are n^2 given constants and so are $b_1, \dots, b_n$, which belong to any of the number sets $\mathbb{Q}, \mathbb{R}$ or $\mathbb{C}$. We have to take a set in which quotient of any member by any nonzero member should also belong to apart from sum, difference of product of any two members if we want the solutions for the unknowns $x_1, \dots, x_n$ in that number set. The above system can be written

in matrix form as $A \underline{x} = \underline{b}$ Where $A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ a_{n1} & \dots & a_{nn} \end{bmatrix}$ is the coefficient

matrix of the system. $\underline{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ is the vector whose components are the n

unknowns and $\underline{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$ is the given right hand side vector and the l.h.s is the

usual matrix product of A and $\underline{x}$. We suppose that A is invertible i.e. $\det(A) = \Delta \neq 0$. Then we know that A^{-1} exists and

$$A^{-1} = \frac{\text{Adj}A}{\Delta} \text{ satisfies } AA^{-1} = I_n = A^{-1}A.$$

Multiplying the above matrix equation $A\mathbf{x}=\mathbf{b}$ by A^{-1} on the left, we get

$$\mathbf{x} = I_n \mathbf{x} = (A^{-1}A) \mathbf{x} = A^{-1} (A \mathbf{x}) = A^{-1}\mathbf{b}. \text{ i.e. } \mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = A^{-1}\mathbf{b} = \frac{1}{\Delta} (\text{Adj}A)\mathbf{b}.$$

Both sides here are $n \times 1$ matrices. Equating the i^{th} components, we get

$$x_i = \frac{1}{\Delta} [(\text{Adj}A)\mathbf{b}]_i = \frac{1}{\Delta} (\tilde{R}_i \cdot \mathbf{b}) \text{ where } \tilde{R}_i = i^{\text{th}} \text{ row of Adj}A$$

$$= i^{\text{th}} \text{ column of cof } A = \begin{bmatrix} A_{1i} \\ A_{2i} \\ \vdots \\ A_{ni} \end{bmatrix}$$

$$\therefore x_i = \frac{(A_{1i}b_1 + A_{2i}b_2 + \dots + A_{ni}b_n)}{\Delta}$$

$$= \frac{1}{\Delta} \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1,i-1} & b_1 & a_{1,i+1} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{n,i-1} & b_n & a_{n,i+1} & \dots & a_{nn} \end{vmatrix}$$

$$(\text{Expanded using the } i^{\text{th}} \text{ column}). = \frac{\Delta_i}{\Delta}$$

where Δ_i is the determinant obtained from $\det A = \Delta$ by replacing

$$\text{its } i^{\text{th}} \text{ column by the r.h.s. column } \mathbf{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$$

$\therefore$ The unique solution to the given system

$$A\mathbf{x}=\mathbf{b} \text{ is } x_1 = \frac{\Delta_1}{\Delta}, x_2 = \frac{\Delta_2}{\Delta}, \dots, x_n = \frac{\Delta_n}{\Delta}$$

whenever the coefficient determinant Δ has nonzero value.

The above result is known as Cramer's rule.

Example: Solve the system

$$x + 3z = 1$$

$$5x + 6y + 9z = 2$$

$$4x + 5y + 6z = 3$$

Solution: We had earlier seen that

$$\Delta = \begin{vmatrix} 1 & 0 & 3 \\ 5 & 6 & 9 \\ 3 & 5 & 6 \end{vmatrix} = -\frac{1}{6} \neq 0. \text{ Here } x = \frac{\Delta_1}{\Delta}$$

$$= \frac{\begin{vmatrix} 1 & 0 & 3 \\ 2 & 6 & 9 \\ 3 & 5 & 6 \end{vmatrix}}{\begin{vmatrix} 1 & 0 & 3 \\ 5 & 6 & 9 \\ 3 & 5 & 6 \end{vmatrix}} = -\frac{1}{6} [36 - 45] + 3(10 - 18) \\ = \frac{+33}{6} = \frac{11}{2}.$$

$$y = \frac{\Delta_2}{\Delta} = \frac{\begin{vmatrix} 1 & 1 & 3 \\ 5 & 2 & 9 \\ 4 & 3 & 6 \end{vmatrix}}{\begin{vmatrix} 1 & 0 & 3 \\ 5 & 6 & 9 \\ 3 & 5 & 6 \end{vmatrix}} = -\frac{1}{6} [(12 - 27) - (30 - 36) + 3(15 - 8)] = -\frac{12}{6} = -2$$

$$z = \frac{\Delta_3}{\Delta} = \frac{\begin{vmatrix} 1 & 0 & 1 \\ 5 & 6 & 2 \\ 4 & 5 & 3 \end{vmatrix}}{\begin{vmatrix} 1 & 0 & 3 \\ 5 & 6 & 9 \\ 3 & 5 & 6 \end{vmatrix}} = -\frac{1}{6} [(18 - 10) + (25 - 24)] = -\frac{9}{6} = -\frac{3}{2}$$

$\therefore$ The system has a unique solution $x = \frac{11}{2}$, $y = -2$, $z = -\frac{3}{2}$

Gaussian Elimination

It is the good old method of solving systems of linear equations by considering linear combinations of given equations in order to eliminate the involved unknowns one by one, until only one of them remains, and having obtained the value of it, go back on the steps and find the values of the other unknowns. Sometimes we find that the above process leads us to false statements such as $0=1$ and conclude that the system has no solutions and call the given system inconsistent. On the other hand, sometimes some of the unknowns cannot be evaluated (called free unknowns) and the other ones can be evaluated in terms of them and these values satisfy all the equations for each value assigned to the free unknowns and consequently the system possesses infinitely many solutions. The steps involved in Gaussian elimination procedure roughly correspond to those of elementary row operations applied on the coefficient matrix specifically to eliminate the unknowns, one by one.

Fact 1: Any solution set of a given linear equations in unknowns $x_1, x_2, \dots, x_n$ (say) also satisfies every linear combination (i.e. sum of constant multiplies of) of the given the given equations.

Fact 2: If the given linear system of equations leads to another such system by a sequence of elementary row operations on the coefficient matrix of the system, then each such operation being reversible, every solution of the new system satisfies the original system and thus the two systems have the same set of solution.

We tackle the system considered earlier, by the present method. Our aim at first is to transform the given system to upper triangular form.

If the system of equations is written as the matrix equations $A\mathbf{x} = \mathbf{b}$, then the matrix $A^+ = [A; \mathbf{b}]$ obtained from A by including the column vector $\mathbf{b}$ as the last column is called the augmented matrix for the given system. It is an $m \times (n + 1)$ matrix clearly if A is of order $m \times n$. We apply the row operations on A^+ rather than on A . The last column does not mix up with the rest under a row operation and the new augmented matrix quickly gives a new system having the same set of solutions as the given one. We choose our elementary row operations in such a way that the new system is easy to solve.

Example: $x + 3z = 1$

$$5x + 6y + 9z = 2$$

$$4x + 5y + 6z = 3$$

$$\text{Here } A^+ = \begin{bmatrix} 1 & 0 & 3 & : & 1 \\ 5 & 6 & 9 & : & 2 \\ 4 & 5 & 6 & : & 3 \end{bmatrix}$$

Apply $R_2 - 5R_1$, $R_3 - 4R_1$ to make the first entries of the II and III column to 0. This corresponds to removal of the unknown x from the last two equations. We get

$$\begin{bmatrix} 1 & 0 & 3 & : & 1 \\ 0 & 6 & -6 & : & -3 \\ 0 & 5 & -6 & : & -1 \end{bmatrix} \xrightarrow{\frac{1}{6}R_2} \begin{bmatrix} 1 & 0 & 3 & : & 1 \\ 0 & 1 & -1 & : & -\frac{1}{2} \\ 0 & 5 & -6 & : & -1 \end{bmatrix} \xrightarrow{R_3 - 5R_2} \begin{bmatrix} 1 & 0 & 3 & : & 1 \\ 0 & 1 & -1 & : & -\frac{1}{2} \\ 0 & 0 & -1 & : & \frac{3}{2} \end{bmatrix}$$

$$\xrightarrow{(-1)R_3} \begin{bmatrix} 1 & 0 & 3 & : & 1 \\ 0 & 1 & -1 & : & -\frac{1}{2} \\ 0 & 0 & 1 & : & -\frac{3}{2} \end{bmatrix}$$

This corresponds to the system

$$x + 3z = 1,$$

$$y - z = -\frac{1}{2},$$

$$z = -\frac{3}{2}$$

Putting the value of z in the II equation,

$$\text{we get, } y = z - \frac{1}{2} = -\frac{3}{2} - \frac{1}{2} = -2$$

$$\text{Putting the value of } z \text{ in (1) } x = 1 - 3z = 1 - 3\left(-\frac{3}{2}\right) = \frac{11}{2}$$

$$\therefore x = \frac{11}{2}, y = -2, z = -\frac{3}{2} \text{ (as obtained earlier)}$$

For a general case

(i). We may make the left corner entry = $a'_{11} \neq 0$

by suitable row interchange if it is 0 initially

(ii). Make it 1 by division by this nonzero entry.

(iii). Make the first column entries of other rows 0 by adding

or subtracting suitable multiples of the first row.

(iv). Repeat the procedure with the leftmost entry of the next equation along with the subsequent equations. The next example illustrates other possible type of outcomes.

Example: $4x + 6y + 8z = 11$

$$5x + 7y + 2z = 13$$

$$6x + 9y + 12z = 15$$

Solution: Here $A^+ = \begin{bmatrix} 4 & 6 & 8 & \vdots & 11 \\ 5 & 7 & 2 & \vdots & 13 \\ 6 & 9 & 12 & \vdots & 15 \end{bmatrix}$

Applying the stated procedure

$$A^- \xrightarrow{\frac{1}{4}R_1} \begin{bmatrix} 1 & \frac{3}{2} & 2 & \frac{11}{4} \\ 5 & 7 & 2 & 13 \\ 6 & 9 & 12 & 15 \end{bmatrix} \xrightarrow{R_2 - 3R_1, R_3 - 6R_1} \begin{bmatrix} 1 & \frac{3}{2} & 2 & \frac{11}{4} \\ 0 & -\frac{1}{2} & -8 & \frac{5}{4} \\ 0 & 0 & 0 & -\frac{3}{2} \end{bmatrix}$$

The last row corresponds to the equation $0 = 0x + 0y + 0z = -\frac{3}{2}$

which is clearly not possible. Therefore the system has no solution.

Example: $y + 3z + 5w = 8,$

$2x + 3y + z + 7w = 10,$

$3x - 6y + 4z + 2w = 2$

Here $A^+ = \begin{bmatrix} 0 & 1 & 3 & 5 & : & 8 \\ 2 & 3 & 1 & 7 & : & 10 \\ 3 & -6 & 4 & 2 & : & 2 \end{bmatrix}$

$$\begin{aligned} A^+ &\xrightarrow{R_{12}} \begin{bmatrix} 2 & 3 & 1 & 7 & 10 \\ 0 & 1 & 3 & 5 & 8 \\ 3 & -6 & 4 & 2 & 2 \end{bmatrix} \xrightarrow{\frac{1}{2}R_1} \begin{bmatrix} 1 & \frac{3}{2} & \frac{1}{2} & \frac{7}{2} & 5 \\ 0 & 1 & 3 & 5 & 8 \\ 3 & -6 & 4 & 2 & 2 \end{bmatrix} \xrightarrow{R_3 - 3R_1} \begin{bmatrix} 1 & \frac{3}{2} & \frac{1}{2} & \frac{7}{2} & 5 \\ 0 & 1 & 3 & 5 & 8 \\ 0 & -\frac{21}{2} & \frac{5}{2} & -\frac{17}{2} & -13 \end{bmatrix} \\ &\xrightarrow{R_3 + \frac{21}{2}R_2} \begin{bmatrix} 1 & \frac{3}{2} & \frac{1}{2} & \frac{7}{2} & 5 \\ 0 & 1 & 3 & 5 & 8 \\ 0 & 0 & 34 & 44 & 75 \end{bmatrix} \xrightarrow{\frac{1}{34}R_3} \begin{bmatrix} 1 & \frac{3}{2} & \frac{1}{2} & \frac{7}{2} & 5 \\ 0 & 1 & 3 & 5 & 8 \\ 0 & 0 & 1 & \frac{22}{17} & \frac{75}{34} \end{bmatrix} \end{aligned}$$

This corresponds to the system of equations

$$x + \frac{3}{2}y + \frac{1}{2}z + \frac{7}{2}w = 5,$$

$$y + 3z + 5w = 8, \quad z + \frac{22}{15}w = \frac{75}{34}$$

$$\text{Getting } z = \frac{75}{34} - \frac{22}{17}w \quad (\forall w).$$

Putting back in the II equation above:

$$\begin{aligned}y &= 8 - 3z - 5w = 8 - 3\left(\frac{75}{34} - \frac{22}{17}w\right) \\&= \left(8 - \frac{225}{34}\right) + \frac{66}{17}w = \frac{47}{34} + \frac{66}{17}w\end{aligned}$$

Putting the values of x and y in the I equation

$$\begin{aligned}x &= 5 - \frac{3}{2}y - \frac{1}{2}z - \frac{7}{2}w \\&= 5 - \frac{3}{2}\left(\frac{47}{34} + \frac{66}{17}w\right) - \frac{1}{2}\left(\frac{75}{34} - \frac{22}{17}w\right) - \frac{7}{2}w \\&= \frac{31}{17} - \frac{295}{34}w\end{aligned}$$

Thus the complete solution is $x = \frac{31}{17} - \frac{295}{34}w$,

$$y = \frac{47}{34} + \frac{66}{17}w, \quad z = \frac{75}{34} - \frac{22}{17}w$$

and $w = w$ for all values of w and hence the system has infinitely many solutions.

Check your progress

1. For the following systems of linear equations, decide whether it has a solution, and then for the solvable system, find all the solutions by Gaussian eliminations.

(i) $x + 3y - 2z = 7$, $5x - 2y + 4z = 12$, $-x + 16y - 16z = 6$

(ii) $x - y + 2z - 3w = 5$, $2x - 3y - 5z + 7w = 9$, $5x + y - w = 8$

(iii) $x + y - z = 3$, $-3x + 2y + 3z = 6$, $2x + 2y - 5z = -7$

2. Apply suitable row operations to transform the matrix A to the matrix I_3 .

(i) $A = \begin{bmatrix} 6 & 18 \\ 4 & 12 \end{bmatrix}$

$$(ii) A = \begin{bmatrix} 1 & 1 & -1 \\ -3 & 2 & 3 \\ -2 & -2 & 5 \end{bmatrix}$$

$$(iii) A = \begin{bmatrix} 2 & -1 & 3 \\ 6 & 0 & 1 \\ 18 & -3 & 11 \end{bmatrix}$$

6.12. Summary

After the study the chapter we will be able to find the Rank of a Matrix, to find Row rank and Column rank of a matrix, elementary matrices, elementary Row and column operations of a matrix, to find Normal form of a matrix, to find Echelon form of matrix, to find Inverse of a non- singular matrix, to find Determinantal rank of a matrix, identify the linear systems which can be solved by using Cramer's rule, and apply the rule to solve them. Also to find Gaussian Elimination to solve systems of linear equations.

6.13 Terminal Questions

- Find rank of matrix $A = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \\ 2 & 3 & 1 \end{bmatrix}$, Answer $\rho(A) = 3$
- Find rank of matrix $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, Answer $\rho(A) = 3$
- Find rank of matrix $A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 6 \end{bmatrix}$, Answer $\rho(A) = 3$
- Find rank of matrix $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 7 & 8 & 10 \end{bmatrix}$, Answer $\rho(A) = 3$
- Find rank of matrix $A = \begin{bmatrix} 0 & 2 & 3 \\ 0 & 4 & 6 \\ 0 & 6 & 9 \end{bmatrix}$, Answer $\rho(A) = 1$
- Solve $x + y + z = 1, x + 2y + z = 2, x + y + 2z = 0$, Answer $x = 1, y = 1, z = -1$
- Show that the system of equations $2x + 6y = -11, 6x + 20y - 6z = -3, 6y - 18z = -1$ are not consistent.
- Test the consistency of the following equation $x - y + z = 5, 2x + y - z = -2, 3x - y = -7$
- Using elementary row operations find the inverse of $\begin{bmatrix} 1 & 2 & 5 \\ 2 & 3 & 1 \\ -1 & 1 & 1 \end{bmatrix}$
- Find the inverse of the matrix $\begin{bmatrix} 1 & 3 & 3 \\ 1 & 4 & 3 \\ 1 & 3 & 4 \end{bmatrix}$

Unit 7

Determinants

Structure

7.1 Introduction

7.2 Objectives

7.3 Determinant of a square matrix

7.4 cofactor of an element of a determinant

7.5 Properties of determinant

7.6 Minor and Complementary minor of determinant

7.7 Evaluation of a determinant

7.8 Laplace expansion of a determinant

7.9 Product of two determinants

7.10 Summary

7.11 Terminal Questions

7.1 INTRODUCTION

This unit introduces the concept of determinants and discusses their important properties. Also these are used to calculate inverses of square matrices wherever they exist.

7.2 Objectives

After studying this unit you should be able to:

- Determinant of a square matrix
- Determine cofactor of an element of a determinant
- Determine Properties of determinant
- Determine Minor and Complementary minor of determinant
- Evaluation of a determinant
- Determine Laplace expansion of a determinant
- Determine product of two determinants

Now let us look at a convenient way of representing a linear system.

Suppose that we are given a square matrix A with entries from a set S in which we can add, subtract and multiply with these operations satisfying usual properties. Recall that a permutation of a finite set say $\{1, 2, \dots, n\}$ is any rearrangement of the elements, $1, 2, \dots, n$ say $(a_1, a_2, \dots, a_n)$. We may view it as a bijective function $p : \{1, 2, \dots, n\} \rightarrow \{1, 2, \dots, n\}$ where $p(j) = a_j$ for $j = 1, \dots, n$. We also know that there are exactly $n!$ permutations of any set containing n distinct elements.

The set of all permutations of $\{1, 2, 3, \dots, n\}$ is denoted by S_n . If we represent a permutation p by

$$p = \begin{pmatrix} 1 & 2 & \dots & n \\ p(1) & p(2) & \dots & p(n) \end{pmatrix},$$

$$\text{then } S_1 = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}, S_2 = \left\{ \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \right\} \text{ and}$$

$$S_3 = \left\{ \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \right\}$$

It is known that every permutation is the composition of such permutations (called transposition) which interchange two elements of $\{1, 2, \dots, n\}$ and leave all other elements unchanged. This composition may even be a singleton composition which is the case for permutation which is themselves

transpositions. For example the second element listed above in S_2 and the last three elements of S_3 above are transpositions (respectively denoted by (1, 2) and (2, 3), (1, 3), (1, 2)). Note that the element $\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$ is the result of interchange of 1 and 2 followed by that of 2 and 3.

$$(1 \ 2 \ 3) \xrightarrow{(1,2)} (2, \ 1, \ 3) \xrightarrow{(2,3)} (3, \ 1, \ 2) = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$$

This kind of representation of a permutation as a composition of transpositions is not unique nor is the number of transposition involved in such decomposition. But there is one aspect of such decomposition and that is the evenness or oddness of the number of transpositions appearing in the decomposition of a particular permutation. We call a permutation an even or an odd permutation according as it can be written as a composition of an even or an odd number of transpositions. It is also known that for $k \geq 2$, among all the $n!$ permutations of $\{1, 2, \dots, n\}$ exactly half are even and the other half odd. We can now state the definition and important properties of a determinant of a square matrix (mostly without proof).

7.3 Determinant of a square matrix:

Definition: Given a square matrix $A = [a_{ij}]$ of order $n \times n$ over a number set S , we call the following element of S , the determinant of A and denote it by $\det(A)$:

$$\det(A) = \sum_{p \in S_n} (-1)^p a_{1,p(1)} a_{2,p(2)} \cdots a_{n,p(n)},$$

$$\text{which is also equal to } \sum_{p \in S_n} (-1)^p a_{p(1),1} a_{p(2),2} \cdots a_{p(n),n}$$

Each of the above sums has $n!$ terms as S_n has $n!$ elements. $(-1)^p$ is taken as $+1$ or -1 according as p is an even or an odd permutation. Note that the determinant is a sum of signed (± 1) sum of products of entries of the square matrix A in such a way that each term is a product of entries coming from different rows and also from different columns because for $i \neq i'$, we have $p(i) \neq p(i')$.

Note1: Determinant $[a_{ij}]$ of the n – square matrix $[a_{ij}]$ is called a determinant of order n .

2. We see that $[a_{ij}]$ contains $n!$ products such that each product contains only one element from each row and each column.

3. A product containing a_{ij} does not contain any other element from the i^{th} row and j^{th} column.

Illustration:

1. $\text{Det} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = +a_{11}a_{22} - a_{12}a_{21}$ because the elements of S_2 are

p_1 and p_2 where $p_1(1)=1, p_1(2)=2, p_2(1)=2, p_2(2)=1$ and p_1, p_2 are respectively even and odd.

We denote the above value by $\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$

$$2. \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{1p_1(1)}a_{2p_1(2)}a_{3p_1(3)} + a_{1p_2(1)}a_{2p_2(2)}a_{3p_2(3)} \\ + a_{1p_3(1)}a_{2p_3(2)}a_{3p_3(3)} - a_{1p_4(1)}a_{2p_4(2)}a_{3p_4(3)} \\ - a_{1p_5(1)}a_{2p_5(2)}a_{3p_5(3)} - a_{1p_6(1)}a_{2p_6(2)}a_{3p_6(3)}$$

Where $p_1, p_2, \dots, p_6$ are the elements of S_3 listed in this order and first three there were even and the last three were odd which explain the signs taken here.

Putting the values of the images $p_i(j)$ for $1 \leq i \leq 6, 1 \leq j \leq 3$ we get

$$\det([a_{ij}]) = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{13}a_{22}a_{31} \\ - a_{12}a_{21}a_{33}.$$

7.4 Cofactor of an element of a determinant: The minor of a given entry a_{ij} of an $n \times n$ determinant $\det([a_{ij}])$ is defined as the $(n-1) \times (n-1)$ determinant obtained by removing the whole row and column in which the entry a_{ij} lies and is denoted by M_{ij} and $A_{ij} = (-1)^{i+j} M_{ij}$ is called the cofactor of the entry a_{ij} . Note that M_{ij} and A_{ij} are values of some determinants and are not matrices. Hence, we may explain $|A|$ in terms of elements of the i^{th} row, by collecting the terms containing $a_{i1}, a_{i2}, a_{i3}, \dots, a_{in}$, and summing them, we get

$$|a_{ij}| = a_{i1}C_{i1} + a_{i2}C_{i2} + a_{i3}C_{i3} + \dots + a_{in}C_{in} = \sum_{j=1}^n a_{ij}C_{ij}$$

Similarly we can expand the elements of the j^{th} column and their corresponding co-factors as $|a_{ij}| = a_{1j} C_{1j} + a_{2j} C_{2j} + a_{3j} C_{3j} + \dots + a_{nj} C_{nj} = \sum_{i=1}^n a_{ij} C_{ij}$

7.5 Properties of determinant

We shall denote the rows and column of the square matrix $A = [a_{ij}]$ by $R_1, R_2, \dots, R_n$ and $C_1, C_2, \dots, C_n$ respectively.

Property1. $\det A = \det A^T$. That is the value of a determinant does not change by writing the rows as columns and vice versa.

Property2. For each $1 \leq i \leq n$, $\det A =$

$$\sum_{j=1}^n a_{ij} A_{ij} = a_{i1} A_{i1} + a_{i2} A_{i2} + \dots + a_{in} A_{in}$$

$$= (-1)^i [-a_{i1} M_{i1} + a_{i2} M_{i2} \dots + (-1)^n a_{in} M_{in}]$$

Which given the usual expression for $\det A$ as a sum of determinants using the entries of the i th row. Similarly a determinant can be expanded using entries of any of its columns: For each $1 \leq j \leq n$ we have

$$\det A = |A| = a_{1j} A_{1j} + a_{2j} A_{2j} + \dots + a_{nj} A_{nj}$$

$$= (-1)^j [-a_{1j} M_{1j} + a_{2j} M_{2j} \dots + (-1)^n a_{nj} M_{nj}]$$

$$\text{e.g. } \begin{vmatrix} 1 & 3 & 5 \\ 7 & 9 & 11 \\ 13 & 15 & 17 \end{vmatrix} = (-1)^2 \left[-3 \begin{vmatrix} 7 & 11 \\ 13 & 17 \end{vmatrix} + 9 \begin{vmatrix} 1 & 5 \\ 13 & 17 \end{vmatrix} - 15 \begin{vmatrix} 1 & 5 \\ 7 & 11 \end{vmatrix} \right]$$

$$= [-3 (7 \times 17 - 11 \times 13) + (17 - 13 \times 5) - 15 (11 - 35)]$$

$$= [-3 (119 - 143) + 9 (-48) - 15 (-24)]$$

$$= [-3 (-24) - 18 (24) + 15 \times 24]$$

= 0 which is the value of this determinant.

Property3. For $1 \leq i \neq k \leq n$, $\sum_{j=1}^n a_{ij} A_{kj} = 0$ i.e. multiplying the elements of any of the row by the cofactors of corresponding elements of another

row gives 0. This result also holds by replacing "rows" by columns i.e. For

$$1 \leq j \neq k \leq n \quad \sum_{i=1}^n a_{ij} A_{ik} = 0$$

Property4. Using the corresponding statements (for rows and columns) of (2) and 3 above we get that for $i \neq k$

$$\begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{i1} & \dots & a_{in} \\ \vdots & & \vdots \\ a_{k1} & \dots & a_{kn} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} = 0 \quad (\text{opening by } k^{\text{th}} \text{ row}) \text{ i.e. if a determinant has}$$

two identical row (or columns) then its value is 0

$\therefore$ we also have

$$\begin{vmatrix} a_{11} & a_{1j} & a_{1k} & a_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{ni} & a_{nj} & a_{nk} & a_{nn} \end{vmatrix} = 0 \quad (\text{opening by } k^{\text{th}} \text{ column})$$

Property5. If we multiply every entry of any row (or column) by any fixed scalar α , the value of the determinant gets multiplied by α . i.e.

$$\begin{vmatrix} a_{11} & \dots & a_{1n} \\ \alpha a_{i1} & \dots & \alpha a_{in} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} = \alpha \begin{vmatrix} a_{11} & \dots & a_{1n} \\ a_{i1} & \dots & a_{in} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & \alpha a_{1j} & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{ni} & \alpha a_{nj} & a_{nn} \end{vmatrix}$$

Property6. Applying the above repeatedly for each of the rows or columns. We get $\text{Det}(\alpha A) = \alpha^n \det(A)$ for any scalar α and any n -square matrix A .

Property7. Interchange of any two distinct rows (or columns) results in multiplication of the determinant by (-1) . i.e.

$$\begin{vmatrix} a_{11} & \dots & a_{in} \\ \vdots & & \vdots \\ a_{i1} & \dots & a_{in} \\ \vdots & & \vdots \\ a_{j1} & \dots & a_{jn} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} = - \begin{vmatrix} a_{11} & \dots & a_{in} \\ \vdots & & \vdots \\ a_{j1} & \dots & a_{jn} \\ \vdots & & \vdots \\ a_{i1} & \dots & a_{in} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{1k} & a_{1j} & a_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{ni} & a_{nk} & a_{nj} & a_{nn} \end{vmatrix}$$

Property 4) clearly follows from (7) as can be seen by taking two identical rows columns.

Property8. If we write any row R_i (or column C_j) of a determinant as a sum of two vectors (with the same number of components of course) in the form

$R_i = R'_i + R''_i$ (or $C_j = C'_j + C''_j$) then the determinant is the sum of two determinants whose i th rows are R'_i and R''_i (whose j th columns are C'_j and C''_j) respectively with everything else same as the original determinant.

$$\text{Thus if } \begin{bmatrix} a_{11} & a_{1j} & a_{1n} \\ a_{i1} & a_{ij} & a_{in} \\ a_{n1} & a_{nj} & a_{nn} \end{bmatrix} = \begin{bmatrix} R_1 \\ R_i \\ R_n \end{bmatrix} = [C_1 C_2 \dots C_n]$$

$$\text{Then } \det A = \begin{vmatrix} a_{11} & a_{1j} & a_{1n} \\ a_{i1} & a_{ij} & a_{in} \\ a_{n1} & a_{nj} & a_{nn} \end{vmatrix} = \det \begin{pmatrix} R_1 \\ R'_i \\ R_n \end{pmatrix} + \det \begin{pmatrix} R_1 \\ R''_i \\ R_n \end{pmatrix}$$

$$= \det ([C_1 \dots C'_j \dots C_n]) + \det ([C_1, \dots C''_j \dots C_n])$$

Example:

$$\begin{vmatrix} 1 & 3 & 5 \\ 0 & 2 & 3 \\ -1 & 0 & 3 \end{vmatrix} + \begin{vmatrix} 1 & 3 & 5 \\ 1 & -2 & 6 \\ -1 & 0 & 3 \end{vmatrix} = \begin{vmatrix} 1 & 3 & 5 \\ (0+1) & (2-2) & (3+6) \\ -1 & 0 & 3 \end{vmatrix} = \begin{vmatrix} 1 & 3 & 5 \\ 1 & 0 & 9 \\ -1 & 0 & 3 \end{vmatrix}$$

$$\text{Example: } \begin{vmatrix} 1 & 5 \\ 3 & 7 \end{vmatrix} = \begin{vmatrix} 3+(-2) & 5 \\ 2+1 & 7 \end{vmatrix} = \begin{vmatrix} 3 & 5 \\ 2 & 7 \end{vmatrix} + \begin{vmatrix} -2 & 5 \\ 1 & 7 \end{vmatrix}$$

Property9. If we add a scalar times any particular row (or column) of a determinant to another row (or column) then the value of the determinant remains unchanged. Actually this result is a straight forward consequence of properties 8, 5 and 4 stated above, for $i \neq j$ if $A = [C_1, \dots C_i, \dots C_n]$,

$$\det([C_1, \dots C_i + \alpha C_j, \dots C_j, \dots C_n])$$

$$= \det ([C_1, \dots C_i, \dots C_j, \dots C_n]) + \det([C_1, \alpha C_j \dots C_j, \dots C_n]) \text{ by (8)}$$

$$= \det (A) + \alpha \det ([C_1, \dots C_j \dots C_j, \dots C_n]) \text{ by (6)}$$

$= \det (A) + \alpha \cdot 0$ by (4) (because the second determinant identical i th and j th c
A similar argument holds for row manipulation instead of columns. This property is often used to facilitate evaluation of determinants where it is used to change a given determinant to a determinant with same value but which is

easy to evaluate in view of it having as many zero entries as possible and if required then property(2) is used for final evaluation

columns) = $\det(A)$.

7.6 Minor and Complementary minor of determinant $|A|$: Let A be an $n \times n$ matrix and D be an $r \times r$ sub-matrix of A obtained by deleting any $(n - r)$ rows and $(n - r)$ columns. Then $|D|$ is called a minor of $|A|$ of order r . The determinant of the submatrix D' of A obtained by deleting the rows and columns of A in which elements of D occur is called the complementary minor of $|D|$.

Minor of an element a_{ij} in $|a_{ij}|$: Let $|a_{ij}|$ be a determinant of order n .

Let M_{ij} be the $(n - 1) \times (n - 1)$ submatrix of $|a_{ij}|$ obtained by deleting the i^{th} row and j^{th} column of $|a_{ij}|$. Then $|M_{ij}|$ is called the minor of the element a_{ij} .

Remark 1: The co-factor C_{ij} of the element a_{ij} in $|a_{ij}|$ is given by $C_{ij} = (-1)^{i+j} |M_{ij}|$.

- 2 If i^{th} row of A , of the determinant $|A| = |a_{ij}|$ of order n be replaced by $(A_i + cA_k)$, where $c \neq 0 \in F$ and A_k is the k^{th} row, then the value of the determinant does not change. This property is also true in the case of column of the determinant.
3. We can change any row (column) of a determinant $|A|$, by adding with its elements k times the corresponding elements of any other row (column), without altering of the value of the determinant $|A|$.

7.7 Evaluation of a determinant:

Example: Evaluation of
$$\begin{vmatrix} a-b-c & 2a & 2a \\ 2b & b-c-a & 2b \\ 2c & 2c & c-a-b \end{vmatrix} = \Delta \text{ (say)}$$

We may of course evaluate it using (2). But it is tiresome and chances of mistakes are higher. A smarter way is as follows. We add II and III rows to the first without changing its value (by a_{ij})

$$\therefore \Delta = \begin{vmatrix} a+b+c & b+c+a & c+a+b \\ 2b & b-c-a & 2b \\ 2c & 2c & c-a-b \end{vmatrix}$$

$$\begin{aligned}
 &= (a+b+c) \begin{vmatrix} 1 & 1 & 1 \\ 2b & b-c-a & 2b \\ 2c & 2c & c-a-b \end{vmatrix} \quad \text{by (5)} \\
 &= (a+b+c) \begin{vmatrix} 1 & 0 & 0 \\ 2b & -b-c-a & 0 \\ 2c & 0 & -c-a-b \end{vmatrix} \quad \text{by (a)}
 \end{aligned}$$

as we have added (-1) times I column to II and III columns

$$\begin{aligned}
 &= (a+b+c) (-1)^2 (a+b+c) (a+b+c) \quad \text{by (2) opening it by I row} \\
 &= (a+b+c)^3.
 \end{aligned}$$

7.8 Laplace's expansion of a determinant: Let A be an $n \times n$ square matrix. We form minors $|D|$ of order r from the elements of the first r columns of A and its corresponding complementary minors $|D'|$ then, $|A| = \sum \pm |D| |D'| \dots (1)$

Where the summation is extended over all the possible minors of $|A|$ of order r formed from the elements of the first r columns and the $+$ or $-$ sign is prefixed according as an even number or an odd number of interchanges of adjacent rows of A is necessary to bring the submatrix D into the first r rows of A . The expansion of $|A|$ as given in (1) is called **Laplace's expansion** of the determinant

Example: Let $|A| = \begin{bmatrix} a_1 & a_2 & a_3 & a_4 \\ b_1 & b_2 & b_3 & b_4 \\ c_1 & c_2 & c_3 & c_4 \\ d_1 & d_2 & d_3 & d_4 \end{bmatrix}$

Expand $|A|$ by Laplace's expansion by the minors of the first two columns.

Solution: The possible minors and their complementary minors are as follows:

$$|D_1| = \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}, \quad |D'_1| = \begin{bmatrix} c_3 & c_4 \\ d_3 & d_4 \end{bmatrix}$$

$$|D_2| = \begin{vmatrix} a_1 & a_2 \\ c_1 & c_2 \end{vmatrix}, \quad |D'_2| = \begin{bmatrix} b_3 & b_4 \\ d_3 & d_4 \end{bmatrix}$$

$$|D_3| = \begin{vmatrix} a_1 & a_2 \\ d_1 & d_2 \end{vmatrix}, \quad |D'_3| = \begin{bmatrix} b_3 & b_4 \\ c_3 & c_4 \end{bmatrix}$$

$$|D_4| = \begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix}, \quad |D'_4| = \begin{bmatrix} a_3 & a_4 \\ d_3 & d_4 \end{bmatrix}$$

$$|D_5| = \begin{vmatrix} b_1 & b_2 \\ d_1 & d_2 \end{vmatrix}, \quad |D'_5| = \begin{bmatrix} a_3 & a_4 \\ c_3 & c_4 \end{bmatrix}$$

$$|D_6| = \begin{vmatrix} c_1 & c_2 \\ d_1 & d_2 \end{vmatrix}, \quad |D'_6| = \begin{bmatrix} a_3 & a_4 \\ b_3 & b_4 \end{bmatrix}$$

Hence,
$$|A| = \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} \begin{bmatrix} c_3 & c_4 \\ d_3 & d_4 \end{bmatrix} - \begin{vmatrix} a_1 & a_2 \\ c_1 & c_2 \end{vmatrix} \begin{bmatrix} b_3 & b_4 \\ d_3 & d_4 \end{bmatrix} + \begin{vmatrix} a_1 & a_2 \\ d_1 & d_2 \end{vmatrix} \begin{bmatrix} b_3 & b_4 \\ c_3 & c_4 \end{bmatrix} - \begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix} \begin{bmatrix} a_3 & a_4 \\ d_3 & d_4 \end{bmatrix} + \begin{vmatrix} b_1 & b_2 \\ d_1 & d_2 \end{vmatrix} \begin{bmatrix} a_3 & a_4 \\ c_3 & c_4 \end{bmatrix} - \begin{vmatrix} c_1 & c_2 \\ d_1 & d_2 \end{vmatrix} \begin{bmatrix} a_3 & a_4 \\ b_3 & b_4 \end{bmatrix}$$

Note: A determinant can be expanded by Laplace's method by minors of any r rows or any r columns, provided proper account is taken of the signs of the products.

Example: Show that
$$\begin{vmatrix} a+b+2c & a & b \\ c & b+c+2a & b \\ c & a & c+a+2b \end{vmatrix} = 2(a+b+c)^3$$

Solution:
$$\begin{vmatrix} a+b+2c & a & b \\ c & b+c+2a & b \\ c & a & c+a+2b \end{vmatrix}$$

Applying $R_1 = R_1 - R_2$ and $R_2 = R_2 - R_3$

$$= \begin{vmatrix} a+b+c & -(a+b+c) & 0 \\ 0 & b+c+a & -(b+c+a) \\ c & a & c+a+2b \end{vmatrix}$$

Taking Common $(a+b+c)$

$$= (a+b+c)^2 \begin{vmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ c & a & c+a+2b \end{vmatrix}$$

Again Applying $C_2 = C_2 + C_1$

$$= (a+b+c)^2 \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ c & a+c & a+c+2b \end{vmatrix}$$

Expanding the determinant along the first row, we get

$$\begin{aligned} &= (a+b+c)^2 \begin{vmatrix} 1 & -1 \\ a+c & c+a+2b \end{vmatrix} \\ &= 2(a+b+c)^3 \end{aligned}$$

Example: Evaluate of $\begin{vmatrix} bc & a & a^2 \\ ca & b & b^2 \\ ab & c & c^2 \end{vmatrix}$

Solution: we have $\Delta = \begin{vmatrix} bc & a & a^2 \\ ca & b & b^2 \\ ab & c & c^2 \end{vmatrix}$

Multiplying R_1, R_2 and R_3 by a, b and c respectively, we get

$$= \frac{1}{abc} \begin{vmatrix} abc & a^2 & a^3 \\ abc & b^2 & b^3 \\ abc & c^2 & c^3 \end{vmatrix}$$

Taking common abc from c_1 , we have

$$= \frac{abc}{abc} \begin{vmatrix} 1 & a^2 & a^3 \\ 1 & b^2 & b^3 \\ 1 & c^2 & c^3 \end{vmatrix}$$

$$= \begin{vmatrix} 1 & a^2 & a^3 \\ 1 & b^2 & b^3 \\ 1 & c^2 & c^3 \end{vmatrix}$$

Example: Prove that $\Delta = \begin{vmatrix} b+c & a+b & a \\ c+a & b+c & b \\ a+b & c+a & c \end{vmatrix} = a^3 + b^3 + c^3 - 3abc$

Solution: Applying $c_1 = c_1 + c_3$, we get

$$\Delta = \begin{vmatrix} a+b+c & a+b & a \\ a+b+c & b+c & b \\ a+b+c & c+a & c \end{vmatrix}$$

$$\Delta = (a+b+c) \begin{vmatrix} 1 & a+b & a \\ 1 & b+c & b \\ 1 & c+a & c \end{vmatrix}$$

Applying $R_2 = R_2 - R_1$, and $R_3 = R_3 - R_1$, we get

$$\Delta = (a+b+c) \begin{vmatrix} 1 & a+b & a \\ 0 & c-a & b-a \\ 0 & c-b & c-a \end{vmatrix}$$

$$\Delta = (a+b+c) \begin{vmatrix} c-a & b-a \\ c-b & c-a \end{vmatrix}$$

$$\Delta = (a+b+c)[(c-a)^2 - (c-b)(b-a)]$$

$$\Delta = a^3 + b^3 + c^3 - 3abc$$

Example: Show that $\begin{vmatrix} 1 & a & b+c \\ 1 & b & c+a \\ 1 & c & a+b \end{vmatrix} = 0$

Solution: We have $\Delta = \begin{vmatrix} 1 & a & b+c \\ 1 & b & c+a \\ 1 & c & a+b \end{vmatrix}$

Applying $c_3 = c_3 + c_2$ we get

$$\Delta = \begin{vmatrix} 1 & a & a+b+c \\ 1 & b & b+c+a \\ 1 & c & a+b+c \end{vmatrix}$$

Taking $(a+b+c)$ common from c_3

$$\Delta = \begin{vmatrix} 1 & a & 1 \\ 1 & b & 1 \\ 1 & c & 1 \end{vmatrix}$$

Since c_1 and c_3 are identical

$$\Delta = 0$$

Check your progress

1. Prove that $\begin{vmatrix} a & b & c & d \\ e & f & g & h \\ 0 & 0 & j & k \\ 0 & 0 & l & m \end{vmatrix} = \begin{vmatrix} a & b \\ e & f \end{vmatrix} \cdot \begin{vmatrix} j & k \\ l & m \end{vmatrix}$

Hint: Expand the determinant by the minor of the first two rows. Out of six possible minors, the only non-zero minor is $\begin{vmatrix} a & b \\ e & f \end{vmatrix}$ and its complementary minor is $\begin{vmatrix} j & k \\ l & m \end{vmatrix}$

7.9 Product of two Determinants: The product of two square determinants of order n is a square determinant of order n . If A and B are both $n \times n$ square matrices then $\det(AB) = \det(A) \cdot \det(B)$ i.e. $|A \cdot B| = |A| \cdot |B|$

Where AB denotes the matrix product of matrices A and B and the product on the right hand side is the product of the scalar (members of S) values of $\det(A)$ and $\det(B)$. As may be checked by directly evaluating the determinants on the two sides. The property stated above looks natural and obvious but has a proof which is not easy. It can also be stated as

$$\det(AB) = \begin{vmatrix} \underline{R_1} \cdot \underline{C'_1} & \underline{R_1} \cdot \underline{C'_2} & \dots & \underline{R_1} \cdot \underline{C'_n} \\ \vdots & & & \\ \underline{R_n} \cdot \underline{C'_1} & \underline{R_n} \cdot \underline{C'_2} & \dots & \underline{R_n} \cdot \underline{C'_n} \end{vmatrix} = \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} \cdot \begin{vmatrix} b_{11} & \dots & b_{1n} \\ \vdots & & \vdots \\ b_{n1} & \dots & b_{nn} \end{vmatrix}$$

in our earlier notation where R_i and R'_i are row vectors of A and B respectively and C_j, C'_j are their columns. Since the values of the determinants do not change by replacing rows by columns and vice versa, the right hand side $\det A \cdot \det B = \det B \cdot \det A$ is also equal to the following determinants :

$$\begin{vmatrix} \underline{R_1 R'_1} & \dots\dots\dots & \underline{R_1 R'_n} \\ \underline{R_n R'_1} & \dots\dots\dots & \underline{R_n R'_n} \end{vmatrix} = \begin{vmatrix} \underline{C_1 C'_1} & \underline{C_2 C'_2} & \dots\dots\dots & \underline{C_1 C'_n} \\ \underline{C_n C'_1} & \dots\dots\dots & \dots\dots\dots & \underline{C_n C'_n} \end{vmatrix} \\ = \begin{vmatrix} \underline{C_1 R'_1} & \dots\dots\dots & \underline{C_n R'_1} \\ \underline{C_n R'_1} & \dots\dots\dots & \underline{C_n R'_n} \end{vmatrix} \text{ etc}$$

Theorem: If A and B be two $(n \times n)$ matrices then $|AB^t| = |A| \cdot |B|$

Proof: Since we know that $|AB^t| = |A| \cdot |B^t| = |A| \cdot |B|$ ($|B| = |B^t|$).

Example: $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} x & y \\ z & w \end{bmatrix} = \begin{bmatrix} x+2z & y+2w \\ 3x+4z & 3y+4w \end{bmatrix}$

Check your progress

1. $A = \begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix}, B = \begin{vmatrix} 1 & a & a^2 \\ 1 & b & b^2 \\ 1 & c & c^2 \end{vmatrix}$

Find AB and BA.

7.10 Summary: After studying this unit you should be able to determinant of a square matrix, determine cofactor of an element of a determinant, determine Properties of determinant, determine Minor and Complementary minor of determinant, evaluation of a determinant, determine Laplace expansion of a determinant and determine product of two determinants.

7.11 Terminal Questions and Answers

1. Evaluate: $\begin{vmatrix} 1 & 3 & -1 \\ -2 & 1 & 2 \\ 4 & 5 & 1 \end{vmatrix}$, Answer 35
2. Evaluate: $\begin{vmatrix} a & b \\ -b & a \end{vmatrix}$, Answer $a^2 + b^2$
3. Evaluate: $\begin{vmatrix} 1 & 2 & 3 \\ 2 & 3 & 6 \\ x & y & z \end{vmatrix}$, Answer $3x - z$
4. Prove that $\begin{vmatrix} \sin 10^\circ & -\cos 10^\circ \\ \sin 80^\circ & \cos 80^\circ \end{vmatrix} = 1$
5. Evaluate: $\begin{vmatrix} 1 & \omega & \omega^2 \\ \omega & \omega^2 & 1 \\ \omega & 1 & \omega \end{vmatrix}$, Answer 0

6. Evaluate: $\begin{vmatrix} 6 & 2 & 3 \\ 2 & 3 & 5 \\ 4 & 2 & 1 \end{vmatrix}$, Answer - 30

7. Prove that $\begin{vmatrix} bc & a & a^3 \\ ca & b & b^3 \\ ab & c & c^3 \end{vmatrix} = \begin{vmatrix} 1 & a^2 & a^3 \\ 1 & b^2 & b^3 \\ 1 & c^2 & c^3 \end{vmatrix}$

8. Evaluate: $\begin{vmatrix} 2 & 0 & 5 \\ 3 & 1 & 1 \\ 1 & 4 & 6 \end{vmatrix}$, Answer 59

9. Evaluate: $\begin{vmatrix} 6 & 3 & 7 \\ 32 & 13 & 37 \\ 10 & 4 & 11 \end{vmatrix}$, Answer 10

10. Evaluate: $\begin{vmatrix} 1 & 4 & 2 \\ 2 & 5 & 3 \\ 3 & 6 & 4 \end{vmatrix}$, Answer 0

Unit-8

Eigen vector and eigen space

Structure

8.1 Introduction

8.2 Objectives

8.3 Adjoint of n –square matrix

8.4 Inverse of a matrix of a non- singular matrix by using adjoint of the matrix

8.5 Characteristic roots or eigen values of a linear transformation

8.6. Eigen vector and eigen space

8.7 Summary

8.8 Terminal Questions and Answers

8.1 INTRODUCTION

This unit introduces the concept of determinants and discusses their important properties. Also these are used to calculate inverses of square matrices wherever they exist.

8.2 Objectives

After studying this unit you should be able to

- Adjoint of n – square matrix.
- Inverse of a matrix of a non-singular matrix by using adjoint of the matrix
- Characteristic roots or eigen values of a linear transformation .
- Eigen vector and eigen space

8.3 Adjoint of an n – square matrix: Let $A = [a_{ij}]$ be an n – square matrix. Let C_{ij} be the co-factor of a_{ij} . Then transpose of the matrix $[C_{ij}]$ is that $[C_{ij}]^t$ is called the adjoint of the matrix A and it is written as $Adj(A)$.

Theorem: If $A = [a_{ij}]$ be an n – square matrix, then $A.Adj(A) = Adj(A).A = |A|.I$

Proof: Let C_{kj} be the co-factor of a_{kj} . Then $Adj(A) = [C_{kj}]^t$

Let $Adj(A) = [C^t_{jk}]$, then $C_{kj} = C^t_{jk}$.

Now, $A.Adj(A) = [a_{ij}].[C^t_{jk}] = [d_{ik}]$, where, $d_{ik} = \sum_{j=1}^n a_{ij}.C^t_{jk} = \sum_{j=1}^n a_{ij}.C_{kj} = |A|$, if $i = k$ or $= 0$, if $i \neq k$

Thus all the principal diagonal element of $[d_{ik}]$ are $|A|$ and rest elements are zero. Hence, $A.Adj(A) =$

$$\begin{bmatrix} |A| & 0 & \dots & 0 \\ \vdots & |A| & \ddots & 0 \\ 0 & \dots & \dots & |A| \end{bmatrix} = |A|.I$$

Similarly we can prove that $Adj(A).A = |A|.I$

Theorem: If $A = [a_{ij}]$ be an n – square matrix, and $|A| \neq 0$ then $A^{-1} = \frac{1}{|A|}.Adj(A)$.

Proof: Since, $A.Adj(A) = Adj(A).A = |A|.I$

Or, $A \cdot \frac{Adj(A)}{|A|} = \frac{Adj(A)}{|A|} \cdot A = I$. Therefore, $A^{-1} = \frac{1}{|A|} \cdot Adj(A)$.

Note: This theorem gives a method of finding of the inverse of a non-singular matrix by using the notion of determinant.

Theorem: If $A = [a_{ij}]$ be an n – square matrix, and $|A| \neq 0$, then prove that

$$(i). |Adj(A)| = |A|^{n-1}, (ii). |C_{ij}| = |A|^{n-1}$$

Proof:(i). From the above theorem we have

$$A \cdot Adj(A) = Adj(A) \cdot A = \begin{bmatrix} |A| & 0 & \cdots & 0 \\ \vdots & |A| & \ddots & 0 \\ 0 & \cdots & \cdots & |A| \end{bmatrix} \text{ Therefore,}$$

$$|A \cdot Adj(A)| = |Adj(A) \cdot A| = |A| \cdot |Adj(A)| = \begin{bmatrix} |A| & 0 & \cdots & 0 \\ \vdots & |A| & \ddots & 0 \\ 0 & \cdots & \cdots & |A| \end{bmatrix} = |A|^n$$

Hence, $|Adj(A)| = |A|^{n-1}$ (Since, $|A| \neq 0$)

(ii). Since, we have, $Adj(A) = [C_{kj}]^t$ and $|Adj(A)| = |C_{ij}| = |C_{ij}^t| = |A|^{n-1}$

Theorem: If $A = [a_{ij}]$ be an n – square matrix, and $|A| \neq 0$, then prove that $Adj(Adj(A)) = |A|^{n-2} A$.

Proof: Let $Adj(A) = B$. Now, $Adj(B) \cdot B = |B| \cdot I$

Or, $Adj(B) \cdot (BA) = |B| \cdot IA$. or $Adj(B) \cdot (Adj(A) \cdot A) = |B| \cdot A$

Or, $Adj(B) \cdot (|A| \cdot I) = |B| \cdot A$. Hence, $|A| \cdot Adj(B) = |Adj(A)| \cdot A = |A|^{n-1} \cdot A$

Therefore, $Adj(B) = |A|^{n-1} \cdot A$. or, $Adj(Adj(A)) = |A|^{n-2} A$.

8.4 Inverse of a matrix of a non- singular matrix by using adjoint of the matrix:

We have already defined the minor and cofactor determinants of any entry a_{ij} of a square matrix or a determinant. For an $n \times n$ matrix $A [a_{ij}]$ we have denoted the cofactor of an entry a_{ij} by A_{ij} . Thus

$$A_{ij} = (-1)^{i+j} \begin{vmatrix} a_{11} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1,n} \\ \vdots & & & & & \\ a_{i-1,1} & & a_{i-1,j-1} & a_{i-1,j+1} & & \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ a_{n,1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{n,n} \end{vmatrix} \quad (1 \leq i, j \leq n)$$

We form an $n \times n$ matrix whose (i, j) entry is A_{ij} , and call it the cofactor matrix of A . $\text{Cof}(A) = [A_{ij}]$, $1 \leq i, j \leq n$. The transpose of this matrix is called the adjugate or the adjoint matrix of A , $\text{Adj}(A) = [(\text{cof}(A))^T] = [A_{ji}]$ i.e. the (i, j) entry of $\text{Adj}(A)$ is A_{ji} = cofactor determinant of the entry a_{ji} of A .

$$\begin{aligned}\text{Now } [A \cdot \text{adj}A]_{ij} &= \sum_{k=1}^n a_{ik} (\text{adj}A)_{kj} \\ &= \sum_{k=1}^n a_{ik} A_{jk} = \begin{cases} 0 & \text{if } i \neq j, \text{ by (2) and (3)} \\ \det A, & \text{if } i = j \end{cases} \\ &= \delta_{ij} \det A = \det A \cdot (I_n)_{ij}.\end{aligned}$$

$$= \begin{bmatrix} \det A & 0 & \dots & 0 \\ 0 & \det A & \dots & 0 \\ 0 & 0 & \dots & \det A \end{bmatrix}_{ij} = (\det A \cdot I_n)_{ij}.$$

$$\text{Similarly } [(\text{adj}A)_{ij}A] = \sum_{k=1}^n (\text{adj}A)_{ik} a_{kj} = \sum_{k=1}^n A_{ki} a_{kj} = \delta_{ij} \det A = (\det A \cdot I_n)_{ij},$$

[by (2) and (3)]. Since the above holds for all $1 \leq i, j \leq n$, we get the matrix equation

$$A \times (\text{adj}A) = \Delta I_n = (\text{adj}A) \cdot A \text{ (where } \Delta = \det A)$$

Now in case $\Delta \neq 0$, by multiplying all the parts by $\frac{1}{\Delta}$,

we get $A \left(\frac{\text{adj}A}{\Delta} \right) = \left(\frac{\text{adj}A}{\Delta} \right) A = I_n$ where $\frac{\text{adj}A}{\Delta}$ stands for the matrix

$$= \begin{bmatrix} \frac{A_{11}}{\Delta} & \frac{A_{21}}{\Delta} & \dots & \frac{A_{n1}}{\Delta} \\ \vdots & & & \\ \frac{A_{1n}}{\Delta} & \dots & \dots & \frac{A_{nn}}{\Delta} \end{bmatrix}$$

since the $\text{Adj}A$ is the transpose of the cofactor matrix

$$\begin{bmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{n1} & \dots & \dots & A_{nn} \end{bmatrix}$$

The above relation shows by uniqueness of A^{-1} (if exists) that if $\Delta = \det A \neq 0$ then A is invertible and $A^{-1} = \frac{1}{\Delta} \text{adj} A$, which shows that $\det A \neq 0$ is a sufficient condition for existence of A^{-1} . It is also a necessary condition because if A is invertible then $AA^{-1} = I_n$ gives

$$\det A \cdot \det A^{-1} = \det I_n = 1 \text{ which shows that } \det A \neq 0.$$

Theorem: An n - square matrix A is non-singular if and only if $|A| \neq 0$.

Proof: Let the matrix A is non-singular, hence its inverse A^{-1} exists, and so, $AA^{-1} = I$, hence, $|AA^{-1}| = |A| \cdot |A^{-1}| = |I| = 1$. So, $|A| \neq 0$.

Again if it possible then we suppose that $|A| = 0$, and let A be non-singular, then

$1 = |I| = |AA^{-1}| = |A| \cdot |A^{-1}| = 0 \cdot |A^{-1}| = 0$, which is not possible. Hence, A is singular. So, if $|A| = 0 \Rightarrow A$ is singular. Therefore, A is non-singular $\Rightarrow |A| \neq 0$.

Conversely, now we want to show that $|A| \neq 0 \Rightarrow A$ is non-singular. If it possible let A be singular, then $\text{rank}(A) = r < n$. so, normal form of the matrix $= \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} \sim A$. Hence, $\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} = |A|(-1)^p$. Hence, $|A| = 0$. Since, $\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} = 0$, which contradicts that $|A| \neq 0$

Theorem: If A and B are two n - square similar matrices, then prove that $|A| = |B|$.

Proof: Since, A and B are two n - square similar matrices, so, there exists a non-singular matrix P such that $B = PAP^{-1}$. Hence, $|B| = |PAP^{-1}| = |P||A||P^{-1}|$

$$|A||P||P^{-1}| = |A| \cdot |PP^{-1}| = |A||I| = |A|$$

Example: Examine invertibility and find A^{-1} in case it exists, for

$$A = \begin{bmatrix} 1 & 0 & 3 \\ 5 & 6 & 9 \\ 4 & 5 & 6 \end{bmatrix}$$

Solution: Note that $\Delta = \det A = \begin{vmatrix} 1 & 0 & 3 \\ 5 & 6 & 9 \\ 4 & 5 & 6 \end{vmatrix} = (6 \times 6 - 5 \times 9) + 3(5 \times 5 - 4 \times 6) = -9 + 3 = -6 \neq 0$

$$\therefore A \text{ is invertible and } \text{Cof } A = \begin{bmatrix} -9 & 6 & 1 \\ 15 & -6 & -5 \\ -18 & 6 & 6 \end{bmatrix}$$

$$\therefore \text{Adj}A = (\text{Cof } A)^T = \begin{bmatrix} -9 & 15 & -18 \\ 6 & -6 & 6 \\ 1 & -5 & 6 \end{bmatrix}$$

$$\therefore A^{-1} = \frac{1}{\Delta} \text{Adj}A$$

$$= \frac{1}{(-6)} \text{Adj}A = \begin{bmatrix} \frac{3}{2} & -\frac{5}{2} & 3 \\ -1 & 1 & -1 \\ -\frac{1}{6} & \frac{5}{6} & -1 \end{bmatrix}$$

$$\text{It may be verified that } AA^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = A^{-1}A.$$

Since $\mathbb{Z}$ is not closed under division by any integer $\neq \pm 1$ and $\Delta = -6$

$\therefore$ we see that although A is a matrix over $\mathbb{Z}$ with $\det A \neq 0$, A^{-1} is a matrix over $\mathbb{Q}$.

Check your progress

1. Examine whether A is invertible and find A^{-1} if it is invertible

$$(i) A = \begin{bmatrix} 6 & 18 \\ 4 & 12 \end{bmatrix}$$

$$(ii) A = \begin{bmatrix} 1 & 1 & -1 \\ -3 & 2 & 3 \\ -2 & -2 & 5 \end{bmatrix}$$

$$(iii) A = \begin{bmatrix} 2 & -1 & 3 \\ 6 & 0 & 1 \\ 18 & -3 & 11 \end{bmatrix}$$

$$(iv) A = \begin{bmatrix} 2 & 1 & 2 \\ 1 & 0 & 5 \\ 2 & 5 & 11 \end{bmatrix}$$

2. If both A and B are $n \times n$ invertible matrices, show that AB is invertible and $(AB)^{-1} = B^{-1}A^{-1}$. Will BA also be invertible.

3. Find the minors M_{23} , M_{22} and cofactors A_{12} and A_{32} in the determinant

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

$$4. \text{ Calculate the determinants } A = \begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix}, B = \begin{vmatrix} 1 & a & a^2 \\ 1 & b & b^2 \\ 1 & c & c^2 \end{vmatrix}$$

Also find AB and BA .

5. For matrices A and B of order $m \times n$ and $n \times m$ respectively, show that

$$(AB)^T = B^T A^T \text{ (Reversal rule).}$$

6. Give examples of two square matrices A and B of same size for which

$$\det(A+B) \neq \det(A) + \det(B).$$

8.5 Characteristic roots or eigen values of a linear

transformation: Let T be a linear transformation on a vector space V over a field F . $\lambda \in F$ is called a characteristic root of T if there exists $0 \neq v \in V$ such that $T(v) = \lambda v$, here, v is called a characteristic vector of t with characteristic root λ . This characteristic root λ is also called latent root or eigen value or proper value of T .

Example: Prove that the identity linear transformation on V has the only characteristic root or eigen value 1, and the zero linear map has the only eigen value 0.

Solution: If $v (\neq 0) \in V$, then.

$$I(v) = \lambda(v); \Rightarrow v = \lambda(v) \Rightarrow 1 = \lambda \quad [\text{Since } v \neq 0]$$

Thus the only eigen value of I is 1.

Similarly Let O be the zero linear map on V .

If $(\neq 0) \in V$, then, $O(v) = \lambda(v)$

$$\Rightarrow 0 = \lambda(v) \Rightarrow 0 = \lambda \quad [\text{Since } v \neq 0]$$

Hence the only eigen value of zero linear transformation is 0.

Example: Let $T: R^2 \rightarrow R^2$ be the linear map which rotates each vector by an angle $\theta = \frac{\pi}{2}$. Prove that T has no eigen value in R .

Solution: If $v = (x, y) (\neq (0, 0)) \in R^2$ then

$$\begin{aligned} T(x, y) &= (x \cos \frac{\pi}{2} - y \sin \frac{\pi}{2}, x \sin \frac{\pi}{2} + y \cos \frac{\pi}{2}); \\ &= (-y, x) \end{aligned}$$

$$\text{Hence, } T(x, y) = \lambda(x, y)$$

$$\Rightarrow (-y, x) = (\lambda x, \lambda y)$$

$$\Rightarrow \lambda x = -y \Rightarrow \lambda x + y = 0; \lambda y = x \Rightarrow x + \lambda y = 0$$

These equations will have a non-zero solution $\neq 0, y \neq 0$,

$$\text{If } \begin{bmatrix} \lambda & 1 \\ -1 & \lambda \end{bmatrix} = 0 \text{ or, if } \lambda^2 + 1 = 0$$

which has no solution in R .

Example: Let T be a linear map on a vector space V . Let V_λ be the collection of all vectors v having the same eigen value λ . Prove that V_λ is a subspace of V called λ -eigen space of V .

Solution: Let $v_1, v_2 \in V_\lambda$ and $\alpha_1, \alpha_2 \in F$, then

$$\begin{aligned} T(\alpha_1 v_1 + \alpha_2 v_2) &= \alpha_1 T(v_1) + \alpha_2 T(v_2) \\ &= \alpha_1 (\lambda v_1) + \alpha_2 (\lambda v_2) \quad [\text{Since } v_1, v_2 \text{ have eigen value } \lambda] \end{aligned}$$

$$\text{or, } T(\alpha_1 v_1 + \alpha_2 v_2) = \lambda (\alpha_1 v_1 + \alpha_2 v_2)$$

Hence, $\alpha_1 v_1 + \alpha_2 v_2$ has also eigen value λ and

$$\text{So, } \alpha_1 v_1 + \alpha_2 v_2 \in V_\lambda.$$

$$\text{Since } T(0) = 0 = \lambda \cdot 0 \text{ So, } 0 \in V_\lambda$$

Hence V_λ is subspace of V .

Example: Let T be a linear transformation on V .

(i) $\lambda \in F$ is an eigen value of T if and only if $(T - \lambda I)$ is Singular.

(ii) T is singular if and only if one of its eigen value is 0.

Solution: (i) If $(T - \lambda I)$ is singular, then there exists a non-zero vector v such that $(T - \lambda I)v = 0$

$$T(v) - \lambda I(v) = 0 \Rightarrow I(v) = \lambda v$$

Hence λ is an eigen value of T .

Now if λ is an eigen value of T , then $\exists v (\neq 0) \in V$

$$\text{Such that } T(v) = \lambda v = \lambda I(v)$$

$$\text{Or, } T(v) - \lambda I(v) = 0$$

or, $(T - \lambda I)v = 0$. Hence $(T - \lambda I)$ is singular.

(ii) Proof being simple, is left for the exercise.

Example: Let T and S be linear transformations on where S is non-Singular. Prove that T and STS^{-1} have the same eigen values, and if v is an eigen vector of T , then $S(v)$ is an eigen vector of STS^{-1} .

Solution: Let $\lambda \in F$ be eigen value of T such that

$$T(v) = \lambda v \dots \dots \dots (1)$$

$$\text{Let, } T_1 = STS^{-1}$$

$$\text{Then, } T_1 S = ST; (T_1 S)(v) = (ST)(v)$$

$$\text{So, } T_1(S(v)) = S(T(v)) = S(\lambda v)$$

$$\text{Hence, } (T_1 S)(v) = \lambda S(v)$$

$$\text{Putting } S(v) = x_1, \text{ we get } T_1(x_1) = \lambda x_1$$

$$\text{Since } v \neq 0, \text{ so } x_1 = S(v) \neq 0$$

Since S is Non Singular. Thus there exists $x_1 = S(v) \neq 0$ such that $T_1(x_1) = \lambda x_1$. Hence, λ is an eigen value of T_1 .

That is λ is an eigen value of $T_1 = STS^{-1}$ and $x_1 = S(v) \neq 0$ is an eigen vector of STS^{-1} .

Example: Show that the non-zero characteristic vectors transformation T corresponding to distinct characteristic of a linear roots are linearly independent.

Solution: Let $v_1, v_2, \dots, v_n$ be non-zero characteristic vectors of the linear map T on V corresponding to distinct characteristic roots $\lambda_1, \lambda_2, \dots, \lambda_n$. We will prove $\{v_1, v_2, \dots, v_n\}$ to be linearly independent by induction.

If $n = 1$, then $\{v_1\}$ is linearly independent since $v_1 \neq 0$.

Hence the Theorem is true for $n = 1$.

Assume the theorem to be true for $(n - 1)$ non-zero characteristic vectors $v_1, v_2, \dots, v_{n-1}$, with distinct eigen values $\lambda_1, \lambda_2, \dots, \lambda_{n-1}$. Let

$$(1) \quad \alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_{n-1} v_{n-1} + \alpha_n v_n = 0 \text{ where } \alpha_i \in F$$

$$\text{Then } T(\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_{n-1} v_{n-1} + \alpha_n v_n) = T(0) = 0$$

$$\text{or } \alpha_1 T(v_1) + \alpha_2 T(v_2) + \dots + \alpha_{n-1} T(v_{n-1}) + \alpha_n T(v_n) = 0$$

$$\text{But } T(v_i) = \lambda_i v_i \text{ for } i = 1, 2, \dots, n. \text{ Hence}$$

$$(2) \quad \alpha_1 \lambda_1 v_1 + \alpha_2 \lambda_2 v_2 + \dots + \alpha_{n-1} \lambda_{n-1} v_{n-1} + \alpha_n \lambda_n v_n = 0$$

Multiplying (1) by λ_n and subtracting from (2) we get

$$(3) \quad \alpha_1 (\lambda_1 - \lambda_n) v_1 + \alpha_2 (\lambda_2 - \lambda_n) v_2 + \dots + \alpha_{n-1} (\lambda_{n-1} - \lambda_n) v_{n-1} = 0$$

Since $\{v_1, v_2, \dots, v_n\}$ is Linearly independent by induction hypothesis, so (3) gives

$$\alpha_1 (\lambda_1 - \lambda_n) = 0, \alpha_2 (\lambda_2 - \lambda_n) = 0, \dots, \alpha_{n-1} (\lambda_{n-1} - \lambda_n) = 0.$$

Since, $(\lambda_i \neq \lambda_n \text{ for } i = 1, 2, \dots, (n - 1))$.

$$\text{Hence, } \alpha_1 = \alpha_2 = \dots = \alpha_{n-1} = 0$$

Therefore from (1) we have $\alpha_n v_n = 0$ or, $\alpha_n = 0$ for $v_n \neq 0$.

Thus, $\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_{n-1} v_{n-1} + \alpha_n v_n = 0$ where $\alpha_i \in F$

$$\Rightarrow \alpha_1 = \alpha_2 = \dots = \alpha_{n-1} = \alpha_n = 0$$

Therefore $\{v_1, v_2, \dots, v_n\}$ is linearly independent.

Example: Let V be a vector space over a field F of dimension n .

Let T and T' be two linear maps on V , associated with the same matrix $A = [a_{ij}]$ with respect to bases $\{x_1, x_2, \dots, x_n\}$ and $\{x'_1, x'_2, \dots, x'_n\}$ respectively. Prove that T and T' have the same eigen values.

Solution: Let $S : V \rightarrow V$ be defined by

$$(1) \quad S(x_j) = x'_j \text{ for } j = 1, 2, \dots, n.$$

It can be easily proved that S is a bijection and hence non-singular and invertible. Now we prove that $T' = STS^{-1}$. Now for $j = 1, 2, \dots, n$, we get $(STS^{-1})(x'_j) = ST(S^{-1}(x'_j))$

$$= ST(x_j), \quad [\text{Since, } (S^{-1}(x'_j)) = x_j]$$

$$= S(T(x_j))$$

$$\text{Hence, } (STS^{-1})(x'_j) = S(\sum_{i=1}^n a_{ij}x_i)$$

(Since Matrix of T relative to basis $\{x_j\}$ is $[a_{ij}]$)

$$= (\sum_{i=1}^n a_{ij}S(x_i)) \quad (\text{Since, } S \text{ is linear})$$

$$= (\sum_{i=1}^n a_{ij}x'_i) \quad (\text{from (1)})$$

$$= T'(x'_j) \quad \{\text{Since matrix of } T' \text{ relative to the basis } \{x'_i\} \text{ is } [a_{ij}]\}$$

Thus, $(STS^{-1})(x'_j) = T'(x'_j)$ for $j = 1, 2, \dots, n$

So, $(STS^{-1})(v) = T'(v)$; for all $v \in V$.

$$\text{Hence, } STS^{-1} = T'$$

Hence T and $STS^{-1} = T'$ have the same eigen values.

We can thus define the eigen value or characteristic root of an n -square matrix $[a_{ij}]$ as follows:

Example: (a). If S and T are Non-Singular Linear Transformations on a vector space of finite dimension. Prove that ST and TS have the same eigen values.

Solution: Let λ be an eigen value of ST

$$\text{So, } \exists v \neq 0, \text{ such that } (ST)(v) = \lambda v$$

$$\text{Or, } S(T(v)) = \lambda v$$

$$\text{Or, } T(v) = S^{-1}(\lambda v) = \lambda S^{-1}(v)$$

$$\text{Or, } T(v) = \lambda v^1$$

$$\text{where } v^1 = S^{-1}(v) \text{ i.e. } S(v^1) = v; v^1 \neq 0$$

$$\text{Now, } (TS)(v^1) = T(S(v^1)) = T(v) = \lambda v^1$$

$$\text{So, } \exists v^1 \neq 0 \text{ such that } (TS)(v^1) = \lambda v^1.$$

Hence, λ is also an eigen value of TS .

So, ST and TS have the same eigen values.

8.6 Eigen vector and eigen space:

Definition: Let $A = [a_{ij}]$ be an n -square matrix over a field F . $\lambda \in F$

is called an eigen value of the matrix $A = [a_{ij}]$ if λ is an eigen value of any linear transformation T on an n -dimensional vector space V , associated with the matrix $A = [a_{ij}]$. If we take $V = F^n$ and consider the linear map $T: F^n \rightarrow F^n$ given by

$$T(X) = [a_{ij}] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \text{ where, } X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in F^n$$

We have proved before that matrix of T relative to the usual basis of F^n is $[a_{ij}]$. Hence, if $\lambda \in F$ be an eigen value of T , then $\lambda \in F$ is also an eigen value of the matrix $[a_{ij}]$ associated with T .

Now $\lambda \in F$ is an eigen value of T , if $\exists X \neq 0$

such that $T(X) = \lambda X$

$$\text{or } T(X) = [a_{ij}] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \lambda x_1 \\ \lambda x_2 \\ \vdots \\ \lambda x_n \end{bmatrix}.$$

Hence we may define eigen value of a matrix $[a_{ij}]$ directly without reference to its associated linear transformation as follows:

Definition: (a) $\lambda \in F$ is called an Eigen characteristic root of the

n -square matrix $[a_{ij}]$ if there exists a non-zero column matrix $\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ such that

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \lambda x_1 \\ \lambda x_2 \\ \vdots \\ \lambda x_n \end{bmatrix}$$

The column matrix $\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ is called the eigen vector of the n –square matrix $[a_{ij}]$ with eigen value $\lambda \in F$

Example: Prove that the eigen values of a diagonal matrix are its diagonal elements. Hence prove that a matrix A similar to a diagonal matrix, has the eigen values as the diagonal elements of the diagonal matrix.

Solution: Let the diagonal matrix $A = \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix}$

If $\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ is an Eigen vector of A with eigen value $\lambda \in F$ then

by Definition we have $\begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \lambda x_1 \\ \lambda x_2 \\ \vdots \\ \lambda x_n \end{bmatrix}$

or, $\begin{bmatrix} \lambda_1 x_1 \\ \lambda_2 x_2 \\ \vdots \\ \lambda_n x_n \end{bmatrix} = \begin{bmatrix} \lambda x_1 \\ \lambda x_2 \\ \vdots \\ \lambda x_n \end{bmatrix}$. Since, $X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \neq 0$, so, there exists $x_i \neq 0$, where $i = 1, 2, \dots, n$ such that $\lambda x_i = \lambda_i x_i$. Or, $\lambda = \lambda_i$ (since, $x_i \neq 0$) Hence, $\lambda_1, \lambda_2, \dots, \lambda_n$ (the diagonal elements) are the eigen values.

Second part obviously follows since a matrix A and the diagonal matrix similar to A are matrices of the same linear transformation T on V relative to different bases and so they have the same eigen values. **Example:** If A and B be n –square matrices over F , then show that AB and BA have the same eigen values.

Solution: If AB has an eigen value 0 then AB is Singular. So either A or B is singular which implies that BA is singular. Therefore BA also must have an eigen value 0.

Now let $\lambda \neq 0$ be an eigen value of AB , then there exists a column $X \neq 0$, such that $(AB)X = \lambda X$

Or, $AY = \lambda X$ where $BX = Y$ (a column matrix)

Now, $(BA)Y = B(AY) = B(\lambda X) = \lambda(BX) = \lambda Y$

Now, $\lambda \neq 0$.

For, $Y = 0 \Rightarrow BX = 0 \Rightarrow (AB)X = 0$

$\Rightarrow X = 0 \Rightarrow X = 0$ since $\lambda \neq 0$.

Thus we get $(BA)Y = \lambda Y$ where $Y \neq 0$

Hence λ , is an eigen value of BA also.

Similarly it can be proved that any non-zero eigen value of BA is also an eigen value of AB . Thus AB and BA have the same eigen values.

Check your progress

1. If v is an eigen vector of a linear map T on V with eigen value λ , then prove that av where $a (\neq 0) \in F$ is also an eigen vector with eigen value λ . Also prove any eigen vector v can not correspond to two different eigen values λ_1 and λ_2 .

2. Give an example of a square matrix of order 3 with entries in Z whose eigen values are 2, 1, 3.

Hint: Apply the theorem and the required Matrix $A = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & \vdots \\ 0 & 0 & 3 \end{bmatrix}$

3. If λ is an eigen value of an invertible linear transformation T , then prove that λ^{-1} is an eigen value of T^{-1} .

Hint: T is Non-singular. Hence $\lambda \neq 0$ so λ^{-1} exists. Now a vector $v \neq 0$ such that $T(v) = \lambda v$

Hence, $T^{-1}(T(v)) = T^{-1}(\lambda v)$

This implies $v = \lambda T^{-1}(v)$.

Or, $T^{-1}(v) = \lambda^{-1} v$.

Theorem: Let $A = [a_{ij}]$ of order $n \times n$. $\lambda \in F$ is a characteristic root of A if and only if $|A - \lambda I_n| = 0$.

Proof: If $\lambda \in F$ is a characteristic root of A , then there exists a non-zero column matrix X , such that $AX = \lambda X$ or, $AX = \lambda I_n X$

or, $AX - \lambda I_n X = 0$, or, $(A - \lambda I_n)X = 0$

which is a system of n - linear homogeneous equation in n unknowns in matrix form. So this system has a non- zero solution X if and only if $|A - \lambda I_n| = 0$. Hence, $\lambda \in F$ is a characteristic root of A if and only if $|A - \lambda I_n| = 0$.

Check your progress

1. Find the eigen values of the Matrix $A = \begin{bmatrix} 1 & 0 & 2 \\ 0 & -1 & 1 \\ 0 & 1 & 0 \end{bmatrix}$.
2. Find the eigen values of the Matrix $A = \begin{bmatrix} 5 & -6 & -6 \\ -1 & 4 & 2 \\ 3 & -6 & -4 \end{bmatrix}$
3. Find the eigen values of the Matrix $A = \begin{bmatrix} 2 & 2 & 1 \\ 1 & 3 & 1 \\ 1 & 2 & 2 \end{bmatrix}$
4. Find the eigen values of the Matrix $A = \begin{bmatrix} 8 & -6 & 2 \\ -6 & 7 & -4 \\ 2 & -4 & 3 \end{bmatrix}$
5. Find the eigen values of the Matrix $A = \begin{bmatrix} 1 & -1 & 0 \\ 1 & 2 & -1 \\ 3 & 2 & -2 \end{bmatrix}$
6. Find the eigen values of the Matrix $A = \begin{bmatrix} 1 & 1 & 1 \\ -1 & -1 & -1 \\ 0 & 0 & 1 \end{bmatrix}$
7. Find the eigen values of the Matrix $A = \begin{bmatrix} 1 & -1 & 2 \\ 2 & -2 & 4 \\ 3 & -3 & 6 \end{bmatrix}$
8. Find the eigen values of the Matrix $A = \begin{bmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{bmatrix}$
9. Find the eigen values of the Matrix $A = \begin{bmatrix} 1 & -1 & -1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix}$
10. Find the eigen values of the Matrix $A = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 2 & 1 \\ 2 & 0 & 3 \end{bmatrix}$

8.7 Summary : After studying this unit you should be able to find adjoint of n –square matrix, determine the inverse of a matrix of a non- singular matrix by using adjoint of the matrix, determine characteristic roots or eigen values of a linear transformation , to find eigen vector and eigen space.

8.8 Terminal Questions and Answers

1. Determine the Eigen values of the matrix
 $A = \begin{bmatrix} 3 & 2 \\ -1 & 0 \end{bmatrix}$, Answer $\lambda = 1, 2$
2. Determine the Eigen values and corresponding Eigen vectors of the matrix

$$A = \begin{bmatrix} 4 & 2 \\ 3 & 3 \end{bmatrix}, \text{ Answer } \lambda = 1, 6 \text{ and } \begin{bmatrix} 2 \\ -3 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

3. Determine the Eigen values of the matrix

$$A = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & -1 \\ 2 & -1 & 0 \end{bmatrix}, \text{ Answer } \lambda = 1, 2, -1 \pm \sqrt{3}$$

4. Determine the Eigen values and corresponding Eigen vectors of the

$$\text{matrix } A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}, \text{ Answer } \lambda = 0, 0, 3 \text{ and } \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

5. Find the characteristic equations of the matrix

$$A = \begin{bmatrix} 10 & -9 \\ 4 & -2 \end{bmatrix}, \text{ Answer } \lambda^2 - 8\lambda + 16 = 0$$

6. Determine the Eigen values of the matrix

$$A = \begin{bmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{bmatrix}, \text{ Answer } \lambda = 8, -1, -1$$

7. Determine the Eigen values of the matrix

$$A = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}, \text{ Answer } \lambda = 2, 2, 2$$

8. Determine the Eigen values of the matrix

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & 3 \\ 0 & 0 & 3 \end{bmatrix}, \text{ Answer } \lambda = 1, 2, 3$$

9. Determine the Eigen values of the matrix

$$A = \begin{bmatrix} 3 & 6 & 6 \\ 0 & 2 & 0 \\ -3 & -12 & -6 \end{bmatrix}, \text{ Answer } \lambda = 0, 2, -3$$

10. Determine the Eigen values of the matrix

$$A = \begin{bmatrix} -8 & -6 & 2 \\ -6 & 7 & -4 \\ 2 & -4 & 3 \end{bmatrix}, \text{ Answer } \lambda = 0, 3, 15$$



**Uttar Pradesh Rajarshi Tondon
Open University**

Bachelor of Science

UGMM - 107

Linear Algebra

BLOCK

3

Unit - 9	Characteristic polynomial of a matrix	183
Unit - 10	Inner Products space	197
Unit - 11	Bilinear, Quadratic & Hermitian Forms	220

Course Design Committee

Prof. Ashutosh Gupta, School of Computer and Information Science, UPRTOU, Prayagraj	Chairman
Prof. Sudhir Srivastav Dept. of Mathematics, DDU Gorakhpur University, Gorakhpur	Member
Prof. P. K. Singh Dept. of Mathematics, Allahabad University, Prayagraj	Member
Prof. Mona Khare Dept. of Mathematics, Allahabad University, Prayagraj	Member
Dr. A. K. Pandey Associate Professor, ECC Prayagraj	Member
Dr. Vikas Singh Academic Consultant, UPRTOU, Prayagraj	Member
Dr. S. S. Tripathi Academic Consultant, UPRTOU, Prayagraj	Member

Course Preparation Committee

Dr. Kamran Alam Khan Assistant Professor (Dept. of Mathematics), Ram Lubhai Sahani Govt. Mahila Degree College, Pilibhit	Author (Unit – 1 - 3)
Dr. Ashok Kumar Pandey Associate Professor, E.C.C, Prayagraj	Author (Unit 4 - 9)
Dr. S. S. Tripathi Academic Consultant School of Science, UPRTOU, Prayagraj	Editor (Unit 1 - 9)
Prof. Ashutosh Gupta Director, School of Computer and Information Science, UPRTOU, Prayagraj	

© UPRTOU , Prayagraj - 2021**© UGMM - 107 Under Algebra****ISBN :****Reprint – 2021 – 100 Copy**

All Rights are reserved. No Part of this work may reproduced in any form, by mimeograph or any other means, without permission in writing from the Uttar Pradesh Rajarshi Tandon Open University.

Printed and Published by Pro. (Dr.) Arun Kumar Gupta Registrar, Uttar Pradesh Rajarshi Tandon Open University, 2021

Printed By. – M/s K.C.Printing & Allied Works, Panchwati, Mathura -281003.

UNIT- 09

Characteristic polynomial of a matrix

Structure

9.1 Introduction

9.2 Objectives

9.3 Characteristic polynomial of a matrix

9.4 Cayley- Hamilton Theorem

9.5.Characteristic polynomial of a linear map

9.6 Polynomial in an n –square matrix

9.7 Minimal polynomial

9.8 Hermitian Matrix

9.9 Summary

9.10 Terminal Questions

9.1 INTRODUCTION

This unit introduces the concept of Characteristic polynomial of a matrix, Cayley- Hamilton Theorem, Characteristic polynomial of a linear map, Polynomial in an n –square matrix, Minimal polynomial, Hermitian Matrix

9.2 Objectives

After studying this unit you should be able to

- Characteristic polynomial of a matrix.
- Cayley- Hamilton Theorem
- Characteristic polynomial of a linear map.
- Polynomial in an n –square matrix
- Minimal polynomial
- Hermitian Matrix

9.3 Characteristic polynomial of a matrix: Suppose that

$A = [a_{ij}]$ of order $n \times n$ matrix, then the determinant $|A - \lambda I_n|$ is a polynomial in $\lambda \in F$ of degree n , is called the characteristic polynomial of matrix A and denoted by $\phi(\lambda)$. Thus the characteristic polynomial of A is given by

$$\phi(\lambda) = |A - \lambda I_n| = \begin{vmatrix} a_{11} - \lambda & a_{12} & a_{1n} \\ a_{21} & a_{22} - \lambda & \vdots \\ a_{n1} & \dots & a_{nn} - \lambda \end{vmatrix}$$

The equation $\phi(\lambda) = |A - \lambda I_n| = 0$ is called the characteristic equation of A .

Note: The characteristic roots of the matrix A are the zeros of its characteristic polynomial $\phi(\lambda) = |A - \lambda I_n|$. This equation is of degree n . Thus a matrix $A = [a_{ij}]$ of order $n \times n$ matrix may have at most n distinct characteristic roots.

Theorem: A linear transformation $T: V \rightarrow V$ can be associated with an n –square diagonal matrix B , if and only if V has a basis consisting of eigen vectors of T . Where $\dim(V) = n$ and the diagonal elements of B are the corresponding eigen values.

Proof: Suppose that the n – distinct eigen vectors $S = \{v_1, v_2, \dots, v_n\}$ be a basis of V . corresponding to distinct given values $\lambda_1, \lambda_2, \dots, \lambda_n$.

$$T(v_1) = \lambda_1 v_1 + 0v_2 + \dots + 0v_{n-1} + 0v_n \text{ where } \lambda_i \in F$$

$$T(v_2) = 0v_1 + \lambda_2 v_2 + \dots + 0v_{n-1} + 0v_n$$

.....

.....

$$T(v_n) = 0v_1 + 0v_2 + \dots + 0v_{n-1} + \lambda_n v_n$$

The matrix with respect to the basis $S = \{v_1, v_2, \dots, v_n\}$

$\begin{bmatrix} \lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_n \end{bmatrix}$ which is a diagonal matrix and the diagonal elements are the eigen values of T .

Note: An n -square matrix A is similar to a diagonal matrix B if and only if A has n linearly independent eigen vectors. The diagonal elements of B are the corresponding eigen values. Suppose P be the matrix whose columns are the n independent eigen vectors of A . Then $B = P^{-1}AP$.

Example: Find the eigen values of the matrix

$$\begin{bmatrix} 0 & 1 & \dots & 0 \\ \vdots & \ddots & 1 & \vdots \\ 1 & \dots & 0 \end{bmatrix}$$

Solution: The characteratic polynomial of the matrix is given by

$$\phi(\lambda) = \begin{vmatrix} -\lambda & 1 & \dots & 0 \\ \vdots & \ddots & 1 & \vdots \\ 1 & \dots & 0 & -\lambda \end{vmatrix} = \lambda^4 - 1$$

$$\lambda^4 - 1 = (\lambda - 1)(\lambda + 1)(\lambda - i)(\lambda + i) = 0$$

Hence, the eigen values are $\pm 1, \pm i$

Example: Find the eigen values and a basis of their eigen spaces of the matrix

$$A = \begin{bmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{bmatrix}$$

Solution: The characteristic polynomial of

$$A = \begin{bmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{bmatrix} \text{ is } \phi(\lambda) = |A - \lambda I_n|$$

$$\begin{bmatrix} 1-\lambda & -3 & 3 \\ 3 & -5-\lambda & 3 \\ 6 & -6 & 4-\lambda \end{bmatrix} = (2+\lambda)^2 \begin{bmatrix} 1 & 0 & 3 \\ 1 & 1 & 3 \\ 0 & 1 & 4-\lambda \end{bmatrix}$$

$$= (2+\lambda)^2(4-\lambda).$$

Hence, eigen values of A are $-2, -2, 4$

We want to find the eigen space V_λ when $\lambda = -2$

Let $X = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ be the eigen vector of A with eigen value $\lambda = -2$. Then

$$(A - \lambda I_n)X = 0$$

Now we put $\lambda = -2$ and $X = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$, we get that

$$\begin{bmatrix} 3 & -3 & 3 \\ 3 & -3 & 3 \\ 6 & -6 & 6 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\text{The coefficient matrix} \sim \begin{bmatrix} 1 & -1 & 1 \\ 1 & -1 & 1 \\ 1 & -1 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

In the echelon form with only one non-zero row. Hence, rank of the coefficient matrix $= \rho = 1$

Hence, the system will have $n - \rho = 3 - 1 = 2$ linearly independent solutions

The reduced form of the system becomes $x - y + z = 0$

We choosing x, y to be arbitrary variables and

$x = 0$ then $y = 1, z = 1$; $x = 1$ then $y = 0, z = -1$. These are two linearly independent solutions. Thus the basis of the eigen space of $\lambda = -2$ is

$$S = \{(0, 1, 1), (1, 0, -1)\}.$$

Similarly $X = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ be the eigen vector of A with eigen value for $\lambda = 4$

$$\begin{bmatrix} 3 & -3 & 3 \\ 3 & -9 & 3 \\ 6 & -6 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

The coefficient matrix $\sim \begin{bmatrix} 1 & 0 & -\frac{1}{2} \\ 0 & 1 & -\frac{1}{2} \\ 0 & 0 & 0 \end{bmatrix}$

In the echelon form with only one non- zero row. Hence, rank of the coefficient matrix $= \rho = 2$

Hence, the system will have $n - \rho = 3 - 2 = 1$ linearly independent solutions

The reduced form of the system becomes $x - \frac{1}{2}z = 0, y - \frac{1}{2}z = 0$

We choosing z to be arbitrary variables and

$z = 2$ then $y = 1, x = 1$; This is the linearly independent solution. Thus the basis of the eigen space of $\lambda = 4$ is $S = \{(1,1,2)\}$.

Hence, the matrix A has three linearly independent solutions, so, three linearly independent vectors $(0,1,1), (1,0,-1), (1,1,2)$.

Example: If A be the given matrix $\begin{bmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{bmatrix}$, find an invertible matrix P such that $P^{-1}AP$ is a diagonal matrix.

Solution: Since $A = \begin{bmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{bmatrix}$ has three linearly independent eigen vectors, so, these three linearly independent vectors are

$(0,1,1), (1,0,-1), (1,1,2)$. So, we have taken $P = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & -1 & 2 \end{bmatrix}$

Then we have $P^{-1}AP = \text{a diagonal matrix} = \begin{bmatrix} -2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 4 \end{bmatrix}$

Thus P is the required invertible matrix.

Theorem: Prove that similar matrices have the same characteristic polynomial and hence the same eigen values.

Proof: Suppose that A and B be two similar matrices. There exists a non-singular matrix S such that $B = S^{-1}AS$.

Now the characteristic polynomial of B is

$$\begin{aligned}
|B - \lambda I_n| &= |S^{-1}AS - S(\lambda I_n)S^{-1}| \\
&= |S^{-1}(A - \lambda I_n)S| = |S||A - \lambda I_n||S^{-1}| \\
&= |A - \lambda I_n||SS^{-1}| = |A - \lambda I_n||I_n| = |A - \lambda I_n| \\
&= \text{Characteristic polynomial of } A.
\end{aligned}$$

Thus A and B have the same characteristic polynomial and hence the same eigen values.

Note: Let $f: V \rightarrow V$ be an endomorphism of a finite dimensional vector space V over a field F . Let $\lambda_1, \lambda_2, \dots, \lambda_n \in F$ be the distinct eigen values of f with corresponding eigen vectors

$x_1, x_2, \dots, x_n \in V$. Let W be the subspace of V generated by $S = \{x_1, x_2, \dots, x_n\}$ then (a). if $f(x) \in W \forall x \in W$. Hence,

$f: W \rightarrow W$ is an endomorphism. (b). $\dim(W) \geq n$.

Theorem: Let A be an $n \times n$ square matrix over a field F . Prove that (i). The sum of the characteristic roots of $A = \text{Trace of } A$.

(ii). The product of the characteristic roots of $A = |A|$.

Proof: Suppose that $A = [a_{ij}]$ of order $n \times n$ matrix, characteristic polynomial of A is $\phi(\lambda) = |A - \lambda I_n| = \begin{vmatrix} a_{11} - \lambda & a_{12} & a_{1n} \\ a_{21} & a_{22} - \lambda & \vdots \\ a_{n1} & \dots & a_{nn} - \lambda \end{vmatrix}$

Hence, $\phi(\lambda) = |A - \lambda I_n| = (a_{11} - \lambda)(a_{22} - \lambda) \dots (a_{nn} - \lambda) +$
terms containing at most $(n - 2)$ factors of the form $a_{ij} - \lambda$.

Therefore, $\phi(\lambda) = |A - \lambda I_n| = (-1)^n \lambda^n + (-1)^{n-1} \lambda^{n-1} (a_{11} + a_{22} + \dots + a_{nn}) + \text{terms of lower degree in } \lambda$.

Hence, the sum of the zeros of $\phi(\lambda) = |A - \lambda I_n| = (a_{11} + a_{22} + \dots + a_{nn}) = \text{Trace of } A$. The sum of the characteristic roots of $A = \text{Trace of } A$

(ii). When we expand $\phi(\lambda) = |A - \lambda I_n|$ as a determinant, then it can be expressed as the sum of the two determinants as follows:

$$\begin{bmatrix} a_{11} & a_{12} & a_{1n} \\ a_{21} & a_{22} & \vdots \\ a_{n1} & \dots & a_{nn} \end{bmatrix} = |A| \text{ and } \begin{bmatrix} -\lambda & 0 & 0 \\ 0 & -\lambda & \vdots \\ 0 & \dots & -\lambda \end{bmatrix} = (-1)^n \lambda^n$$

Hence, the characteristic equation of A is $\phi(\lambda) = |A - \lambda I_n| = 0$

Or, $\phi(\lambda) = |A - \lambda I_n| = (-1)^n \lambda^n + \dots + |A| = 0$

Hence, product of the characteristic roots of $A = \frac{(-1)^n |A|}{(-1)^n} = |A|$.

Theorem: If A and B are two n – square similar matrices, then prove that $|A| = |B|$.

Proof: Since, A and B are two n – square similar matrices, so, there exists a non-singular matrix P such that $B = PAP^{-1}$. Hence, $|B| = |PAP^{-1}| = |P||A||P^{-1}|$

$$|A||P||P^{-1}| = |A| \cdot |PP^{-1}| = |A||I| = |A|$$

Theorem: Prove that the matrix A and its transpose A^t have the same eigen values.

Proof: Since, $(A - \lambda I_n)^t = A^t - \lambda I_n$

Also, $|A - \lambda I_n| = |(A - \lambda I_n)^t| = |A^t - \lambda I_n|$

Hence, A and its transpose A^t have the same eigen values

Check your progress

1. Find the eigen values and a basis of their eigen spaces of the matrix

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

2. Find the eigen values and a basis of their eigen spaces of the matrix

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

3. Find the eigen values and a basis of their eigen spaces of the matrix

$$A = \begin{bmatrix} -2 & 3 & 1 \\ 2 & -1 & 3 \\ 6 & -2 & 2 \end{bmatrix}$$

4. Find the eigen values and a basis of their eigen spaces of the matrix

$$A = \begin{bmatrix} 3 & 10 & 5 \\ -2 & -3 & -4 \\ 3 & 5 & 7 \end{bmatrix}$$

5. Find the eigen values and a basis of their eigen spaces of the matrix

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

6. Prove that the characteristic roots of a real symmetric matrix are all real.

7. Prove that the characteristic roots of a triangular matrix are the elements in the principal diagonal.

8. If λ is an eigen value of an invertible linear transformation T , then prove that λ^{-1} is an eigen value of T^{-1} .

9. Find the characteristic roots and corresponding characteristic vectors of the

$$\text{matrix } A = \begin{bmatrix} 3 & 1 & 1 \\ 2 & 4 & 2 \\ 1 & 1 & 3 \end{bmatrix}$$

10. Find the characteristic roots and corresponding characteristic vectors of the

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

11. Find the characteristic roots and corresponding characteristic vectors of the

$$\text{matrix } A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & -4 & 2 \\ 0 & 0 & 7 \end{bmatrix}$$

9.4 Cayley- Hamilton Theorem: Every square matrix is a root of its characteristic polynomial.

Proof: Let $A = [a_{ij}]$ be an n – square matrix over a field F . Let its characteristic polynomial be $\phi(\lambda) = |A - \lambda I_n| = (-1)^n \{\lambda^n + a_{n-1}\lambda^{n-1} + \dots + a_1\lambda + a_0\}$.

Suppose that the adjoint of $(A - \lambda I_n)$ be B . Now the elements of B are the cofactors of the elements of $|A - \lambda I_n|$ and hence polynomials in λ of degree $(n - 1)$ or less.

Hence, the matrix B can be expressed as the sum of n matrices of order n as follows:

$$B = B_{n-1}\lambda^{n-1} + \dots + B_1\lambda + B_0$$

Where $B_0, B_1, \dots, B_{n-1}$ are n – square matrices independent of λ .

By the property of adjoint of the matrix we have, $A(\text{adj}A) = |A|I_n$

So, $(A - \lambda I_n)B = |A - \lambda I_n|I_n$

$$= (A - \lambda I_n)(B_{n-1}\lambda^{n-1} + \dots + B_1\lambda + B_0) = (-1)^n \{\lambda^n + a_{n-1}\lambda^{n-1} + \dots + a_1\lambda + a_0\}I_n.$$

$$-I_n B_{n-1}\lambda^n + (AB_{n-1} - I_n B_{n-2})\lambda^{n-1} + (AB_{n-2} - I_n B_{n-3})\lambda^{n-2} + \dots + (AB_1 - I_n B_0)\lambda + AB_0 = (-1)^n \{\lambda^n + a_{n-1}\lambda^{n-1} + \dots + a_1\lambda + a_0\}I_n.$$

Now we equating the coefficients of powers of λ on both sides of this equation

$$\text{We get } B_{n-1} = (-1)^n I_n;$$

$$(-1)^n a_{n-1} I_n = AB_{n-1} - B_{n-2};$$

$$(-1)^n a_{n-2} I_n = AB_{n-2} - B_{n-3}$$

.....

$$(-1)^n a_1 I_n = AB_1 - B_0$$

$$(-1)^n a_0 I_n = AB_0$$

Now we adding all these above equations , we have

$$(-1)^n \{A^n + a_{n-1}A^{n-1} + \dots + a_1A + a_0I_n\} = 0$$

$$\text{Or, } \phi(A) = 0.$$

Thus A is a root of the characteristic polynomial $\phi(\lambda)$ of A .

Inverse of a Non- singular Matrix by Cayley- Hamilton

Theorem : Inverse of a non- singular matrix A can be computed by Cayley- Hamilton theorem .

Example: By using of Cayley- Hamilton theorem find the inverse of the matrix

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

Solution: Characteristic polynomial of A is given by

$$\begin{aligned} \phi(x) &= \begin{vmatrix} 1-x & 2 & 3 \\ 0 & 5-x & 0 \\ 2 & 4 & 3-x \end{vmatrix} \\ &= -(x^3 - 9x^2 + 17x + 15) \end{aligned}$$

By Cayley- Hamilton theorem A is a zero of $\phi(x)$, so,

$$A^3 - 9A^2 + 17A + 15I_3 = 0$$

$$\text{Or, } A^3 - 9A^2 + 17A = -15I_3$$

Now we multiply this by A^{-1} , we get

$$A^2 - 9A + 17I_3 = -15A^{-1}$$

$$\text{Therefore, } A^{-1} = \frac{-1}{15}(A^2 - 9A + 17I_3)$$

$$\begin{aligned} \text{Now, } A^2 - 9A + 17I_3 &= \begin{bmatrix} 7 & 24 & 12 \\ 0 & 25 & 0 \\ 8 & 36 & 15 \end{bmatrix} + \begin{bmatrix} -9 & -18 & -27 \\ 0 & -45 & 0 \\ -18 & -36 & -27 \end{bmatrix} + \\ &\begin{bmatrix} 17 & 0 & 0 \\ 0 & 17 & 0 \\ 0 & 0 & 17 \end{bmatrix} \\ &= \begin{bmatrix} 15 & 6 & -15 \\ 0 & -3 & 0 \\ -10 & 0 & 4 \end{bmatrix} \end{aligned}$$

$$\text{Therefore, } A^{-1} = -\frac{1}{15} \begin{bmatrix} 15 & 6 & -15 \\ 0 & -3 & 0 \\ -10 & 0 & 4 \end{bmatrix} = \begin{bmatrix} -1 & -\frac{2}{5} & 1 \\ 0 & \frac{1}{5} & 0 \\ \frac{2}{3} & 0 & -\frac{1}{3} \end{bmatrix}$$

Check your progress

12. Find the characteristic equation of the matrix $A = \begin{bmatrix} 2 & -1 & 1 \\ -1 & 2 & -1 \\ 1 & -1 & 2 \end{bmatrix}$ and also verify that it is satisfied by A and hence, find A^{-1} .

13. Show the matrix $A = \begin{bmatrix} 1 & 2 & 0 \\ 2 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$ satisfies Cayley – Hamilton theorem. Hence or otherwise obtain the value of A^{-1} and A^{-2} .

14. Verify Cayley – Hamilton theorem for the matrix $A = \begin{bmatrix} 0 & 0 & 1 \\ 3 & 1 & 0 \\ -2 & 1 & 4 \end{bmatrix}$

15. Show that the matrix $A = \begin{bmatrix} 0 & c & -b \\ -c & 0 & a \\ b & -a & 0 \end{bmatrix}$ satisfies Cayley – Hamilton theorem.

16. Find the characteristic equation of the matrix $A = \begin{bmatrix} 0 & 1 & 2 \\ 0 & -3 & 0 \\ 1 & 1 & -1 \end{bmatrix}$ and also verify that it is satisfied by A and hence, find A^{-1} .

17. Verify Cayley – Hamilton theorem for the matrix $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & -1 & 4 \\ 3 & 1 & 1 \end{bmatrix}$

9.5 Characteristic polynomial of a linear map: Suppose

$T: V \rightarrow V$ be a linear transformation. The characteristic polynomial of T is defined to be the Characteristic polynomial of any matrix of T , such matrices are similar and having the same Characteristic polynomial.

Definition: Let $T: V \rightarrow V$ be a linear transformation on a vector space V over a field F . Let $f(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0$ be a polynomial in x over F . Now we define $f(T) = a_n T^n + a_{n-1} T^{n-1} + \dots + a_1 T + a_0 I$.

This $f(T)$ is also a linear transformation on V . This $f(T)$ is called a polynomial in T over the field F .

Theorem: Let $T: V \rightarrow V$ be a linear transformation on a vector space V over a field F . And λ is a characteristic root of T . If $f(x) \in F[x]$ then $f(\lambda)$ is a characteristic root of $f(T)$.

Proof: Since, λ is a characteristic root of T , so, there exists a non-zero $v \in V$ such that $T(v) = \lambda v$. Now $T^2(v) = T(T(v)) = T(\lambda v) = \lambda T(v) = \lambda^2 v$.

We continue this process we get $T^k(v) = \lambda^k v$ for all positive integers k .

If $f(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0 \in F[x]$.

Then $f(T) = a_n T^n + a_{n-1} T^{n-1} + \dots + a_1 T + a_0 I$

So, $f(T)(v) = (a_n T^n + a_{n-1} T^{n-1} + \dots + a_1 T + a_0 I)(v)$

$= a_n T^n(v) + a_{n-1} T^{n-1}(v) + \dots + a_1 T(v) + a_0 I(v)$

$= a_n \lambda^n(v) + a_{n-1} \lambda^{n-1}(v) + \dots + a_1 \lambda(v) + a_0 v$

$(a_n \lambda^n + a_{n-1} \lambda^{n-1} + \dots + a_1 \lambda + a_0)(v)$

$= f(\lambda)(v)$

Therefore, $f(\lambda)$ is a characteristic root of $f(T)$ and v is an eigen vector of $f(T)$ with eigen value $f(\lambda)$.

9.6 Polynomial in an n -square matrix: Let A be an n -square matrix over a field F . Let $f(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0 \in F[x]$. Then $f(A) = a_n A^n + a_{n-1} A^{n-1} + \dots + a_1 A + a_0 I_n \in F[x]$ is called a polynomial in A over F .

Note 1: If λ is a characteristic root of the matrix A . If $f(x) \in F[x]$ then $f(\lambda)$ is a characteristic root of $f(A)$.

2. Let V be a vector space over F of dimension n . Let $T: V \rightarrow V$ be a linear transformation on a vector space V over a field F , there exists a non-zero polynomial $f(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0 \in F[x]$ such that $f(T) = 0$.

9.7 Minimal polynomial: Let $T: V \rightarrow V$ be a linear transformation on a vector space V over a field F , a non-zero monic polynomial $m(x)$ is said to be the minimal polynomial of T if

(i). $m(T) = 0$

(ii). If $f(T) = 0$ for some polynomial $f(x)$, then $m(x)$ divides $f(x)$. that is $m(x)$ is the polynomial of lowest degree such that $m(A) = 0$.

Such polynomial exists and is unique.

Similarly, $m(x)$ is said to be the minimal polynomial of a square matrix A . If $m(x)$ is the polynomial of lowest degree such that $m(A) = 0$

Note: If $\lambda \in F$ is an eigen value of a linear transformation T on a vector space V , then λ is a zero the minimal polynomial $m(x)$ of T .

If A is a matrix of a linear transformation T and $\phi(x)$ be the characteristic polynomial of A , then $\phi(x)$ is also the characteristic polynomial of T .

Example: Find the minimal polynomial $m(x)$ of the matrix

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

Solution: Characteristic polynomial of A is

$$\begin{aligned} \phi(x) = \det(A - xI) &= \begin{vmatrix} 2-x & 1 & 0 \\ 0 & 2-x & 0 \\ 0 & 0 & 2-x \end{vmatrix} \\ &= (2-x)^3 \end{aligned}$$

Since, $m(x)$ and $\phi(x)$ have the same zeros, hence, $(2-x)$ must be a factor of $m(x)$ and $m(x)$ must also divide $\phi(x)$. So, $m(x)$ must be either $(2-x)$ or, $(2-x)^2$ or, $(2-x)^3$

But $m(x) \neq (2-x)$, since, $m(A) \neq 0$ so,
 $m'(x) = (2-x)^2 = \phi'(x)$. Since, $m'(A) = \phi'(A) = 0$. Hence, $m'(x) = (2-x)^2$ is the required minimal polynomial of A .

Remark: Let $M = \begin{bmatrix} A_1 & B \\ 0 & A_2 \end{bmatrix}$, where A_1, A_2 and B are square matrices. The characteristic polynomial of M is the product of the characteristic polynomials of A_1 and A_2 .

9.8 Hermitian Matrix: Let an n -square matrix $A = [a_{ij}]$ over the field C of complex numbers, then $\bar{A} = [\bar{a}_{ij}]$, where $\bar{a}_{ij}$ is the complex conjugate of a_{ij} is called the conjugate of A .

$$\text{Now } (\bar{A})^t = [\bar{a}_{ji}]^t = [\bar{a}_{ji}]. \quad \bar{A}^t = [\bar{a}_{ji}]$$

Hence, $(\bar{A})^t = \bar{A}^t = A^*$

Thus, the matrix A is said to be Hermitian if $A = A^*$. That is if $a_{ij} = \overline{a_{ji}}$.

Note: All the diagonal elements of the Hermitian matrix are real. Any real symmetric matrix is a Hermitian matrix.

Example: $A = \begin{bmatrix} 2 & 1+i & 3-i \\ 1-i & 3 & 2+3i \\ 3+i & 2-3i & 4 \end{bmatrix}$ is a Hermitian matrix of order 3.

Remark: If A and B are Hermitian matrices of the same order, then

- (i). $A + B$ is Hermitian
- (ii). AB is Hermitian if $AB = BA$
- (iii). A^t is Hermitian.
- (iv). A^{-1} is Hermitian if A is a non-singular matrix.
- (v). The characteristic roots of a complex Hermitian matrix are all real.
- (vi). Let A be any complex square matrix then $H = \bar{A}^t A$ is Hermitian
- (vii). We can define a matrix B is Hermitian congruent to A if there exists a non-singular matrix Q such that $B = \bar{Q}^t A Q$. That is the Hermitian congruence is an equivalence relation.

Example: Show that matrix $A = \begin{bmatrix} 1 & 1-i & 2 \\ 1+i & 3 & i \\ 2 & -i & 0 \end{bmatrix}$ is Hermitian Matrix.

Solution: We have $A = \begin{bmatrix} 1 & 1-i & 2 \\ 1+i & 3 & i \\ 2 & -i & 0 \end{bmatrix}$

$$\bar{A} = \begin{bmatrix} 1 & 1-i & 2 \\ 1+i & 3 & -i \\ 2 & -i & 0 \end{bmatrix}$$

$$(\bar{A})^t = \begin{bmatrix} 1 & 1-i & 2 \\ 1+i & 3 & i \\ 2 & -i & 0 \end{bmatrix}$$

$$A^t = A$$

It is Hermitian Matrix.

9.9 Summary: In this unit we have studied about the concept of Characteristic polynomial of a matrix, Cayley- Hamilton Theorem,

Characteristic polynomial of a linear map, Polynomial in an n –square matrix,
Minimal polynomial, Hermitian Matrix.

9.10 Terminal Questions

1. Show that the

$$\text{Matrix } A = \begin{bmatrix} 4 & 2-i & 5+2i \\ 2+i & 1 & 2-5i \\ 5-2i & 2+5i & 2 \end{bmatrix} \text{ is Hermitian}$$

2. What is Cayley -Hamilton theorem?

3. Find the characteristic equation of the matrix

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

4. Find the characteristic equation of the matrix $A = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 2 & 1 \\ 2 & 0 & 3 \end{bmatrix}$

5. Show that the matrix $A = \begin{bmatrix} 0 & 0 & 1 \\ 3 & 1 & 0 \\ 2 & 1 & 4 \end{bmatrix}$ satisfies Cayley-Hamilton theorem.

UNIT-10

Inner Products space

Structure

- 10.1. Introduction
- 10.2. Objectives
- 10.3. Definition and Examples of inner product space
- 10.4. Length of a vector
- 10.5. Cauchy Schwarz inequality
- 10.6. Distance between two vectors
- 10.7. Angle between two vectors
- 10.8. Orthogonal and Ortho normal sets
- 10.9. Gram- Schmidt orthogonalization process
- 10.10. Orthogonal complement of a subset.
- 10.11 Summary
- 10.12 Terminal Questions

10.1. INTRODUCTION

This unit we introduce inner product space, Length of a vector, Cauchy Schwarz inequality. We also define Distance between two vectors, Angle between two vectors and Orthogonal and Ortho normal sets.

10.2. Objectives

After studying this unit we should be able to:

- Define an Inner product space;
- Define Length of a vector
- Define Cauchy Schwarz inequality
- Define Distance between two vectors;
- Define Angle between two vectors
- Define Orthogonal and Ortho normal sets
- Gram- Schmidt orthogonalization process
- Define Orthogonal complement of a subset.

10.3. Definition and Examples of inner product space

We will now consider vector spaces V over the real field R or the complex field C , called respectively a real vector space and a complex vector space. Let us recall some properties of dot product in three dimensional real vectors.

Given the vectors $v = (x_1, x_2, x_3)$ and $w = (y_1, y_2, y_3)$ where $x_i, y_i \in R$ the dot product (v, w) was defined as

$$v \cdot w = x_1 y_1 + x_2 y_2 + x_3 y_3$$

representing the three dimensional real vector space by V , we find that the dot product can be considered as a map from $V \times V \rightarrow R$ given by if $(v, w) \in V \times V$.

$$(v, w) \rightarrow x_1 y_1 + x_2 y_2 + x_3 y_3 \in R$$

the image of (v, w) is written as $v \cdot w$.

The *length of the vector* or $|v|$ is given by $|v| = \sqrt{v \cdot v}$

And the angle θ between v and w is given by

$$\cos \theta = \frac{v \cdot w}{\sqrt{v \cdot v} \sqrt{w \cdot w}}$$

We will now generalize this concept for any real or complex vector space.

Definition : Let V be a vector space over the field F which is either the real field R or the complex field C .

A map from $V \times V \rightarrow F$ is called an inner product in V , if it satisfies the following conditions:

[Image of $(v, w) \in V \times V$ will be written as $\langle u, v \rangle$].

- (i). $\langle au_1 + bu_2, v \rangle = a\langle u_1, v \rangle + b\langle u_2, v \rangle$ Where $a, b \in F$ and $u_1, u_2 \in V$
- (ii). $\langle u, v \rangle = \overline{\langle v, u \rangle}$ (complex conjugate of $\langle u, v \rangle$)
- (iii). $\langle u, v \rangle \geq 0$ and $\langle u, u \rangle = 0$ iff $u = 0$.

Note : Condition (i) states that the map is linear in the first co-ordinate. Condition (ii) gives the *conjugate symmetry* of the map,

Since $\langle u, u \rangle = \overline{\langle u, u \rangle}$ by (ii), so $\langle u, u \rangle$ is real

Hence the inequality in (iii) makes sense. Due to the condition (iii) the inner product is called *positive definite*

$$\begin{aligned}
 \text{Note : } \langle u, av_1 + bv_2 \rangle &= \overline{\langle av_1 + bv_2, u \rangle} && \text{by (ii)} \\
 &= \overline{a\langle v_1, u \rangle + b\langle v_2, u \rangle} && \text{by (i)} \\
 &= \overline{a}\overline{\langle v_1, u \rangle} + \overline{b}\overline{\langle v_2, u \rangle} \\
 &= \overline{a}\langle u, v_1 \rangle + \overline{b}\langle u, v_2 \rangle && \text{by (ii)}
 \end{aligned}$$

Thus the map is conjugate linear in the second co-ordinate.

Note : If $F = R$, then the conjugate signs in (ii) or in the previous note may be ignored, since if $x \in R$, then $x = \bar{x}$. Thus the inner product in a real vector space is linear both in the first and second co-ordinates and condition (ii) becomes:

$\langle u, v \rangle = \langle v, u \rangle$. i.e. the inner product is symmetrical.

Definition : A vector space V with an inner product defined is called an inner product space.

When $F = R$, then V is called Euclidean space and when $F = C$, it is called Unitary space.

Example : Let C^n be the complex vector space.

If $u = (z_1, z_2, \dots, z_n)$ and $v = (w_1, w_2, \dots, w_n)$ are elements of C^n then define $\langle u, v \rangle$ as

$$\langle u, v \rangle = z_1 \overline{w_1} + z_2 \overline{w_2} + \dots + z_n \overline{w_n}.$$

This defines an inner product in C^n for if $u_1 = (a_1, a_2, \dots, a_n) \in C^n$ and $a, b \in C$, then

$$\begin{aligned}
 \langle au + bu_1, v \rangle &= (az_1 + ba_1)\overline{w_1} + (az_2 + ba_2)\overline{w_2} + \dots + (az_n + ba_n)\overline{w_n} \\
 &= a(z_1\overline{w_1} + z_2\overline{w_2} + \dots + z_n\overline{w_n}) + b(a_1\overline{w_1} + a_2\overline{w_2} + \dots + a_n\overline{w_n}) \\
 &= a\langle u, v \rangle + b\langle u_1, v \rangle
 \end{aligned}$$

Thus (i) holds.

Again $\langle v, u \rangle = w_1\overline{z_1} + w_2\overline{z_2} + \dots + w_n\overline{z_n}$ Therefore ,

$$\begin{aligned}
 \overline{\langle v, u \rangle} &= \overline{w_1\overline{z_1} + w_2\overline{z_2} + \dots + w_n\overline{z_n}} \\
 &= \langle u, v \rangle
 \end{aligned}$$

Hence (ii) holds. Now,

$$\begin{aligned}
 \langle u, u \rangle &= z_1\overline{z_1} + z_2\overline{z_2} + \dots + z_n\overline{z_n} \\
 &= |z_1|^2 + |z_2|^2 + \dots + |z_n|^2 \geq 0
 \end{aligned}$$

Since $|z_1|^2 + |z_2|^2 + \dots + |z_n|^2 = 0$ iff every $z_i = 0$

Hence $\langle u, u \rangle = 0$ if and only if $u = (0, 0, \dots, 0) = 0$. Thus (iii) satisfied.

Check your progress

1. The dot product in $\mathbb{R}^n$ is given by :

If $u = (a_1, a_2, \dots, a_n), v = (b_1, b_2, \dots, b_n) \in \mathbb{R}^n$, then

$$u \cdot v = a_1b_1 + a_2b_2 + \dots + a_nb_n$$

Then, $\langle u, v \rangle = u \cdot v$ is an inner product in $\mathbb{R}^n$. Proof is exactly similar to that of the previous example.

2. Let V be the vector space of real continuous functions on the interval $[a, b]$. Then if

$$f, g \in V, \langle f, g \rangle = \int_a^b f(t)g(t)dt$$

is an inner product in V . For if $h \in V$, and $\alpha, \beta \in \mathbb{R}$, then

$$\begin{aligned}
 \langle \alpha f + \beta h, g \rangle &= \int_a^b (\alpha f(t) + \beta h(t))g(t)dt \\
 &= \alpha \int_a^b f(t)g(t)dt + \beta \int_a^b h(t)g(t)dt \\
 &= \alpha \langle f, g \rangle + \beta \langle h, g \rangle
 \end{aligned}$$

Thus (i) is satisfied.

Since V is real vector space, we have to prove that

$\langle f, g \rangle = \langle g, f \rangle$ to prove (ii). Which is obviously true, since

$$\int_a^b f(t) g(t) dt = \int_a^b g(t) f(t) dt$$

Now $\langle f, f \rangle = \int_a^b (f(t))^2 dt \geq 0$ since, $(f(t))^2 \geq 0$.

And $\langle f, f \rangle = 0$ if and only if $\int_a^b (f(t))^2 dt = 0$ i.e. if and only if $f(t) = 0$

Thus (iii) is satisfied.

3. Let $V = \left\{ (a_1, a_2, a_3, \dots) \text{ where } a_i \in \mathbb{R} \mid \sum_{i=1}^{\infty} a_i^2 \text{ is convergent} \right\}$

Now V is the vector space over $\mathbb{R}$ with addition and scalar multiplication defined component wise.

The following is an inner product in V :

$$\langle (a_1, a_2, \dots), (b_1, b_2, \dots) \rangle = a_1 b_1 + a_2 b_2 + \dots$$

The inner product is well defined, since the sum $a_1 b_1 + a_2 b_2 + \dots$

Converges absolutely for any pair of elements in V . We can be easily verified that conditions (i),(ii),(iii) hold. This inner product space is called l_2 -space or *Hilbert space*.

Example: Let V be an inner product space. Prove

1. $\langle 0, v \rangle = 0 \forall v \in V$ and $\langle u, 0 \rangle = 0 \forall u \in V$
2. $\langle u, v \rangle = 0 \forall v \in V$ implies that $u=0$

Proof.

$$1. \langle 0, v \rangle = \langle 0u, v \rangle = 0 \langle u, v \rangle = 0 \quad \text{by (iii).}$$

$$\text{Again } \langle u, v \rangle = \langle u, 0v \rangle = \bar{0} \langle u, v \rangle = 0$$

$$2. \text{ If } \langle u, v \rangle = 0 \quad v \in V \text{ then, } \langle u, u \rangle = 0 \text{ which implies } u = 0.$$

Example: If V be a finite dimensional real or complex vector space, then an inner product can be defined in it as follows:

Let $\{e_1, e_2, \dots, e_n\}$ be a basis of V . If $u, v \in V$, then

$$u = a_1 e_1 + a_2 e_2 + \dots + a_n e_n$$

$$v = b_1 e_1 + b_2 e_2 + \cdots + b_n e_n$$

Where, a_i, b_i are uniquely determined scalars.

$$\text{Define } \langle u, v \rangle = a_1 \bar{b}_1 + a_2 \bar{b}_2 + \cdots + a_n \bar{b}_n$$

which can be easily proved to be an inner-product in V .

10.4. Length of a vector (Norm or length of a vector):

Let V be an inner product space, then the norm or the length of a vector v denoted by $\|v\|$ is defined as

$$\|v\| = \sqrt{\langle v, v \rangle} \text{ (a non-negative real number)}$$

If, $\|v\| = 1$, then v is said to be *normalized* or *unit vector*.

A non-zero vector u can be normalized by setting $v = \frac{u}{\|u\|}$ which is a unit vector. For

$$\begin{aligned} \langle v, v \rangle &= \left\langle \frac{u}{\|u\|}, \frac{u}{\|u\|} \right\rangle = \frac{1}{\|u\| \|u\|} \langle u, u \rangle \\ &= \frac{1}{\|u\|^2} \|u\|^2 = 1. \text{ Hence, } \|v\| = 1 \end{aligned}$$

Example: In the Euclidean space R^n with inner product as

defined as if $v = (x_1, x_2, \dots, x_n) \in R^n$, then

$$\|v\| = \sqrt{(x_1^2 + x_2^2 + \cdots + x_n^2)}$$

Example: In the three dimensional Euclidean space if a, b are two vectors then $\langle a, b \rangle = a \cdot b$

$$\text{So, } \|a\|^2 = \langle a, a \rangle = a \cdot a = |a|^2$$

Thus, $\|a\|^2 = |a|^2$ (magnitude of a).

Example: In the unitary space C^n with inner product defined as, if $v = (z_1, z_2, \dots, z_n) \in C^n$, then

$$\begin{aligned} \|v\| &= \sqrt{(z_1 \bar{z}_1 + z_2 \bar{z}_2 + \cdots + z_n \bar{z}_n)} \\ &= \sqrt{(|z_1|^2 + |z_2|^2 + \cdots + |z_n|^2)}. \text{ Since, } z \bar{z} = |z|^2 \end{aligned}$$

10.5. Cauchy Schwarz inequality:.

Theorem: For any two vectors u, v of an inner product space V .

$$|\langle u, v \rangle| \leq \|u\| \|v\|.$$

We will require the following Lemma for the proving of this theorem :

Lemma : If a, b, c are real numbers such that $a > 0$ and $a\lambda^2 + b\lambda + c \geq 0$ for all real numbers λ , then $b^2 \leq ac$

$$\text{Proof: } a\lambda^2 + 2b\lambda + c = \frac{1}{a}(a\lambda + b)^2 + \left(c - \frac{b^2}{a}\right) \geq 0$$

For all real λ , in particular case it is true for $\lambda = -\frac{b}{a}$

$$\text{Thus, } c - \frac{b^2}{a} \geq 0 \text{ or } ac - b^2 \geq 0 \quad [\text{since } a > 0]$$

$$\text{Or, } b^2 \leq ac.$$

Proof of Theorem: If either of u, v be zero, the inequality is satisfied, both sides being equal to zero.

So, let $u \neq 0, v \neq 0$.

First we suppose $\langle u, v \rangle$ is real.

Then for any real λ , $0 \leq \langle \lambda u + v, \lambda u + v \rangle$

$$\text{Or, } 0 \leq \lambda^2 \langle u, u \rangle + 2\lambda \langle u, v \rangle + \langle v, v \rangle$$

Putting $\langle u, u \rangle = a, \langle u, v \rangle = b, \langle v, v \rangle = c$ we get

$$0 \leq a\lambda^2 + 2a\lambda + c \text{ for all real numbers } \lambda$$

and $a > 0$, since $u \neq 0$. Hence by the above lemma $b^2 \leq ac$

$$\text{or, } \langle u, v \rangle^2 \leq \langle u, u \rangle \langle v, v \rangle$$

$$\text{or, } \langle u, v \rangle \leq \|u\| \|v\|$$

$$\text{Hence, } |\langle u, v \rangle| \leq \|u\| \|v\|.$$

Now suppose $\langle u, v \rangle = \alpha$ which is not real, then it certainly is not 0, since $u \neq 0, v \neq 0$, so that $\frac{u}{\alpha}$ is meaningful,

$$\text{Now, } \left\langle \frac{u}{\alpha}, v \right\rangle = \frac{1}{\alpha} \langle u, v \rangle = \frac{1}{\langle u, v \rangle} \langle u, v \rangle = 1$$

So $\left\langle \frac{u}{\alpha}, v \right\rangle$ is real, this has been discussed above

$$1 = \left| \left\langle \frac{u}{\alpha}, v \right\rangle \right| \leq \left\| \frac{u}{\alpha} \right\| \|v\|$$

$$1 \leq \frac{1}{|\alpha|} \|u\| \|v\|$$

$$\left(\left\| \frac{u}{\alpha} \right\| = \sqrt{\left\langle \frac{u}{\alpha}, \frac{u}{\alpha} \right\rangle} = \sqrt{\frac{1}{\alpha \bar{\alpha}} \langle u, u \rangle} = \sqrt{\frac{1}{\alpha^2} \|u\|^2} = \frac{1}{\alpha} \|u\| \right)$$

Hence, $|\alpha| \leq \|u\| \|v\|$.

So, $|\langle u, v \rangle| \leq \|u\| \|v\|$.

Another easy Proof of Cauchy-Schwartz inequality:

Suppose $u \neq 0, v \neq 0$, then

$$\begin{aligned} 0 &\leq \| (v, v)u - \langle u, v \rangle v \|^2 = \langle (v, v)u - \langle u, v \rangle v, (v, v)u - \langle u, v \rangle v \rangle \\ &= \langle \alpha u - \beta v, \alpha u - \beta v \rangle \text{ putting } \alpha = \langle v, v \rangle, \beta = \langle u, v \rangle. \\ &= \alpha \bar{\alpha} \langle u, u \rangle - \alpha \bar{\beta} \langle u, v \rangle - \beta \bar{\alpha} \langle v, u \rangle + \beta \bar{\beta} \langle v, v \rangle \\ &= |\alpha|^2 \|u\|^2 - \langle v, v \rangle \overline{\langle u, v \rangle} \langle u, v \rangle - \langle u, v \rangle \overline{\langle v, v \rangle} \langle v, u \rangle + \langle u, v \rangle \langle u, v \rangle \|v\|^2 \\ &= \|v\|^4 \|u\|^2 - \|v\|^2 \langle v, u \rangle \langle u, v \rangle - \|v\|^2 \langle v, u \rangle \langle u, v \rangle + \|v\|^2 \langle v, u \rangle \langle u, v \rangle \\ &= \|v\|^4 \|u\|^2 - \|v\|^2 \overline{\langle u, v \rangle} \langle u, v \rangle \\ &= \|v\|^4 \|u\|^2 - \|v\|^2 |\langle u, v \rangle|^2 \\ &= \|v\|^2 [\|v\|^2 \|u\|^2 - \|v\|^2 |\langle u, v \rangle|^2] \end{aligned}$$

Hence, $0 \leq \|v\|^2 \|u\|^2 - \|v\|^2 |\langle u, v \rangle|^2$

Or, $|\langle u, v \rangle|^2 \leq \|u\|^2 \|v\|^2$. Hence, $|\langle u, v \rangle| \leq \|u\| \|v\|$.

Corollary : In the previous theorem, equality holds, that is

$|\langle u, v \rangle| \leq \|u\| \|v\|$ if and only if u and v are linearly dependent.

Proof : Let $\{u, v\}$ be linearly dependent.

Now if $u = 0$ or $v = 0$, then $\{u, v\}$ is linearly dependent and then the equality holds trivially.

So, let $u \neq 0, v \neq 0$ and $v = \alpha u$ where $\alpha (\neq 0) \in F$.

Then $\langle u, v \rangle = \langle u, \alpha u \rangle = \bar{\alpha} \langle u, u \rangle = \bar{\alpha} \|u\|^2$

Hence, $|\langle u, v \rangle| \leq |\bar{\alpha}| \|v\|$

And, $\|u\| \|v\| = \|u\| \|\alpha u\| = \|u\| |\alpha| \|u\|$

$$= |\bar{\alpha}| \|u\|^2 \quad [\text{Since } |\alpha| = |\bar{\alpha}| = |\langle u, v \rangle|]$$

Now let $|\langle u, v \rangle| \leq \|u\| \|v\|$.

Let $u \neq 0, v \neq 0$ and let $w = v - \frac{\langle v, u \rangle}{\|u\|^2} u$, then

$$\begin{aligned} \langle w, w \rangle &= \langle v - \frac{\langle v, u \rangle}{\|u\|^2} u, v - \frac{\langle v, u \rangle}{\|u\|^2} u \rangle \\ &= \langle v, v \rangle - \frac{\langle v, u \rangle}{\|u\|^2} \langle v, u \rangle - \frac{\langle v, u \rangle}{\langle v, u \rangle} \langle u, v \rangle + \\ &\quad \frac{\langle v, u \rangle \overline{\langle v, u \rangle}}{\|u\|^2 \|u\|^2} \langle u, u \rangle \\ &= \|v\|^2 - \frac{\langle v, u \rangle \langle v, u \rangle}{\|u\|^2} - \frac{\langle v, u \rangle \overline{\langle v, u \rangle}}{\|u\|^2} + \frac{\langle v, u \rangle \overline{\langle v, u \rangle}}{\|u\|^2} \\ &= \|v\|^2 - \frac{|\langle v, u \rangle|^2}{\|u\|^2} = \|v\|^2 - \frac{\|u\|^2 \|v\|^2}{\|u\|^2} = 0 \end{aligned}$$

Hence, $w = 0$. So, $v - \frac{\langle v, u \rangle}{\|u\|^2} u = 0$. Therefore, $\{u, v\}$ linearly independent.

Example: In the unitary space C^n , if

$u = (z_1, z_2, \dots, z_n)$ and $v = (w_1, w_2, \dots, w_n)$, then

$$|\langle u, v \rangle| = |z_1 \bar{w}_1 + z_2 \bar{w}_2 + \dots + z_n \bar{w}_n|$$

$$\|u\| = \sqrt{|z_1|^2 + |z_2|^2 + \dots + |z_n|^2}$$

$$\|v\| = \sqrt{|w_1|^2 + |w_2|^2 + \dots + |w_n|^2}$$

Hence by Cauchy-Schwarz inequality we get

$$(1). \quad \frac{|z_1 \bar{w}_1 + z_2 \bar{w}_2 + \dots + z_n \bar{w}_n|}{\sqrt{|z_1|^2 + |z_2|^2 + \dots + |z_n|^2} \sqrt{|w_1|^2 + |w_2|^2 + \dots + |w_n|^2}} \leq 1$$

In the real Euclidean inner product space R^n , the above inequality takes the following form :

$$(2). \quad (\alpha_1 \beta_1 + \alpha_2 \beta_2 + \dots + \alpha_n \beta_n)^2 \leq (\alpha_1^2 + \alpha_2^2 + \dots + \alpha_n^2)(\beta_1^2 + \beta_2^2 + \dots + \beta_n^2) \text{ where, } \alpha_i, \beta_i \in R.$$

(This inequality was given by Cauchy in 1821.)

Example: Let f and g be real continuous functions defined on $[0, 1]$. Then

$\langle f, g \rangle = \int_0^1 f(t)g(t)dt$ is an inner product by previous example.

$$\text{Now } \|f\|^2 = \langle f, f \rangle = \int_0^1 (f(t))^2 dt$$

$$\|g\|^2 = \langle g, g \rangle = \int_0^1 (g(t))^2 dt$$

Now by Schwarz -Cauchy inequality we get

$$|\langle f, g \rangle|^2 \leq \|f\|^2 \|g\|^2$$

$$\text{Therefore, } \left| \int_0^1 f(t)g(t)dt \right|^2 \leq \int_0^1 (f(t))^2 dt \int_0^1 (g(t))^2 dt$$

Check your progress

1. The norm of a vector in an inner product space satisfies the following conditions :

- (1). $\|v\| \geq 0$ and $\|v\| = 0$ if and only if $v = 0$.
- (2). $\|kv\| = k\|v\|$ where k is any scalar.
- (3). $\|u + v\| \leq \|u\| + \|v\|$ (Triangle inequality)

Solution: (1). Follows from condition I_3 of an inner product space.

$$(2). \|kv\|^2 = \langle kv, kv \rangle = k\bar{k}\langle v, v \rangle = |k|^2\|v\|^2.$$

$$\text{Hence } \|kv\| = |k|\|v\|.$$

$$(3). \|u + v\|^2 = \langle u + v, u + v \rangle$$

$$= \langle u, u \rangle + \langle u, v \rangle + \langle v, u \rangle + \langle v, v \rangle$$

$$= \|u\|^2 + \langle u, v \rangle + \overline{\langle u, v \rangle} + \|v\|^2$$

$$= \|u\|^2 + 2R(\langle u, v \rangle) + \|v\|^2 \quad [\text{since } z + \bar{z} = 2R(z)]$$

$R(z)$ = real part of

$Z]$

$$\leq \|u\|^2 + 2|\langle u, v \rangle| + \|v\|^2 \quad [\text{since } R(z) \leq |z|]$$

$$\leq \|u\|^2 + 2\|u\|\|v\| + \|v\|^2 \quad [\text{by Cauchy-Schwarz inequality}]$$

$$= (\|u\| + \|v\|)^2$$

$$\text{Hence, } \|u + v\| \leq \|u\| + \|v\|.$$

Example: If x, y are vector in an product space V then prove that

$$\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2$$

$$\text{Hint: } \|x + y\|^2 = \langle x + y, x + y \rangle = \|x\|^2 + \langle x, y \rangle + \langle y, x \rangle + \|y\|^2$$

$$\|x - y\|^2 = \langle x - y, x - y \rangle = \|x\|^2 - \langle x, y \rangle - \langle y, x \rangle + \|y\|^2$$

10.6. Distance between two vectors:

In an inner product space V , the distance between two vectors x, y denoted by $d(x, y)$ is defined as

$$d(x, y) = \|x - y\| = \sqrt{\langle x - y, x - y \rangle}$$

Note: The distance $d(x, y)$ between vectors x, y of an inner product space V , defined above, satisfies the following conditions of a metric space :

$$(M_1) \quad d(x, y) \geq 0 \text{ and } d(x, y) = 0 \text{ if and only if } x = y$$

$$(M_2) \quad d(x, y) = d(y, x)$$

$$(M_3) \quad d(x, y) \leq d(x, z) + d(z, y)$$

Proof: (M₁) follows easily by previous concepts of this chapter

$$\begin{aligned}\text{Now } d(x, y) &= \|x - y\| = \|-(y - x)\| \\ &= |-1| \|y - x\| = \|y - x\| \\ &= d(y, x)\end{aligned}$$

Again, $x - y = (x - z) + (z - y)$

Hence, $\|x - y\| = \|(x - z) + (z - y)\| \leq \|x - z\| + \|z - y\|$

Hence $d(x, y) \leq d(x, z) + d(z, y)$

10.7. Angle between two vectors in a real inner product space:

If x, y are two non - zero vectors of a real inner product space V , the angle θ between x and y is defined by :

$$\cos \theta = \frac{\langle x, y \rangle}{\|x\| \|y\|}$$

Since by Cauchy-Schwarz inequality, $\frac{|\langle x, y \rangle|}{\|x\| \|y\|} \leq 1$

Therefore, $|\cos \theta| \leq 1$.

Hence, angle θ defined above exists.

If $\theta = \pi/2$, Then $\cos \theta = 0$

Or, $\langle x, y \rangle = 0$. Thus, x is perpendicular to y if $\langle x, y \rangle = 0$

We will now generalize this concept of orthogonality of two vectors of a real or complex inner product space.

10.8. Orthogonal and Ortho normal sets:

Definition: In a real or complex inner product space , two vectors x and y are said to be orthogonal if $\langle x, y \rangle = 0$. And we write $x \perp y$.

Note: (i). $0 \in V$ is orthogonal to *for every* $x \in V$, that is $\langle 0, x \rangle = 0$

(ii). If x is orthogonal to *for every* $y \in V$, then $x = 0$

For then $\langle x, x \rangle = 0$ and so $x = 0$.

Example: If x, y are two vectors in a real inner product space V , then $x \perp y$ if and only if $\|x - y\|^2 = \|x\|^2 + \|y\|^2$

Solution: Let $x \perp y$, then $\langle x, y \rangle = \langle y, x \rangle = 0 \dots \dots (1)$

Now, $\|x - y\|^2 = \langle x - y, x - y \rangle$

$$= \langle x, x \rangle - \langle x, y \rangle - \langle y, x \rangle + \langle y, y \rangle = \langle x, x \rangle + \langle y, y \rangle$$

$$= \|x\|^2 + \|y\|^2$$

Now, $\|x - y\|^2 = \|x\|^2 + \|y\|^2$

$$\Rightarrow \langle x - y, x - y \rangle = \langle x, x \rangle + \langle y, y \rangle$$

$$\Rightarrow \langle x, x \rangle - \langle x, y \rangle - \langle y, x \rangle + \langle y, y \rangle = \langle x, x \rangle + \langle y, y \rangle$$

$$\Rightarrow \langle x, y \rangle + \langle y, x \rangle = 0$$

$$\Rightarrow 2 \langle x, y \rangle = 0$$

$$\Rightarrow \langle x, y \rangle = 0 \text{ [Since } \langle y, x \rangle = \langle x, y \rangle \text{]}$$

$$\Rightarrow x \perp y.$$

The above result is known as Pythagorean theorem. If V is an unitary space, then the above result is not necessarily true. From the above results we get

$$\|x - y\|^2 = \|x\|^2 + \|y\|^2 - 2 \langle x, y \rangle$$

$$= \|x\|^2 + \|y\|^2 - 2\|x\| \|y\| \frac{\langle x, y \rangle}{\|x\| \|y\|}$$

$$= \|x\|^2 + \|y\|^2 - 2\|x\| \|y\| \cos \theta$$

Where θ is the angle between the vectors x and y . Thus we get the well known trigonometrical theorem.

In a triangle ABC , $AC^2 = AB^2 + BC^2 - 2AB \cdot BC \cdot \cos B$

For in the the three dimensional Euclidean space, $\overline{AB} = x$, $\overline{BC} = y$, then

$$\overline{AC} = x - y$$

$$\text{And } \|x\| = |x| = AB; \|y\| = |y| = BC, \|x - y\| = |x - y| = AC$$

10.9. Orthogonal and Ortho normal sets:

Let $S = \{v_1, v_2, \dots, v_i, \dots, v_j, \dots\}$ be a set of vectors in an inner

Product space V . Then S is said to be orthogonal if $\langle v_i, v_j \rangle = 0$

For $i \neq j$ that is, any two distinct vectors of S are orthogonal.

If further, $\langle v_i, v_j \rangle = 1 \forall v_i \in S$

That is $\|v_i\| = 1 \quad \forall v_i \in S$

Then the set S is called an orthogonal set.

Thus we can say that the set S is orthogonal if $\langle v_i, v_j \rangle = \delta_{ij}$ where $v_i, v_j \in S$

Example: The usual basis of the Euclidean n -space R^n viz

$\{e_1, e_2, \dots, e_i, \dots, e_n\}$ where, $e_i = (0, 0, \dots, 1, \dots, 0)$, i^{th} ... Coordinate is 1

and the rest entries are zeros, is ortho-normal. For

if $i \neq j$, then $\langle e_i, e_j \rangle = e_i \cdot e_j = 0$

And $\langle e_i, e_j \rangle = e_i \cdot e_j = 1$ for every i

Example: Let V be the vector space of real continuous function

And the interval $[-\pi, \pi]$ with inner product defined by :

if $f, g \in V$, then, $\langle f, g \rangle = \int_{-\pi}^{\pi} f(x)g(x)dx$.

The following is an orthogonal subset of V :

$\{1, \cos x, \cos 2x, \dots, \cos nx, \dots, \sin x, \sin 2x, \dots, \sin nx, \dots\}$

$$\text{For, } \langle \cos mx, \cos nx \rangle = \int_{-\pi}^{\pi} \cos mx \cos nx \, dx = 0 \quad (m \neq n)$$

$$\text{Again, } \langle \cos mx \sin nx \rangle = \int_{-\pi}^{\pi} \cos mx \sin nx \, dx = 0$$

Hence the given set is orthogonal.

But it is not ortho-normal since,

$$\langle \cos mx \sin nx \rangle = \int_{-\pi}^{\pi} \cos^2(mx) \, dx \neq 1$$

The above orthogonal set plays a fundamental role in the theory of

fourier series and other theories also.

Example: If the set $\{v_1, v_2, \dots, v_r\}$ is ortho-normal in an inner

Product space V , then for any $v \in V$, the vector

$$w = v - \langle v, v_1 \rangle v_1 - \langle v, v_2 \rangle v_2 - \dots - \langle v, v_r \rangle v_r$$

is orthogonal to each v_i . For,

$$\begin{aligned} \langle w, v_i \rangle &= \langle v - \langle v, v_1 \rangle v_1 - \langle v, v_2 \rangle v_2 - \dots - \langle v, v_r \rangle v_r, v_i \rangle \\ &= \langle v, v_i \rangle - \langle v, v_i \rangle \langle v_i, v_i \rangle \quad [\text{Since } \langle v_i, v_j \rangle = 0 \text{ for } i \neq j] \\ &= \langle v, v_i \rangle - \langle v, v_i \rangle \quad [\text{Since } \langle v_i, v_i \rangle = 1] \\ &= 0. \text{ Thus } w \text{ is orthogonal to each } v_i. \end{aligned}$$

Check your progress

1. Any orthogonal Set S of non-zero vectors in an inner product space V is linearly independent.

Solution: Let $\{v_1, v_2, \dots, v_m\}$ be a finite subset of S .

$$(1) \text{ Let } a_1 v_1 + a_2 v_2 + \dots + a_k v_k + \dots + a_m v_m = 0 \text{ where } a_i \in F$$

Now for $k = 1, 2, \dots, m$ we have

$$\begin{aligned} \langle a_1 v_1 + a_2 v_2 + \dots + a_k v_k + \dots + a_m v_m, v_k \rangle &= a_k \|v_k\|^2 \\ [\text{Since } \langle v_i, v_k \rangle &= 0 \text{ for } i \neq k] \end{aligned}$$

$$\text{Therefore, } \langle 0, v_k \rangle = a_k \|v_k\|^2 \quad [\text{From (1)}]$$

$$\text{Or, } 0 = a_k \|v_k\|^2. \text{ Hence } a_k = 0, \text{ since } v_k \neq 0$$

$$\text{Thus, } a_1 v_1 + \dots + a_k v_k + \dots + a_m v_m = 0 \Rightarrow a_k = 0 \text{ for } k = 1, 2, \dots, m.$$

Hence, the set $\{v_1, v_2, \dots, v_m\}$ is linearly independent. Thus every finite subset of S is L.I. and so, S is linearly independent.

2. Let V be an inner product space and let $S = \{v_1, v_2, \dots, v_m\}$ be an orthogonal set of non-zero vectors in V . Then any vector $v \in S_j$ (that a subspace of V generated by S), can be expressed as:

$$v = \frac{\langle v, v_1 \rangle}{\|v_1\|^2} v_1 + \frac{\langle v, v_2 \rangle}{\|v_2\|^2} v_2 + \dots + \frac{\langle v, v_m \rangle}{\|v_m\|^2} v_m.$$

Solution: Now $v \in [S]$ can be expressed as

$$v = a_1 v_1 + a_2 v_2 + \dots + a_m v_m \text{ where } a_i \in F.$$

Now for $i = 1, 2, \dots, m$ we have

$$\langle v, v_i \rangle = \langle a_1 v_1 + a_2 v_2 + \dots + a_i v_i + \dots + a_m v_m, v_i \rangle$$

$$= a_i \langle v_i, v_i \rangle \quad [\text{Since } \langle v_i, v_j \rangle = 0 \text{ for } i \neq j]$$

$$= a_i \|v_i\|^2$$

$$\text{Hence, } a_i = \frac{\langle v, v_i \rangle}{\|v_i\|^2} \text{ for } i = 1, 2, \dots, m$$

$$\text{Therefore, } v = \frac{\langle v, v_1 \rangle}{\|v_1\|^2} v_1 + \dots + \frac{\langle v, v_m \rangle}{\|v_m\|^2} v_m.$$

10.10. Gram- Schmidt orthogonalization process:

Every finite- dimensional inner product space V has an orthogonal basis.

Proof: We know that every finite-dimensional vector space V has a basis.

Let $\{v_1, v_2, \dots, v_n\}$ be a basis of V . We will now construct an orthogonal basis $\{u_1, u_2, \dots, u_n\}$.

We put $u_1 = \frac{v_1}{\|v_1\|}$ which is a unit vector.

Let $w_2 = v_2 - \langle v_2, u_1 \rangle u_1$ [$w_2 \neq 0$, since u_1, v_2 are linearly independent]

Put $u_2 = \frac{w_2}{\|w_2\|}$ then u_2 is a unit vector.

$$\text{Now } \langle w_2, u_1 \rangle = \langle v_2 - \langle v_2, u_1 \rangle u_1, u_1 \rangle$$

$$= \langle v_2, u_1 \rangle - \langle v_2, u_1 \rangle \|u_1\|^2 = 0.$$

Hence $w_2 \perp u_1$, so $u_2 \perp u_1$.

Consequently, the set $\{u_1, u_2\}$ is orthonormal.

Proceeding thus, in general after getting the orthonormal set $\{u_1, u_2, \dots, u_i\}$ we put

$$w_{i+1} = v_{i+1} - \langle v_{i+1}, u_1 \rangle u_1 - \langle v_{i+1}, u_2 \rangle u_2 - \dots - \langle v_{i+1}, u_i \rangle u_i$$

$$\text{And } u_{i+1} = \frac{w_{i+1}}{\|w_{i+1}\|}.$$

It is clear that $w_{i+1} \neq 0$, since v_{i+1} can not be expressed as a linear combination of $v_1, v_2, \dots, v_i$.

and hence v_{i+1} can not be expressed as a linear combination of $u_1, u_2, \dots, u_i$.

Thus, the set $\{u_1, u_2, \dots, u_{i+1}\}$ is an orthogonal set of non- zero vectors.

Now by induction we get an orthogonal set $\{u_1, u_2, \dots, u_n\}$ which is linearly independent and hence is a basis of V . Further we get

$$u_1 = \frac{1}{\|v_1\|} v_1 = a_{11} v_1,$$

$$u_2 = \frac{1}{\|w_2\|} \{v_2 - \frac{\langle v_2, u_1 \rangle}{\|v_1\|} v_1\} = a_{21} v_1 + a_{22} v_2$$

.....

.....

$$u_i = a_{i1} v_1 + a_{i2} v_2 + \dots + a_{in} v_n$$

Thus the matrix of $\{u_i\}$ with respect to $\{v_i\}$ is the triangular matrix.

$$\begin{bmatrix} a_{11} & \cdots & a_{n1} \\ \vdots & \ddots & \vdots \\ 0 & \cdots & a_{nn} \end{bmatrix}$$

Example: Transform the following basis of Euclidean space $\mathbb{R}^3$ to orthogonal basis: $S = \{v_1 = (1,1,1), v_2 = (0,1,1), v_3 = (0,0,1)\}$

Solution: Put $u_1 = \frac{v_1}{\|v_1\|} = \frac{(1,1,1)}{\sqrt{1^2+1^2+1^2}} = (\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}})$

Let $w_2 = v_2 - \langle v_2, u_1 \rangle u_1$

$$\begin{aligned} &= (0,1,1) - (0 \cdot \frac{1}{\sqrt{3}} + 1 \cdot \frac{1}{\sqrt{3}} + 1 \cdot \frac{1}{\sqrt{3}}) (\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}) \\ &= (0,1,1) - \frac{2}{\sqrt{3}} (\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}) = (-\frac{2}{3}, \frac{1}{3}, \frac{1}{3}) \end{aligned}$$

Put $u_2 = \frac{w_2}{\|w_2\|} = \frac{1}{\sqrt{\frac{4}{9} + \frac{1}{9} + \frac{1}{9}}} (-\frac{2}{3}, \frac{1}{3}, \frac{1}{3}) = (-\frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}})$

Let $w_3 = v_3 - \langle v_3, u_1 \rangle u_1 - \langle v_3, u_2 \rangle u_2$

$$\begin{aligned} &= (0,0,1) - (0 \cdot \frac{1}{\sqrt{3}} + 1 \cdot \frac{1}{\sqrt{3}} + 1 \cdot \frac{1}{\sqrt{3}}) (\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}) - \frac{1}{\sqrt{6}} (-\frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}) \\ &= (0,0,1) - (\frac{1}{3}, \frac{1}{3}, \frac{1}{3}) - (-\frac{2}{6}, \frac{1}{6}, \frac{1}{6}) = (0, -\frac{1}{2}, \frac{1}{2}) \end{aligned}$$

Put $u_3 = \frac{w_3}{\|w_3\|} = \frac{1}{\sqrt{\frac{1}{4} + \frac{1}{4}}} (0, -\frac{1}{2}, \frac{1}{2}) = (0, -\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$

Hence the required ortho-normal basis of V is

$$\left\{ u_1 = \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}} \right), u_2 = \left(-\frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}} \right), u_3 = \left(0, -\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right) \right\}$$

Check your progress

1. Find an ortho-normal basis of the subspace W of C^2 generated by:

$$\{v_1 = (1, i, 0), v_2 = (1, 2, 1 - i)\}.$$

$$\text{Solution: Put } u_1 = \frac{v_1}{\|v_1\|} = \frac{(1, i, 0)}{\sqrt{(1, 1 + i(-i) + 0, 0)}} = \left(\frac{1}{\sqrt{2}}, \frac{i}{\sqrt{2}}, 0\right)$$

$$\text{Let } w_2 = v_2 - \langle v_2, u_1 \rangle u_1 = (1, 2, 1 - i) - \left(1 \cdot \frac{1}{\sqrt{2}} + 2 \cdot \frac{i}{\sqrt{2}}\right) \left(\frac{1}{\sqrt{2}}, \frac{i}{\sqrt{2}}, 0\right)$$

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

$$\text{Hence, } \|w_2\|^2 = \left(\frac{1+2i}{2}\right)\left(\frac{1-2i}{2}\right) + \left(\frac{2-i}{2}\right)\left(\frac{2+i}{2}\right) + (1-i)(1+i)$$

$$= \frac{5}{4} + \frac{5}{4} + 2 = \frac{18}{4}.$$

$$\text{Put, } u_2 = \frac{w_2}{\|w_2\|} = \frac{2}{\sqrt{18}} \left(\frac{1+2i}{2}, \frac{2-i}{2}, 1-i\right) = \left(\frac{1+2i}{\sqrt{18}}, \frac{2-i}{\sqrt{18}}, \frac{2-i}{\sqrt{18}}\right)$$

Hence, the required ortho-normal basis of W is

$$\{u_1 = \left(\frac{1}{\sqrt{2}}, \frac{i}{\sqrt{2}}, 0\right), u_2 = \left(\frac{1+2i}{\sqrt{18}}, \frac{2-i}{\sqrt{18}}, \frac{2-i}{\sqrt{18}}\right)\}.$$

2. Let V be the set of polynomials in x over R of degree 2 or less. In V we define an inner product by: if $p(x), q(x) \in V$, then

$$\langle p(x), q(x) \rangle = \int_{-1}^1 p(x)q(x)dx. \text{ Find an ortho-normal basis of } V.$$

Solution: Let us start with the basis $S = \{v_1 = 1, v_2 = x, v_3 = x^2\}$

$$\text{Put } u_1 = \frac{v_1}{\|v_1\|} = 1 / \sqrt{\int_{-1}^1 dx} = \frac{1}{\sqrt{2}}$$

$$\text{Let } w_2 = v_2 - \langle v_2, u_1 \rangle u_1 = x - \left(\int_{-1}^1 x \cdot \frac{1}{\sqrt{2}} dx\right) \frac{1}{\sqrt{2}} = x$$

$$\text{Put } u_2 = \frac{w_2}{\|w_2\|} = \frac{x}{\left(\int_{-1}^1 x^2 dx\right)^{\frac{1}{2}}} = \frac{\sqrt{3}}{\sqrt{2}} x$$

$$\text{Let } w_3 = v_3 - \langle v_3, u_1 \rangle u_1 - \langle v_3, u_2 \rangle u_2$$

$$= x^2 - \left(\int_{-1}^1 x^2 \cdot \frac{1}{\sqrt{2}} dx\right) \frac{1}{\sqrt{2}} - \left(\int_{-1}^1 x^2 \cdot \frac{\sqrt{3}}{\sqrt{2}} x dx\right) \frac{\sqrt{3}}{\sqrt{2}} x = x^2 - \frac{1}{3} - 0$$

$$\text{Put } u_3 = \frac{w_3}{\|w_3\|} = \frac{x^2 - \frac{1}{3}}{\left(\int_{-1}^1 \left(x^2 - \frac{1}{3}\right)^2 dx\right)^{\frac{1}{2}}} = \frac{\sqrt{10}}{4} (3x^2 - 1)$$

Hence, the required ortho-normal basis is

$$\left\{ u_1 = \left(\frac{1}{\sqrt{2}} \right), u_2 = \left(\frac{\sqrt{3}}{\sqrt{2}} x \right), u_3 = \frac{\sqrt{10}}{4} (3x^2 - 1) \right\}.$$

3. Let $\{e_1, e_2, \dots, e_n\}$ be an orthonormal basis of an inner product space V . Prove that:

$$(i) \text{ if } v \in V, \text{ then } v = \sum_{i=1}^n \langle v, e_i \rangle e_i$$

$$(ii) \left\langle \sum_{i=1}^n a_i e_i, \sum_{j=1}^n b_j e_j \right\rangle = \sum_{i=1}^n a_i \bar{b}_i \quad \text{where } a_i, b_j \in F$$

$$(iii) \text{ If } u, v \in V, \text{ then, } \langle u, v \rangle = \sum_{i=1}^n \langle u, e_i \rangle \overline{\langle v, e_i \rangle}$$

Solution:

$$(i) \text{ Let } v = \sum_{i=1}^n a_i e_i, \text{ Now for } i = 1, 2, \dots, n$$

$$\langle v, e_i \rangle = \left\langle \sum_{i=1}^n a_i e_i, e_i \right\rangle = a_i \langle e_i, e_i \rangle = a_i$$

$$\text{Hence, } v = \sum_{i=1}^n \langle v, e_i \rangle e_i$$

$$(ii) \left\langle \sum_i^n a_i e_i, \sum_j^n b_j e_j \right\rangle = \sum_{i,j} a_i \bar{b}_j \langle e_i, e_j \rangle = \sum_{i,j} a_i \bar{b}_j \delta_{ij} = \sum a_i \bar{b}_i$$

$$(iii). \text{ By (i) } u = \sum_{i=1}^n \langle u, e_i \rangle e_i, \quad v = \sum_{i=1}^n \langle v, e_i \rangle e_i$$

$$\begin{aligned} \text{Hence, } \langle u, v \rangle &= \left\langle \sum_i \langle u, e_i \rangle e_i, \sum_j \langle v, e_j \rangle e_j \right\rangle \\ &= \sum_i \langle u, e_i \rangle \overline{\langle v, e_i \rangle} \quad \text{by (ii)} \end{aligned}$$

x. Orthogonal complement of a subset

Let V be an inner product space and W a subset of V .

The orthogonal complement of W , denoted by $W^\perp$ (read 'W perp') is the set of those vectors of V which are orthogonal to *for all* $w \in W$. Thus

$$W^\perp = \{v \in V \mid \langle v, w \rangle = 0 \forall w \in W\}$$

Note: If W be a subset of an inner product space V , then $W^\perp$ is a subspace of V .

Clearly, $0 \in W^\perp$. Now let $u, v \in W^\perp$ and $a, b \in F$, and let $w \in W$, then

$$\begin{aligned} \langle au + bv, w \rangle &= a \langle u, w \rangle + b \langle v, w \rangle \\ &= a0 + b0 \text{ [since } v \perp w, u \perp w \text{]} \\ &= 0 \end{aligned}$$

Thus $au + bv$ is orthogonal to w for all $w \in W$.

So, $au + bv \in W^\perp$ which is therefore a subspace of V .

Theorem: If W be a subspace of a finite dimensional inner product space V , then V is the direct sum of W and $W^\perp$. That is $V = W \oplus W^\perp$

Proof: W must be finite dimensional and hence possesses an ortho-normal basis. Let $\{w_1, w_2, \dots, w_m\}$ be an orthonormal basis of W . Let $v \in V$.

$$\text{Let } w^* = v - \sum_{i=1}^m \langle v, w_i \rangle w_i$$

By the previous example, w^* is orthogonal to each w_i for $i = 1, 2, \dots, m$.

Let w^* is orthogonal to every $w \in W$ and therefore, $w^* \in W^\perp$.

$$\text{Again, } \sum_{i=1}^m \langle v, w_i \rangle w_i = w \in W.$$

Thus, $v = w + w^*$ where $w \in W$ and $w^* \in W^\perp$

Hence, $V = W + W^\perp$

Now let $v \in W \cap W^\perp$, then $v \in W$ and $v \in W^\perp$ and so, $\langle v, v \rangle = 0$

Hence, $v = 0$. So, $W \cap W^\perp = \{0\}$(2)

From (1) and (2) we get $V = W \oplus W^\perp$.

Corollary: If W be a subspace of a finite dimensional inner product space V , then $W = W^{\perp\perp}$ (orthogonal component of $W^\perp$)

Proof: $w \in W \Rightarrow \langle w, v \rangle = 0 \forall v \in W^\perp$.

$\Rightarrow w$ is orthogonal to every $v \in W^\perp \Rightarrow w \in W^{\perp\perp}$

(1) Hence, $W \subseteq W^{\perp\perp}$.

By Previous theorem we have $V = W \oplus W^\perp$ and also $V = W^{\perp\perp} \oplus W^\perp$

By Previous Corollary we have

$$\dim V = \dim W + \dim W^\perp = \dim W^\perp + \dim W^{\perp\perp}$$

Which gives that $\dim W = \dim W^{\perp\perp}$

By (1) $W \subseteq W^{\perp\perp}$. Hence, $W = W^{\perp\perp}$

Bessel's Inequality

Theorem: Let $v_1, v_2, \dots, v_k, \dots, v_n$ be an orthogonal set of non-zero vectors in an inner product space V . If u is any vector then $\sum_{k=1}^n \frac{|\langle u, v_k \rangle|^2}{\|v_k\|^2} \leq \|u\|^2$

Proof: Let $W = [\{v_1, v_2, \dots, v_k, \dots, v_n\}]$ be the sub space of V generated by $\{v_1, v_2, \dots, v_n\}$ and $W^\perp$ be the orthogonal complement of W , then $V = W \oplus W^\perp$
Hence if $u \in V$, then $u = w + \delta$
where $w \in W$ and $\delta \in W^\perp$

(1) So, $\langle w, \delta \rangle = 0$. Now $\langle u, u \rangle =$

$$\begin{aligned} \langle w + \delta, w + \delta \rangle &= \langle w, w \rangle + \langle w, \delta \rangle + \langle \delta, w \rangle + \langle \delta, \delta \rangle \\ &= \langle w, w \rangle + \langle \delta, \delta \rangle \quad [\text{Since } \langle w, \delta \rangle = 0] \end{aligned}$$

$$(2) \text{ Hence, } \|u\|^2 = \|w\|^2 + \|\delta\|^2$$

Let $\gamma = \sum_{k=1}^n \frac{\langle u, v_k \rangle}{\|v_k\|^2} v_k$. Then, $\gamma \in W$, hence

$$(3) \|u\|^2 = \|\gamma\|^2 + \|\delta\|^2$$

Where $u = \gamma + \delta$ and $\delta \in W^\perp$

$$\begin{aligned} \text{Now } \langle \gamma, \gamma \rangle &= \left\langle \frac{\langle u, v_1 \rangle}{\|v_1\|^2} v_1 + \frac{\langle u, v_2 \rangle}{\|v_2\|^2} v_2 + \dots \dots \dots + \right. \\ &\quad \left. \frac{\langle u, v_n \rangle}{\|v_n\|^2} v_n, \frac{\langle u, v_1 \rangle}{\|v_1\|^2} v_1 + \dots \dots \dots \frac{\langle u, v_n \rangle}{\|v_n\|^2} v_n \right\rangle \end{aligned}$$

$$\text{or, } \|\gamma\|^2 = \frac{\langle u, v_1 \rangle}{\|v_1\|^2} \cdot \frac{\langle u, v_1 \rangle}{\|v_1\|^2} \cdot \|v_1\|^2 + \dots \dots \dots +$$

$$\frac{\langle u, v_n \rangle}{\|v_n\|^2} \cdot \frac{\langle u, v_n \rangle}{\|v_n\|^2} \|v_n\|^2 = \frac{|\langle u, v_1 \rangle|^2}{\|v_1\|^4} \|v_1\|^2 + \dots \dots \dots + \frac{|\langle u, v_n \rangle|^2}{\|v_n\|^4} \|v_n\|^2$$

$$\text{Or, } \|\gamma\|^2 = \frac{|\langle u, v_1 \rangle|^2}{\|v_1\|^4} \cdot \|v_1\|^2 + \dots \dots \dots + \frac{|\langle u, v_n \rangle|^2}{\|v_n\|^4} \cdot \|v_n\|^2$$

$$\text{But } \|\gamma\|^2 \leq \|u\|^2 \quad [\text{By (3)}]$$

Hence, $\sum_{k=1}^n \frac{|\langle u, v_k \rangle|^2}{\|v_k\|^2} \leq \|u\|^2$.

Corollary: Let $\{v_1, v_2, \dots, v_n\}$ be an orthonormal subset of an inner product space V . If u is any vector in V , then $\sum_{k=1}^n |\langle u, v_k \rangle|^2 \leq \|u\|^2$

Note: This Corollary is another form of Bessel's Inequality.

Theorem: An ortho-normal subset $\{v_1, v_2, \dots, v_n\}$ of an inner product space V is an ortho-normal basis of V if and only if

$$\|u\|^2 = \sum_{k=1}^n |\langle u, v_k \rangle|^2 \text{ for every } u \in V$$

Proof: Let $\{v_1, v_2, \dots, v_n\}$ be an orthonormal basis of V .

$\therefore$ if $u \in V$. Then $u = \sum_{k=1}^n \langle u, v_k \rangle v_k$

$$\text{Or, } \sum_{k=1}^n a_k v_k \quad \dots (1)$$

Where, $a_k = \langle u, v_k \rangle$, $k = 1, 2, \dots, n$ (2)

From (1), $\|u\|^2 = \langle u, u \rangle = a_1 \bar{a}_1 + a_2 \bar{a}_2 + \dots + a_n \bar{a}_n$

$$\text{Or, } \|u\|^2 = |a_1|^2 + |a_2|^2 + \dots + |a_n|^2$$

$$\text{Or, } \|u\|^2 = |\langle u, v_1 \rangle|^2 + |\langle u, v_2 \rangle|^2 + \dots + |\langle u, v_n \rangle|^2$$

$$\therefore \|u\|^2 = \sum_{k=1}^n |\langle u, v_k \rangle|^2$$

Converse, Let $\|u\|^2 = \sum_{k=1}^n |\langle u, v_k \rangle|^2$, $\forall u \in V$.

Since, $\{v_1, v_2, \dots, v_n\}$ is an orthonormal set so it is linearly independent. In order to prove $\{v_1, v_2, \dots, v_n\}$ is a basis of V , we have to show that every $u \in V$ is a linear combination of elements of $\{v_1, v_2, \dots, v_n\}$.

$$\text{Let } z = u - \sum_{k=1}^n \langle u, v_k \rangle v_k$$

$$= u - \sum_{k=1}^n a_k v_k \quad \text{when } a_k = \langle u, v_k \rangle.$$

$$\text{Therefore, } \|z\|^2 = \langle u - \sum_{k=1}^n a_k v_k, u - \sum_{k=1}^n a_k v_k \rangle$$

$$= \langle u, u \rangle - \langle u, \sum_{k=1}^n a_k v_k \rangle - \sum_{k=1}^n a_k \langle v_k, u \rangle + \sum_{k=1}^n a_k \bar{a}_k \langle v_k, v_k \rangle$$

$$= \langle u, u \rangle - \sum_{k=1}^n \bar{a}_k a_k - \sum_{k=1}^n a_k \bar{a}_k + \sum_{k=1}^n |a_k|^2$$

$$= \|u\|^2 - \sum_{k=1}^n |a_k|^2$$

$$= \|u\|^2 - \sum_{k=1}^n |\langle u, v_k \rangle|^2$$

$$\therefore \|z\|^2 = 0 \Rightarrow z = 0$$

$$\Rightarrow u = \sum_{k=1}^n \langle u, v_k \rangle v_k = a_1 v_1 + a_2 v_2 + \cdots + a_n v_n$$

$\Rightarrow u$ is a linear combination of $v_1, v_2, \dots, v_n$.

Check your progress

1. If V be an inner product space and S, S_1, S_2 are subsets of V , then

- (i). $S_1 \subseteq S_2 \Rightarrow S_2^\perp \subseteq S_1^\perp$
- (ii). $S^\perp = [S]^\perp$
- (iii). $[S] \subseteq S^{\perp\perp}$
- (iv). $[S] = S^{\perp\perp}$ if V is finite dimensional

Solution.

- (i). Let $S_1 \subseteq S_2$
 $v \in S_2^\perp \Rightarrow v \perp x \forall x \in S_2 \Rightarrow v \perp x \forall x \in S_1$ [since $S_1 \subseteq S_2$]
 $\Rightarrow v \in S_1^\perp$ Hence, $S_2^\perp \subseteq S_1^\perp$. Since, $S \subseteq [S]$,
 so by (i) $[S]^\perp \subseteq S^\perp$. Now let $v \in S^\perp$. Then $\langle u, v \rangle = 0 \forall u \in S$
- (ii). If $w \in [S]$, then $w = \sum a_i u_i$ where $a_i \in F, u_i \in S$. We have
 $\langle v, w \rangle = \langle v, \sum a_i u_i \rangle = \sum \bar{a}_i \langle v, u_i \rangle = 0$ [since, by (2) $\langle v, u_i \rangle = 0 \forall u_i \in S$]

Thus, $v \perp w \forall w \in [S]$. That $v \in [S]^\perp$. Hence, $v \in S^\perp \Rightarrow v \in [S]^\perp$, so,
 $S^\perp \subseteq [S]^\perp$

From (1) and (2) we get $S^\perp = [S]^\perp$

- (iii). Let $v \in [S]$, then v is a linear combination of finite number of elements of S .
 If $w \in S^\perp$, then w is orthogonal to every vector in S and so w is orthogonal to v . Thus v is orthogonal to $\forall w \in S^\perp$.
 Hence, $v \in S^{\perp\perp}$
 Thus, $v \in [S] \Rightarrow v \in S^{\perp\perp}$
 Therefore, $[S] \subseteq S^{\perp\perp}$
- (iv). From (ii) we get $S^\perp \subseteq [S]^\perp$
 Therefore $[S]^{\perp\perp} = [S]^\perp{}^\perp = [S]$,
 Since, $W = [S]$ is a subspace of V and by previous Corollary
 $W = W^{\perp\perp}$ if V is finite dimensional.

2. A real matrix is orthogonal if and only if its rows (columns) form orthogonal set of elements of R^n

10.11 Summary: In this unit we have studied about the concept of an Inner product space, length of a vector, Cauchy Schwarz inequality, we

introduce distance between two vectors, angle between two vectors and Orthogonal and Ortho normal sets

10.12 Terminal questions

Q.1- Consider the inner product on $\mathbb{R}^3$ as

$$\langle \vec{x}, \vec{y} \rangle = x_1 y_1 + x_2 y_2 + 2x_3 y_3 + x_2 y_3 + x_3 y_2$$

$$\text{where } \vec{x} = (x_1, x_2, x_3) \text{ \& } \vec{y} = (y_1, y_2, y_3).$$

Show that it is an inner product space.

Q.2 – Find the orthonormal basis of the subspace generated by the sub set

$$\{(1,1,1), (0,1,1), (2,1,1)\} \text{ in the inner product space .}$$

Q.3- Show that the orthogonal complement of the vector $(1,1,-1)$ is the plane $x + y - z = 0$ which is a two dimensional subspace of $\mathbb{R}^3$.

Q.4- Consider the usual inner product space $\mathbb{C}^3$. Find an ortho normal basis for the subspace spanned by vectors $(1,0,i), (0,1,i)$

UNIT-11

Bilinear, Quadratic & Hermitian Forms

Structure

11.1.1. Introduction

11.1.2. Objectives

11.1.3. Bilinear form on a vector space

11.1.4. Matrix of a bilinear form on V

11.1.5. Congruent matrix

11.1.6. Symmetric, skew- symmetric and alternating bilinear forms

11.1.7. Quadratic forms

11.1.8. Matrix of quadratic form

11.1.9. Normal (canonical form)

11.1.10. Complex hermitian form

11.11 Summary

11.12 Terminal Questions

11.1.1. INTRODUCTION

This unit introduces the concept of Bilinear form on a vector space, Matrix of a bilinear form on V , Congruent matrix, Symmetric, skew- symmetric and alternating bilinear forms, Quadratic forms, Matrix of quadratic form, Normal (canonical form) Also Complex hermitian form.

11.1.2. Objectives

After studying this unit you should be able to

- Define the Bilinear form on a vector space
- Matrix of a bilinear form on V ;
- Congruent matrix
- Symmetric, skew- symmetric and alternating bilinear forms

- Quadratic forms
 - Matrix of quadratic form
 - Normal (canonical form)
 - Complex hermitian form

11.1.3. Bilinear form on a vector space.

Definition: Let V be a vector space over a field F . A bilinear form on V is a map $f: V \times V \rightarrow F$ satisfying the following conditions:

- (i) $f(\alpha_1 u_1 + \alpha_2 u_2, v) = \alpha_1 f(u_1, v) + \alpha_2 f(u_2, v)$
 (ii) $f(u, \alpha_1 v_1 + \alpha_2 v_2) = \alpha_1 f(u, v_1) + \alpha_2 f(u, v_2), \alpha_1, \alpha_2 \in F$ and $u, u_1, u_2, v, v_1, v_2 \in V$

Note: These two conditions (i) and (ii) in the above definition express that f is linear both in the first and second variables and can be combined into the single condition:

$$f(\alpha_1 u_1 + \alpha_2 u_2, \beta_1 v_1 + \beta_2 v_2) = \alpha_1 \beta_1 f(u_1, v_1) + \alpha_1 \beta_2 f(u_1, v_2) + \alpha_2 \beta_1 f(u_2, v_1) + \alpha_2 \beta_2 f(u_2, v_2)$$

Where $\alpha_1, \alpha_2 \in F, \beta_1, \beta_2 \in F$ and $u_1, u_2, v_1, v_2 \in V$

Example1: let $f: R^n \times R^n \rightarrow R$ be defined as follows:

If $x = (\alpha_1, \alpha_2, \dots, \alpha_n)$ and $y = (\beta_1, \beta_2, \dots, \beta_n)$ in R^n , then

$$f(x, y) = \alpha_1 \beta_1 + \alpha_2 \beta_2 + \alpha_3 \beta_3 + \dots + \alpha_n \beta_n.$$

Then prove that f is a bilinear form on R^n

Solution: Let $a, b \in R$, and $z = (\gamma_1, \gamma_2, \dots, \gamma_n) \in R^n$ then

$$\begin{aligned} f(ax + bz, y) &= f((a\alpha_1 + b\gamma_1, a\alpha_2 + b\gamma_2, \dots, a\alpha_n + b\gamma_n), (\beta_1, \beta_2, \dots, \beta_n)) \\ &= (a\alpha_1 + b\gamma_1)\beta_1 + (a\alpha_2 + b\gamma_2)\beta_2 + \dots + (a\alpha_n + b\gamma_n)\beta_n \\ &= a(\alpha_1\beta_1 + \alpha_2\beta_2 + \alpha_3\beta_3 + \dots + \alpha_n\beta_n) \\ &\quad + b(\gamma_1\beta_1 + \gamma_2\beta_2 + \gamma_3\beta_3 + \dots + \gamma_n\beta_n) = af(x, y) + bf(z, y) \end{aligned}$$

Thus f is linear in the first variable.

Similarly it can be proved that f is also linear in the second variable.

Example2: Let $A = (a_{ij})$ be an n -square matrix over the field F . We define $f: F^n \times F^n \rightarrow F$ as follows:

If $X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ and $Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \in F^n$, then

$$f(X, Y) = X^t A Y$$

$$= (x_1, x_2, \dots, x_n) \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \sum_{i,j=1}^n a_{ij} x_i y_j$$

$$= a_{11}x_1y_1 + a_{12}x_1y_2 + \dots + a_{nn}x_ny_n$$

Prove that f is a bilinear form on F^n .

Solution: Let $Z = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix} \in F^n$ and $a, b \in F$, then

$$f(aX + bZ, Y) = (aX + bZ)^t A Y$$

$$= (aX^t + bZ^t) A Y = a(X^t A Y) + b(Z^t A Y)$$

$$= af(X, Y) + bf(Z, Y)$$

Thus f is linear in the first variable. Similarly f is also linear in the second variable.

Example3: Let ϕ_1, ϕ_2 be linear functionals on a vector space V over a field F . Define $f: V \times V \rightarrow F$ by

$$f(x, y) = \phi_1(x)\phi_2(y), \text{ where } x, y \in V,$$

Prove that f is a bilinear form on V

Solution: Let $x, y, z \in V$ and $a, b \in F$, then

$$f(ax + bz, y) = \phi_1(ax + bz)\phi_2(y) = [a\phi_1(x) + b\phi_1(z)]\phi_2(y)$$

$= a\phi_1(x)\phi_2(y) + b\phi_1(z)\phi_2(y) = af(x, y) + bf(z, y)$. Thus f is linear in the first variable. Similarly f is linear in the second variable.

Theorem: The set $B(V)$ of all bilinear forms on $V(F)$ is a vector space over F with respect to addition and scalar multiplication defined as:

$$(1) \quad (f + g)(u, v) = f(u, v) + g(u, v).$$

$$(2) \quad (kf)(u, v) = kf(u, v)$$

for any $f, g \in B(V)$ and $k \in F$

Conditions (1) and (2) above can be combined in a single condition:

$$(\alpha f + \beta g)(u, v) = \alpha f(u, v) + \beta g(u, v), \text{ where } \alpha, \beta \in F$$

Proof: Can do as an exercise.

Theorem: Let V be a vector space of dimension n over F . Let $S = \{x_1, x_2, \dots, x_n\}$ be a basis of V . Then a map $f: S \times S \rightarrow F$ given by

$f(x_i, x_j) = a_{ij}$ where $x_i, x_j \in S$ and $a_{ij} \in F$ can be extended to a unique bilinear form on V whose restriction on $S \times S$ is f .

Proof: Let $x = \sum_{i=1}^n a_i x_i$ $y = \sum_{j=1}^n b_j x_j \in V$

Define $\phi: V \times V \rightarrow F$ by $\phi(x, y) = \sum_{i,j=1}^n a_i b_j a_{ij}$

Let $z = \sum_{i=1}^n c_i x_i \in V$ and $\alpha, \beta \in F$

Then $\phi(\alpha x + \beta z, y) = \phi(\alpha(\sum_{i=1}^n a_i x_i) + \beta(\sum_{i=1}^n c_i x_i), \sum_{j=1}^n b_j x_j)$

$$= \phi\left(\left(\sum_{i=1}^n (\alpha a_i + \beta c_i) x_i\right), \sum_{j=1}^n b_j x_j\right) = \sum_{i,j=1}^n (\alpha a_i + \beta c_i) b_j a_{ij}$$

$$= \alpha \left(\sum_{i=1}^n a_i b_j a_{ij}\right) + \beta \left(\sum_{i=1}^n c_i b_j a_{ij}\right) = \alpha \phi(x, y) + \beta \phi(z, y)$$

Hence ϕ is linear in the first variable. Similarly ϕ is linear in the second variable. Thus ϕ is a bilinear form on V .

Now $\phi(x_i, x_j) = 1 \cdot 1 \cdot a_{ij} = f(x_i, x_j)$; for all $(x_i, x_j) \in S \times S$.

Hence restriction of ϕ on $S \times S$ is f . It can be easily proved that ϕ is unique.

The next theorem gives the dimension of $B(V)$.

Theorem: Let V be a vector space of dimension n over a field F . $S = \{x_1, x_2, \dots, x_n\}$ be a basis of V and $\{x_1^*, x_2^*, \dots, x_n^*\}$ be its dual basis.

Let $f_{ij}: V \times V \rightarrow F$ (where $i, j = 1, 2, \dots, n$) be defined by

$f_{ij}(u, v) = x_i^*(u) x_j^*(v)$. Then $\{f_{ij}: i, j = 1, 2, \dots, n\}$ is a basis of $B(V)$ Thus dimension $B(V) = n^2$.

Proof: Here we see that every f_{ij} is a bilinear form on V .

Let $f: \{x_i\} \times \{x_i\} \rightarrow F$ be defined by

$$f(x_i, x_j) = a_{ij} \text{ where } x_i, x_j \in \{x_i\}, a_{ij} \in F, i, j = 1, 2, \dots, n$$

Then f can be extended to a unique bilinear form on V by the previous Theorem.

Now we prove that $f = \sum_{i,j=1}^n a_{ij} f_{ij}$. Now if $x_s, x_t \in \{x_i\}$ then $(\sum_{i,j=1}^n a_{ij} f_{ij})(x_s, x_t) = \sum_{i,j=1}^n a_{ij} f_{ij}(x_s, x_t)$

$$\sum_{i,j=1}^n a_{ij} x_i^*(x_s) x_j^*(x_t) = \sum_{i,j=1}^n a_{ij} \delta_{is} \delta_{jt} = \sum_{i,t=1}^n a_{it} \delta_{is} = a_{st} = f(x_s, x_t)$$

From this it can be proved that $(\sum_{i,j=1}^n a_{ij} f_{ij})(u, v) = f(u, v), \forall (u, v) \in V \times V$.

Hence, $f = \sum_{i,j=1}^n a_{ij} f_{ij}$. Therefore, $\{f_{ij}: i, j = 1, 2, \dots, n\}$ spans $B(V)$. Now we prove that $\{f_{ij}: i, j = 1, 2, \dots, n\}$ is linearly independent.

Let $\sum_{i,j=1}^n a_{ij} f_{ij} = 0$ (zero bilinear form on V)

Now, $0(u, v) = 0$ (zero of F) $\forall (u, v) \in V \times V$.

$$\begin{aligned} \text{Hence, } 0 &= 0(x_s, x_t) = (\sum_{i,j=1}^n a_{ij} f_{ij})(x_s, x_t) \\ &= \sum_{i,j=1}^n a_{ij} f_{ij}(x_s, x_t) \end{aligned}$$

$$\sum_{i,j=1}^n a_{ij} x_i^*(x_s) x_j^*(x_t) = \sum_{i,j=1}^n a_{ij} \delta_{is} \delta_{jt} = \sum_{i,t=1}^n a_{it} \delta_{is} = a_{st} = 0, \forall s = 1, 2, \dots, n, t = 1, 2, 3, \dots, n.$$

Hence, the set $\{f_{ij}: i, j = 1, 2, \dots, n\}$ is linearly independent and so it is a basis of $B(V)$. Since $\{f_{ij}: i, j = 1, 2, \dots, n\}$ contains n^2 elements. So, dimension of $B(V) = n^2$.

Theorem: Corresponding to a bilinear form f on a vector space $V(F)$ of dimension n , there exists an n -square matrix $A = [a_{ij}]$ with respect to an ordered basis $S = \{e_1, e_2, \dots, e_n\}$ of V such that $f(e_i, e_j) = a_{ij}$ where $e_i, e_j \in \{e_i\}, a_{ij} \in F, i, j = 1, 2, \dots, n$.

Proof: Let $u, v \in V$ and suppose that $u = \sum_{i=1}^n a_i e_i$ and $v = \sum_{i=1}^n b_i e_i$, then $f(u, v) = f(a_1 e_1 + a_2 e_2 + \dots + a_n e_n, b_1 e_1 + b_2 e_2 + \dots + b_n e_n) = a_1 b_1 f(e_1, e_1) + a_1 b_2 f(e_1, e_2) + \dots + a_n b_n f(e_n, e_n) = \sum_{i,j=1}^n a_i b_j f(e_i, e_j)$.

Let $f(e_i, e_j) = a_{ij}$ where $e_i, e_j \in \{e_i\}, a_{ij} \in F, i, j = 1, 2, \dots, n$

Then, $f(u, v) = \sum_{i,j=1}^n a_i b_j f(e_i, e_j) = X^t A Y$ where $A = [a_{ij}]$

$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in F^n$, and $Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \in F^n$. Thus f is completely determined by the n^2 values of a_{ij} . Thus the matrix $A = ((a_{ij}))$, where

$f(e_i, e_j) = a_{ij}$ where $e_i, e_j \in \{e_i\}$, $a_{ij} \in F$, $i, j = 1, 2, \dots, n$ corresponds to the bilinear form f on V .

Conversely, given an $n \times n$ square matrix $A = ((a_{ij}))$ where

$a_{ij} \in F$, $i, j = 1, 2, \dots, n$ and let the map $f: V \times V \rightarrow F$ be defined as

If $u, v \in V$ and suppose that $u = \sum_{i=1}^n a_i e_i$ and $v = \sum_{i=1}^n b_i e_i$, then

$f(u, v) = \sum_{i,j=1}^n a_i b_j a_{ij} = X^t A Y$ where $A = [a_{ij}]$

$X = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \in F^n$, so, $X^t = [a_1, a_2, \dots, a_n]$ and $Y = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \in F^n$

Then f is a bilinear form on V as proved in above.

Thus, the bilinear form f corresponds to the matrix $A = [a_{ij}]$.

11.1.4. Matrix of a bilinear form on V .

Let $S = \{e_1, e_2, \dots, e_n\}$ be a basis of V . Let f be a bilinear form on V . Then the matrix of f relative to the ordered basis $\{e_i\}$ is defined as:

Matrix of $f = A = ((a_{ij}))$ of order $n \times n$ where $f(e_i, e_j) = a_{ij}$ where $e_i, e_j \in \{e_i\}$, $a_{ij} \in F$, $i, j = 1, 2, \dots, n$ corresponds to the bilinear form f on V .

We next find, how does a matrix of a bilinear form change when a new basis is chosen.

Theorem: Let A be the matrix of a bilinear form f on V relative to the basis $S = \{e_1, e_2, \dots, e_n\}$. Then the matrix of f relative to a new basis $S' = \{e'_1, e'_2, \dots, e'_n\}$ is B given by $B = P^t A P$

Where P is the transition matrix from basis $\{e_i\}$ to the basis $\{e'_i\}$

Proof: Let $x = \sum_{i=1}^n x_i e_i$ and $y = \sum_{i=1}^n y_i e_i$ be elements of V .

Co-ordinate vectors of x and y relative to basis $\{e_i\}$, written as column matrices are

$$X = \begin{bmatrix} x_1 \\ x_2 \\ x_n \end{bmatrix} \quad \text{and} \quad Y = \begin{bmatrix} y_1 \\ y_2 \\ y_n \end{bmatrix} \in F^n$$

Let $x = \sum_{i=1}^n x'_i e'_i$ and $y = \sum_{i=1}^n y'_i e'_i$ be elements of V .

Then Co-ordinate vectors of x and y relative to the new basis $\{e'_i\}$, written as column matrices are

$$X = \begin{bmatrix} x'_1 \\ x'_2 \\ x'_n \end{bmatrix} \quad \text{and} \quad Y = \begin{bmatrix} y'_1 \\ y'_2 \\ y'_n \end{bmatrix} \in F^n$$

Since P is the transition matrix from basis $\{e_i\}$ to basis $\{e'_i\}$ so,

$$PX' = X \text{ and } PY' = Y.$$

Since A is the matrix of f relative to basis $\{e_i\}$ so $f(x, y) = X^t A Y$.
 $= (PX')^t A (PY') = (X')^t (P^t A P) Y' = (X')^t (B) Y'$

$$\text{where } B = P^t A P$$

Hence, $B = P^t A P$ is the matrix of f relative to the new basis $\{e'_i\}$.

11.1.5. Congruent matrix: A matrix B is said to be congruent to a matrix A if there exists a non-singular matrix P such that $B = P^t A P$.

It is clear that matrices of the same bilinear form relative to different bases are congruent.

Example: Let f be the bilinear form on R^3 defined by

$$f((x_1, x_2, x_3), (y_1, y_2, y_3)) = 3x_1 y_1 - 2x_1 y_2 + 5x_2 y_1 + 7x_2 y_2 - 8x_2 y_3 + 4x_3 y_2 - x_3 y_3. \text{ Find matrices of } f \text{ in the bases}$$

$$(i) \{ e_1 = (1, 0, 0), e_2 = (0, 1, 0), e_3 = (0, 0, 1) \}$$

$$(ii) \{ e'_1 = (1, 1, 0), e'_2 = (1, 0, 1), e'_3 = (0, 0, 1) \}$$

and verify that they are congruent.

Solution: It can be easily verified that f is a bilinear form on R^3 .

{Proof is left as an exercise}.

Now the matrix A of f relative to the basis $\{e_i\}$ is given by

$$A = ((a_{ij}))$$

Where, $f(e_i, e_j) = a_{ij}$ where $e_i, e_j \in \{e_i\}$, $a_{ij} \in F$, $i, j = 1, 2, 3$ corresponds to the bilinear form f on V .

$$a_{11} = f(e_1, e_1) = f((1, 0, 0), (1, 0, 0)) = 3.1.1 = 3$$

$$a_{12} = f(e_1, e_2) = f((1, 0, 0), (0, 1, 0)) = -2.1.1 = -2.$$

$$a_{13} = f(e_1, e_3) = f((1, 0, 0), (0, 0, 1)) = 0$$

These are nothing but a_{11} is the coefficient of $x_1 y_1$ in

$$f((x_1, x_2, x_3), (y_1, y_2, y_3)) = 3. \text{ Similarly,}$$

$$a_{12} \text{ is the coefficient of } x_1 y_2 \text{ in } f((x_1, x_2, x_3), (y_1, y_2, y_3)) = -2$$

$$a_{13} \text{ is the coefficient of } x_1 y_3 \text{ in } f((x_1, x_2, x_3), (y_1, y_2, y_3)) = 0$$

$$a_{21} \text{ is the coefficient of } x_2 y_1 \text{ in } f((x_1, x_2, x_3), (y_1, y_2, y_3)) = 5$$

$$a_{22} \text{ is the coefficient of } x_2 y_2 \text{ in } f((x_1, x_2, x_3), (y_1, y_2, y_3)) = 7$$

$$a_{23} \text{ is the coefficient of } x_2 y_3 \text{ in } f((x_1, x_2, x_3), (y_1, y_2, y_3)) = -8$$

$$a_{31} \text{ is the coefficient of } x_3 y_1 \text{ in } f((x_1, x_2, x_3), (y_1, y_2, y_3)) = 0$$

$$a_{32} \text{ is the coefficient of } x_3 y_2 \text{ in } f((x_1, x_2, x_3), (y_1, y_2, y_3)) = 4$$

$$a_{33} \text{ is the coefficient of } x_3 y_3 \text{ in } f((x_1, x_2, x_3), (y_1, y_2, y_3)) = -1$$

Hence we have the required matrix $A = \begin{bmatrix} 3 & -2 & 0 \\ 5 & 7 & -8 \\ 0 & 4 & -1 \end{bmatrix}$

(II). The matrix B relative to the basis $\{e'_i\}$ is given by $B =$

$$\left((b_{ij}) \right). \text{ Where, } f(e'_i, e'_j) = b_{ij} \text{ where } e'_i, e'_j \in \{e'_i\}, b_{ij} \in F, i, j = 1, 2, 3$$

corresponds to the bilinear form f on V .

$$b_{11} = f(e'_1, e'_1) = f((1, 1, 0), (1, 1, 0)) = 13$$

$$b_{12} = f(e'_1, e'_2) = f((1, 1, 0), (1, 0, 1)) = 0.$$

$$b_{13} = f(e'_1, e'_3) = f((1, 1, 0), (0, 0, 1)) = -8. \text{ Similarly,}$$

$$b_{21} = f(e'_2, e'_1) = f((1, 0, 1), (1, 1, 0)) = 5$$

$$b_{22} = f(e'_2, e'_2) = f((1, 0, 1), (1, 0, 1)) = 2$$

$$b_{23} = f(e'_2, e'_3) = f((1, 0, 1), (0, 0, 1)) = -1$$

$$b_{31} = f(e'_3, e'_1) = f((0, 0, 1), (1, 1, 0)) = 4$$

$$b_{32} = f(e'_3, e'_2) = f((0,0,1), (1,0,1)) = -1$$

$$b_{33} = f(e'_3, e'_3) = f((0,0,1), (0,0,1)) = -1$$

$$\text{Hence we have the required matrix } B = \begin{bmatrix} 13 & 0 & -8 \\ 5 & 2 & -1 \\ 4 & -1 & -1 \end{bmatrix}$$

$$\text{Now, } (1,1,0) = 1(1,0,0) + 1(0,1,0) + 0(0,0,1)$$

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

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

Hence the transition matrix P from the basis $\{e_i\}$ to the basis $\{e'_i\}$ is given by

$$P = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}, \text{ so, } P^t = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\begin{aligned} \text{Now, } P^t A P &= \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 3 & -2 & 0 \\ 5 & 7 & -8 \\ 0 & 4 & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 3 & 0 \\ 12 & -3 & -8 \\ 4 & -1 & -1 \end{bmatrix} = \begin{bmatrix} 13 & 0 & -8 \\ 5 & 2 & -1 \\ 4 & -1 & -1 \end{bmatrix} = B \end{aligned}$$

Therefore, matrix B is congruent to the matrix A .

11.1.6. Symmetric, skew-symmetric and alternating bilinear forms:

1. A bilinear form f on V is said to be symmetric if

$$f(u, v) = f(v, u) \text{ for } \forall u, v \in V.$$

2. A bilinear form f on V is said to be skew-symmetric if

$$f(u, v) = -f(v, u) \text{ for } \forall u, v \in V.$$

3. A bilinear form f on V is said to be alternating if

$$f(v, v) = 0 \text{ for } \forall v \in V$$

Note: If characteristic $F \neq 2$, i.e. $1 + 1 \neq 0$ in F , then the above relation (2) implies that $f(u, v) = -f(v, u) \Rightarrow (1 + 1)f(v, v) = 0 \Rightarrow f(v, v) = 0$ for $\forall u, v \in V$ (since $1 + 1 \neq 0$).

That is relation (2) $\Rightarrow$ relation (3) $\Rightarrow$. Also, $f(v, v) = 0$ for $\forall v \in V$

This implies that $f(x + y, x + y) = 0, \forall x, y \in V$.

$$\Rightarrow f(x, x) + f(x, y) + f(y, x) + f(y, y) = 0$$

$$\Rightarrow f(x, y) + f(y, x) = 0, \text{ Since, } f(x, x) = f(y, y) = 0$$

$$\Rightarrow f(x, y) = -f(y, x)$$

That is relation(3) $\Rightarrow$ relation(2)

Remark: We see that when characteristic $F \neq 2$, i.e. $1 + 1 \neq 0$ in F , then skew – symmetric and alternating bilinear forms are equivalent.

Note: If A be a matrix of a symmetric bilinear form f on V , then A is symmetric or $A = A^t$. If A be a matrix of a symmetric bilinear form f on V relative to the basis $S = \{e_1, e_2, \dots, e_n\}$ and if $u = \sum_{i=1}^n a_i x_i$ $v = \sum_{j=1}^n b_j x_j \in V$ then

$$f(u, v) = X^t A Y = (X^t A Y)^t = Y^t A^t X. \text{ Since, } X^t A Y = (X^t A Y)^t$$

Now, $f(v, u) = Y^t A X$. Since, $f(u, v) = f(v, u)$ for $\forall u, v \in V$

So, $Y^t A^t X = Y^t A X$ and since this is true for all vectors X, Y , so, $A^t = A$.

Theorem: If A be a symmetric matrix over F whose characteristic $\neq 2$, then there exists a non-singular matrix P such that $P^t A P$ is diagonal, that is A is congruent to a diagonal matrix.

Proof: Since a non-singular matrix P is a product of elementary matrices, to getting the diagonal form $P^t A P$ is by a sequence of elementary row operations and the same sequence of elementary column operations to reduce it to the diagonal form. These same elementary operations on I will give P^t . For let $E_1, E_2, \dots, E_r$ be the elementary matrices of the sequence of elementary row operations and let $C_1, C_2, \dots, C_r$ be the elementary matrices of the corresponding sequence of elementary column operations to reduce A to the diagonal form B . Then

$$B = E_1, E_2, \dots, E_r A C_1, C_2, \dots, C_r$$

$$\text{or, } B = (E_1, E_2, \dots, E_r I) A (I C_1, C_2, \dots, C_r)$$

$$= P^t A P, \text{ where, } (E_1, E_2, \dots, E_r I) = P^t \text{ and}$$

$$(I C_1, C_2, \dots, C_r) = P$$

Example: Find the diagonal matrix congruent to the symmetric matrix

$$A = \begin{bmatrix} 1 & 2 & -3 \\ 2 & 5 & -4 \\ -3 & -4 & 8 \end{bmatrix} \text{ and express the diagonal form as } P^t A P.$$

Solution: The matrix $(A, I) = \begin{bmatrix} 1 & 2 & -3 \\ 2 & 5 & -4 \\ -3 & -4 & 8 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

Now we apply the row operations $R_2 \rightarrow R_2 - 2R_1, R_3 \rightarrow R_3 + 3R_1,$

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

Now we apply the row operations $C_2 \rightarrow C_2 - 2C_1, C_3 \rightarrow C_3 + 3C_1$

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

Now we apply the row operations $R_3 \rightarrow R_3 - 2R_2,$

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

Now we apply the row operations $C_3 \rightarrow C_3 - 2C_2$

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

Hence the diagonal form $B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -5 \end{bmatrix}$

And $P^t = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 7 & -2 & 1 \end{bmatrix}$

We can verify that $P^t A P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -5 \end{bmatrix}$

Check your progress

For the following real symmetric matrix A , find a non-symmetric matrix P such that $P^t A P$ is diagonal.

1. $A = \begin{bmatrix} 0 & 1 & 1 \\ 1 & -2 & 2 \\ 1 & 2 & -1 \end{bmatrix}$
2. $A = \begin{bmatrix} 1 & -3 & 2 \\ -3 & 7 & -5 \\ 2 & -5 & 8 \end{bmatrix}$

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

Theorem: If characteristic $F \neq 2$, every bilinear form on $V(F)$ can be represented uniquely as the sum of a symmetric bilinear form and a skew symmetric bilinear form.

Proof: Let f be a bilinear form on V . We define bilinear form f_s and f_{ss} on V as follows:

$$(a). f_s(x, y) = \frac{1}{2}(f(x, y) + f(y, x)) \text{ and}$$

$$(b). f_{ss}(x, y) = \frac{1}{2}(f(x, y) - f(y, x))$$

Since, $1 + 1 \neq 0$, so $\frac{1}{2}$ has a meaning. We can easily prove that f_s and f_{ss} are bilinear forms on V .

Since, $f_s(x, y) = f_s(y, x)$, so, f_s is a symmetric bilinear form.

Also, $f_{ss}(x, x) = \frac{1}{2}(f(x, x) - f(x, x)) = 0$. so, f_{ss} is a skew-symmetric bilinear form on V .

$$\text{Now, } f_s(x, y) + f_{ss}(x, y) = f(x, y)$$

$$\text{Or, } (f_s + f_{ss})(x, y) = f(x, y). \text{ Hence, } f = (f_s + f_{ss}).$$

Thus f has been expressed as a sum of a symmetric bilinear form and a skew-symmetric bilinear forms on V .

Claim: show that this representation is unique. For this we suppose that if there is another representation $f = (f'_s + f'_{ss})$, where f'_s is symmetric and f'_{ss} skew-symmetric. Then

$$f(x, y) + f(y, x) = (f'_s(x, y) + f'_{ss}(x, y) + f'_s(y, x) + f'_{ss}(y, x))$$

$$= [f'_s(x, y) + f'_s(y, x)] + [f'_{ss}(x, y) + f'_{ss}(y, x)]$$

$$= 2f'_s(x, y) + 0 \text{ (since, } f'_s \text{ is symmetric and } f'_{ss} \text{ skew-symmetric)}$$

$$\text{Therefore, } f'_s(x, y) = \frac{1}{2}(f(x, y) + f(y, x)) = f_s(x, y).$$

Hence, $f'_s = f_s$. Similarly we can prove that $f'_{ss} = f_{ss}$. Hence, the representation is unique.

11.1.7. Quadratic forms

Let V be a vector space over the field F . A map $q: V \rightarrow F$ is called a quadratic form if there exists a symmetrical bilinear form f on V such that $q(x) = f(x, x), \forall x \in V$.

Note: q is called real quadratic form if $F = \mathbb{R}$ and complex quadratic form if $F = \mathbb{C}$.

Theorem: If characteristic $F \neq 2$, every quadratic form q on $V(F)$ determines a unique symmetrical bilinear form f_s on V which is obtained from q according to $f_s(x, y) = \frac{1}{2}(q(x+y) - q(x) - q(y))$. This formula is called the polar form of symmetrical bilinear form f_s .

Proof: Let f be the bilinear form on V such that $f(x, x) = q(x), \forall x \in V$. Also, we know that f can be expressed uniquely as $f = f_s + f_{ss}$

Where, $f_s(x, y) = \frac{1}{2}(f(x, y) + f(y, x))$ is a symmetrical bilinear form and $f_{ss}(x, y) = \frac{1}{2}(f(x, y) - f(y, x))$ is a skew-symmetrical bilinear form. Hence, $f(x, x) = q(x), \forall x \in V$

Or, $q(x) = f_s(x, x) + f_{ss}(x, x) = f_s(x, x)$. Since, $f_{ss}(x, x) = 0$

Thus, q on $V(F)$ determines a unique symmetrical bilinear form f_s on V .

Now, $q(x+y) = f_s(x+y, x+y) = f_s(x, x) + f_s(x, y) + f_s(y, x) + f_s(y, y)$. Since, $f_s(x, y) = f_s(y, x)$. Therefore,

$$q(x+y) = f_s(x, x) + 2f_s(x, y) + f_s(y, y) = q(x) + 2f_s(x, y) + q(y)$$

Therefore, $2f_s(x, y) = q(x+y) - q(x) - q(y)$

Or, $f_s(x, y) = \frac{1}{2}[q(x+y) - q(x) - q(y)]$. Since, $1 + 1 \neq 0$

Theorem: Let q be a quadratic form on an n -dimensional vector space V . Let f be the unique symmetric bilinear form associated with q . Let the matrix of f relative to the basis $S = \{e_1, e_2, \dots, e_n\}$ be the symmetric matrix $A = (a_{ij})$ then if $x = \sum_{i=1}^n x_i e_i \in V$ then

$$q(x) = \sum_{i=1}^n a_{ii} x_i^2 = (x_1, x_2, x_3, \dots, x_n) \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

Proof: $q(x) = f(x, x) = f(\sum_{i=1}^n x_i e_i, \sum_{j=1}^n x_j e_j) = \sum_{i,j=1}^n x_i x_j f(e_i, e_j)$

$$\sum_{i,j=1}^n x_i x_j a_{ij}. \text{ Since, } f(e_i, e_j) = a_{ij}$$

Now, we put $X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$, then we can write

$$q(x) = \sum_{i,j=1}^n a_{ij} x_i x_j = (x_1, x_2, x_3, \dots, x_n) \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$$\text{Or, } q(x) = X^t A X.$$

11.1.8. Matrix of a quadratic form: If f be the unique symmetric bilinear form associated with the quadratic form q then the symmetric matrix A of the symmetric bilinear form f relative to the basis $\{e_i\}$ is called the matrix of the quadratic form q relative to the basis $\{e_i\}$.

Note: We have $q(x) = \sum_{i,j=1}^n a_{ij} x_i x_j$

$$= a_{11}x_1^2 + a_{22}x_2^2 + a_{33}x_3^2 + \dots + a_{nn}x_n^2 + 2 \sum_{i < j=1}^n a_{ij} x_i x_j.$$

The above expression in the variable x_i is termed as the quadratic polynomial of the symmetric matrix $A = (a_{ij})$.

11.1.9. Normal (Canonical) form: Let f be the unique symmetric bilinear form associated with the quadratic form q . If we have a basis $\{e'_i\}$ such that

$a'_{ij} = f(e'_i, e'_j) = 0$, when $i \neq j$, then the matrix of f becomes diagonal and the quadratic polynomial of the matrix of f is given by $q(x) = a'_{11}x_1^2 + a'_{22}x_2^2 + a'_{33}x_3^2 + \dots + a'_{nn}x_n^2$

Such a form of q is called normal or canonical form of q . The basis $\{e'_i\}$ is called an orthogonal basis for f . Such a basis exists when $\text{char}(F) \neq 2$.

Example: Find the matrix of the quadratic form q on R^3 given by

$$q(x_1, x_2, x_3) = x_1^2 - x_3^2 + 3x_1x_2 - 6x_2x_3$$

Solution: Let f be the symmetric bilinear form on R^3 associated with q . Hence

$$f(x, y) = \frac{1}{2} [q(x+y) - q(x) - q(y)]$$

Where $x = (x_1, x_2, x_3)$ and $y = (y_1, y_2, y_3)$

$$\text{Or, } f(x, y) = \frac{1}{2}[(x_1 + y_1)^2 - (x_3 + y_3)^2 + 3(x_1 + y_1)(x_2 + y_2) - 6(x_2 + y_2)(x_3 + y_3) - (x_1^2 - x_3^2 + 3x_1x_2 - 6x_2x_3) - (y_1^2 - y_3^2 + 3y_1y_2 - 6y_2y_3)]$$

After simplification we get that

$$\begin{aligned} f((x_1, x_2, x_3), (y_1, y_2, y_3)) \\ = x_1y_1 - x_3y_3 + \frac{3}{2}x_2y_1 + \frac{3}{2}x_1y_2 - 3x_2y_3 - 3x_3y_2 \end{aligned}$$

Hence, matrix of f relative to the standard basis of R^3 is

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

Which is therefore, also the matrix of q relative to the standard basis of R^3 .

Note: The symmetric matrix $A = (a_{ij})$ of q has the diagonal entries a_{ii} equal to the coefficient of x_i^2 in $q(x_1, x_2, x_3)$ and has the entries a_{ij} and a_{ji} each equal to half the coefficient of $x_i x_j$ in $q(x_1, x_2, x_3)$. Thus we get a method of writing the matrix of the quadratic form q relative to the usual basis from the coefficients in $q(x)$.

Example: Reduce the quadratic form on R^4 to normal form:

$$q(x_1, x_2, x_3, x_4) = 4x_1x_2 + 4x_1x_3 + 2x_1x_4 + 2x_2x_4 + 2x_3x_4$$

Solution: Coefficient of $x_1x_2 = a_{12} = 4 \neq 0$

Now we apply the substitution: $x_1 = y_1 + y_2$; $x_2 = y_1 - y_2$; $x_3 = y_3$, $x_4 = y_4$

$$\text{Then, } q(x) = q(x_1, x_2, x_3, x_4) = 4(y_1 + y_2)(y_1 - y_2) + 4(y_1 + y_2)y_3 + 2(y_1 + y_2)y_4 + 2(y_1 - y_2)y_4 + 2y_3y_4$$

$$= 4y_1^2 - 4y_2^2 + 4y_1y_3 + 4y_2y_3 + 4y_1y_4 + 2y_3y_4$$

It is a quadratic form, so we can write in this form as

$$q(x) = (4y_1^2 + 4y_1y_3 + 4y_1y_4) - 4y_2^2 + 4y_2y_3 + 2y_3y_4$$

$$\begin{aligned} q(x) = 4 \left(y_1^2 + y_1y_3 + y_1y_4 + \frac{1}{4}y_3^2 + \frac{1}{4}y_4^2 + \frac{1}{2}y_3y_4 \right) - y_3^2 - y_4^2 \\ - 2y_3y_4 - 4y_2^2 + 4y_2y_3 + 2y_3y_4 \end{aligned}$$

$$= (y_1 + \frac{1}{2}y_3 + \frac{1}{2}y_4)^2 - y_3^2 - y_4^2 - 4y_2^2 + 4y_2y_3$$

Now, we substitute, $(y_1 + \frac{1}{2}y_3 + \frac{1}{2}y_4) = z_1, y_2 = z_2, y_3 = z_3, y_4 = z_4$

So, we get $q(x) = 4z_1^2 - z_3^2 - z_4^2 - 4z_2^2 + 4z_2z_3$

$$= 4z_1^2 - (4z_2^2 - 4z_2z_3) - z_3^2 - z_4^2$$

$$= 4z_1^2 - 4\left(z_2^2 - z_2z_3 + \frac{1}{4}z_3^2\right) - z_4^2$$

$$= 4z_1^2 - 4\left(z_2 - \frac{1}{2}z_3\right)^2 - z_4^2 = 4X_1^2 - 4X_2^2 - X_4^2$$

Where, $z_1 = X_1, z_2 - \frac{1}{2}z_3 = X_2, z_4 = X_4$

Now we can find the matrix of this quadratic form q as in the previous example.

Example: Reduce the quadratic form on R^3 to normal form by diagonalisation method: $q(x_1, x_2, x_3) = 2x_1x_2 + 4x_1x_3 + 2x_2x_3$

Solution: the polar form of the symmetric bilinear form f associated with q is given by

$$f(x, y) = \frac{1}{2}[q(x+y) - q(x) - q(y)].$$

Where $x = (x_1, x_2, x_3)$ and $y = (y_1, y_2, y_3)$,

$$(x+y) = (x_1 + y_1, x_2 + y_2, x_3 + y_3)$$

$$\text{Hence, } f((x_1, x_2, x_3), (y_1, y_2, y_3)) = \frac{1}{2}[2(x_1 + y_1)(x_2 + y_2) + 4(x_1 + y_1)(x_3 + y_3) + 2(x_2 + y_2)(x_3 + y_3) - (2x_1x_2 + 4x_1x_3 + 2x_2x_3) - (2y_1y_2 + 4y_1y_3 + 2y_2y_3)]$$

$$= x_1y_2 + y_1x_2 + 2(x_1y_3 + y_1x_3) + x_2y_3 + y_2x_3$$

We consider the standard basis of R^3 is $S = \{e_1 = (1,1,0), e_2 = (1,0,1), e_3 = (0,0,1)\}$

$$\text{Now, } f(e_1, e_1) = f((1,1,0), (1,1,0)) = 2,$$

$$f(e_2, e_1) = f((1,0,1), (1,1,0)) = 4,$$

$$f(e_3, e_1) = f((0,0,1), (1,1,0)) = 3,$$

$$f(e_1, e_2) = f((1,1,0), (1,0,1)) = 4$$

$$f(e_2, e_2) = f((1, 0, 1), (1, 1, 0)) = 4$$

$$f(e_3, e_2) = f((0, 0, 1), (1, 0, 1)) = 2$$

$$f(e_1, e_3) = f((1, 1, 0), (0, 0, 1)) = 3$$

$$f(e_2, e_3) = f((1, 0, 1), (0, 0, 1)) = 2$$

$$f(e_3, e_3) = f((0, 0, 1), (0, 0, 1)) = 0$$

Hence matrix of f relative to the basis S is $A = \begin{bmatrix} 2 & 4 & 3 \\ 4 & 4 & 2 \\ 3 & 2 & 0 \end{bmatrix}$

Now we diagonalise A :

$$A \sim \begin{bmatrix} 2 & 4 & 3 \\ 0 & -4 & -4 \\ 0 & -4 & -\frac{9}{2} \end{bmatrix} (R_2 \rightarrow R_2 - 2R_1, R_3 \rightarrow R_3 - \frac{3}{2}R_1)$$

$$\sim \begin{bmatrix} 2 & 0 & 0 \\ 0 & -4 & -4 \\ 0 & -4 & -\frac{9}{2} \end{bmatrix} (C_2 \rightarrow C_2 - 2C_1, C_3 \rightarrow C_3 - \frac{3}{2}C_1)$$

$$\sim \begin{bmatrix} 2 & 0 & 0 \\ 0 & -4 & -4 \\ 0 & 0 & -\frac{1}{2} \end{bmatrix} (R_3 \rightarrow R_3 - R_2)$$

$$\sim \begin{bmatrix} 2 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & -\frac{1}{2} \end{bmatrix} (C_3 \rightarrow C_3 - C_2)$$

$$\sim \begin{bmatrix} 2 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & -1 \end{bmatrix} (R_3 \rightarrow 2R_3)$$

$$\sim \begin{bmatrix} 2 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & -2 \end{bmatrix} (C_3 \rightarrow 2C_3)$$

Hence, the normal form of $q = 2X_1^2 - 4X_2^2 - 2X_3^2$

Example: Reduce the quadratic form on R^3 to normal form by completing square method: $q(x_1, x_2, x_3) = 2x_1x_2 + 4x_1x_3 + 2x_2x_3$

Solution: Now we take the substitution: $x_1 = y_1 + y_2$; $x_2 = y_1 - y_2$; $x_3 = y_3$, we get

$$\begin{aligned}
q(x) &= q(x_1, x_2, x_3) = 2(y_1^2 - y_2^2) + 4(y_1 + y_2)y_3 + 2(y_1 - y_2)y_3 \\
&= (2y_1^2 + 6y_1y_3) - 2y_2^2 + 2y_2y_3 \\
&= 2\left(y_1^2 + 3y_1y_3 + \frac{9}{4}y_3^2\right) - \frac{9}{2}y_3^2 - 2y_2^2 + 2y_2y_3 \\
&= 2\left(y_1 + \frac{3}{2}y_3\right)^2 - \frac{9}{2}y_3^2 - 2y_2^2 + 2y_2y_3
\end{aligned}$$

Now we put $z_1 = \left(y_1 + \frac{3}{2}y_3\right)$, $z_2 = y_2$, $z_3 = y_3$, we get

$$\begin{aligned}
q(x) &= 2z_1^2 - 2z_2^2 - \frac{9}{2}z_3^2 + 2z_2z_3 \\
&= 2z_1^2 - 2\left(z_2^2 + \frac{1}{4}z_3^2 - z_2z_3\right) - 4z_3^2 \\
&= 2z_1^2 - 2\left(z_2 - \frac{1}{2}z_3\right)^2 - 4z_3^2
\end{aligned}$$

We put $z_1 = X_1$, $z_3 = X_2$, $z_2 - \frac{1}{2}z_3 = X_3$, we get

$$q(x) = 2X_1^2 - 4X_2^2 - 2X_3^2$$

11.1.10. Complex Hermitian form: Let V be vector space over a complex field C . A map $f: V \times V \rightarrow C$ is said to be a Complex Hermitian form on V , if the following conditions are satisfied:

- (i). $f(a_1u_1 + a_2u_2, v) = a_1f(u_1, v) + a_2f(u_2, v)$
- (ii). $f(u, v) = \overline{f(v, u)}$

Where $u_1, u_2, v \in V$ and $a_1, a_2 \in C$

Note1: From the condition (i) and (ii) we get if $v_1, v_2, u \in V$ and $a_1, a_2 \in C$, then $f(u, a_1v_1 + a_2v_2) = \overline{f(a_1v_1 + a_2v_2, u)}$

$$= \overline{a_1f(v_1, u) + a_2f(v_2, u)} = \overline{a_1}f(u, v_1) + \overline{a_2}f(u, v_2)$$

Thus f is conjugate linear in the second variable.

2. $f(u, u) = \overline{f(u, u)}$. Hence, $f(u, u)$ is real for all $u \in V$.

Definition: The complex hermitian form $f: V \times V \rightarrow C$ is said to be non-negative semi-definite if $f(u, u) \geq 0, \forall u \in V$. f is said to be positive definite if $f(u, u) > 0, \forall u \neq 0 \in V$

Note: A positive definite complex hermitian form on V is an inner product in V . Thus $f: C^n \times C^n \rightarrow C$ defined by

$f((z_1, z_2, \dots, z_n), (w_1, w_2, \dots, w_n)) = z_1 \overline{w_1} + z_2 \overline{w_2} + \dots + z_n \overline{w_n}$ is a positive definite hermitian form on C^n , since f is an inner product in C^n .

Theorem: Let $A = ((a_{ij}))$ be an n – square hermitian matrix over C . Then

$f(X, Y) = X^t A \bar{Y}$ where $X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$, $Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$ are elements of C^n is a hermitian form on C^n .

Proof: Let $X_1, X_2, Y \in C^n$ and $a_1, a_2 \in C$, then

$$\begin{aligned} f(a_1 X_1 + a_2 X_2, Y) &= (a_1 X_1 + a_2 X_2)^t A \bar{Y} = (a_1 X_1^t + a_2 X_2^t) A \bar{Y} \\ &= a_1 (X_1^t A \bar{Y}) + a_2 (X_2^t A \bar{Y}) = a_1 f(X_1, Y) + a_2 f(X_2, Y). \end{aligned}$$

Hence, f is linear in the first variable. Also,

$$\overline{f(Y, X)} = \overline{Y^t A \bar{X}} = \overline{(Y^t A \bar{X})^t} \text{ (since, } Y^t A \bar{X} \text{ being a scalar equal to its transpose).}$$

$$= \overline{\bar{X}^t A^t \bar{Y}} = X^t \overline{A^t} \bar{Y} = X^t A \bar{Y}. \text{ (Since, } A = \overline{A^t})$$

$$= f(X, Y). \text{ Hence, } f \text{ is a hermitian form on } C^n.$$

Theorem: Corresponding to a complex hermitian form f on a vector space V over C of dimension n , there exists an n – square hermitian matrix

$A = ((a_{ij}))$ relative to an ordered basis $S = \{e_1, e_2, \dots, e_n\}$ of V , such that $f(e_i, e_j) = a_{ij}$ for $i, j = 1, 2, \dots, n$

Proof: Let $x = \sum_{i=1}^n a_i e_i, y = \sum_{j=1}^n b_j e_j \in V$, then

$$\begin{aligned} f(x, y) &= f(\sum_{i=1}^n a_i e_i, \sum_{j=1}^n b_j e_j) \\ &= f(a_1 e_1 + a_2 e_2 + \dots + a_n e_n, b_1 e_1 + b_2 e_2 + \dots + b_n e_n) \\ &= a_1 \bar{b}_1 f(e_1, e_1) + a_1 \bar{b}_2 f(e_1, e_2) + \dots + a_n \bar{b}_n f(e_n, e_n) \\ &= \sum_{i,j=1}^n a_i \bar{b}_j f(e_i, e_j) = \sum_{i,j=1}^n (a_i \bar{b}_j a_{ij}) = X^t A \bar{Y} \end{aligned}$$

where $X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$, $Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$ and $A = ((a_{ij}))$ be an n – square hermitian matrix over C . Thus f is determined by the matrix $A = ((a_{ij}))$.

Now, $f(e_j, e_i) = a_{ji}$. Therefore, $\overline{f(e_j, e_i)} = \overline{a_{ji}}$

$= f(e_i, e_j) = a_{ij}$. Hence, $A = \overline{A^t} = A^*$.

Thus A is hermitian. The matrix $A = ((a_{ij}))$ where $f(e_i, e_j) = a_{ij}$ is called the matrix of the hermitian form f on V relative to the basis $\{e_i\}$.

Theorem: Let f a hermitian form on V . Let $A = ((a_{ij}))$ be the matrix of the hermitian form f on V relative to the basis $S = \{e_1, e_2, \dots, e_n\}$ of V . If P is the transition matrix of f relative to a new basis $\{e'_i\}$. Then $B = P^t A \bar{P}$ is the matrix of f relative to a new basis $\{e'_i\}$.

Proof: Let $x = \sum_{i=1}^n a_i e_i, y = \sum_{j=1}^n b_j e_j \in V$, then

$$f(x, y) = (a_1, a_2, \dots, a_n) ((a_{ij})) \begin{bmatrix} \overline{b_1} \\ \overline{b_2} \\ \overline{b_n} \end{bmatrix}$$

Since, P is the transition matrix from $\{e_i\}$ to $\{e'_i\}$.

So, we get $\begin{bmatrix} a_1 \\ a_2 \\ a_n \end{bmatrix} = P \begin{bmatrix} a'_1 \\ a'_2 \\ a'_n \end{bmatrix}$ where $x = \sum_{i=1}^n a'_i e'_i$ and

$$\begin{bmatrix} b_1 \\ b_2 \\ b_n \end{bmatrix} = P \begin{bmatrix} b'_1 \\ b'_2 \\ b'_n \end{bmatrix} \text{ where } y = \sum_{j=1}^n b'_j e'_j$$

Hence, $(a_1, a_2, \dots, a_n) = (a'_1, a'_2, \dots, a'_n) P^t$ and

$$\begin{bmatrix} \overline{b_1} \\ \overline{b_2} \\ \overline{b_n} \end{bmatrix} = \bar{P} \begin{bmatrix} \overline{b'_1} \\ \overline{b'_2} \\ \overline{b'_n} \end{bmatrix}. \text{ Hence, we get } f(x, y) = (a'_1, a'_2, \dots, a'_n) P^t ((a_{ij})) \bar{P} \begin{bmatrix} \overline{b'_1} \\ \overline{b'_2} \\ \overline{b'_n} \end{bmatrix}$$

$$= (a'_1, a'_2, \dots, a'_n) B \begin{bmatrix} \overline{b'_1} \\ \overline{b'_2} \\ \overline{b'_n} \end{bmatrix}, \text{ Where } B = P^t A \bar{P}$$

Hence, $B = P^t A \bar{P}$ is the matrix of f relative to a new basis $\{e'_i\}$.

Example: Let $H = \begin{bmatrix} 1 & 1+i & 2i \\ 1-i & 4 & 2-3i \\ -2i & 2+3i & 7 \end{bmatrix}$ be a hermitian matrix. Find a non-singular matrix P such that $P^t H \bar{P}$ is diagonal.

Solution: The matrix $(H, I) = \begin{bmatrix} 1 & 1+i & 2i \\ 1-i & 4 & 2-3i \\ -2i & 2+3i & 7 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

Now we apply hermitian row operations $R_2 \rightarrow R_2 + (-1+i)R_1, R_3 \rightarrow R_3 +$

$$2iR_1, \begin{bmatrix} 1 & 1+i & 2i \\ 0 & 2 & -5i \\ 0 & 5i & 3 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -1+i & 1 & 0 \\ 2i & 0 & 1 \end{bmatrix}$$

Now we apply hermitian row operations $C_2 \rightarrow C_2 + (-1-i)C_1, C_3 \rightarrow C_3 -$
 $2iC_1$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & -5i \\ 0 & 5i & 3 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -1+i & 1 & 0 \\ 2i & 0 & 1 \end{bmatrix}$$

Now we apply hermitian row operations $R_3 \rightarrow 2R_3 - 5iR_2,$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & -5i \\ 0 & 0 & -19 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -1+i & 1 & 0 \\ 5+9i & -5i & 2 \end{bmatrix}$$

Now we apply hermitian column operations $C_3 \rightarrow 2C_3 + 5iC_2$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -38 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -1+i & 1 & 0 \\ 5+9i & -5i & 2 \end{bmatrix}$$

Hence the diagonal form $B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -38 \end{bmatrix} = P^t H \bar{P}$

And $P^t = \begin{bmatrix} 1 & 0 & 0 \\ -1+i & 1 & 0 \\ 5+9i & -5i & 2 \end{bmatrix}$

We can verify that $P^t H P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -38 \end{bmatrix}$

The matrix B is called congruent to H .

Check your progress

Find the symmetric matrix of the following quadratic polynomials relative to the usual basis:

1. $q(x, y, z) = 3x^2 + 4xy - z^2 + 8xz - 6yz + z^2$
2. $q(x, y, z) = xz + x^2 - 2yz - 2xy$

Reduce the quadratic form on R^3 to normal form by diagonalisation method/
completing square method:

3. $q(x, y, z) = xy - 2yz - 4xz$
4. $q(x, y, z) = x^2 + y^2 + z^2 + xy + 2yz + 3xz$
5. $q(x, y, z) = 2x^2 + y^2 + 5z^2 - 8xy + 14yz - 16xz$

Write down the polar form and the associated matrix of the following quadratic forms:

6. $q(x) = x_1x_3 + x_4^2$
7. $q(x) = 2x_1x_2 - x_3x_4$
- 8.
9. $q(x) = x_1^2 + 6x_1x_2 - 2x_2^2 - 2x_2x_3 + x_3^2$

For each of the following matrices A , find a non-singular matrix P such that P^tAP is diagonal.

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

For each of the following hermitian matrices H , find a non-singular matrix P such that $P^tH\bar{P}$ is diagonal.

1. $H = \begin{pmatrix} 1 & i \\ -i & 2 \end{pmatrix}$
2. $H = \begin{pmatrix} 1 & 2+3i \\ 2-3i & -1 \end{pmatrix}$
3. $H = \begin{bmatrix} 1 & i & 2+i \\ -i & 2 & 1-i \\ 2-i & 1+i & 2 \end{bmatrix}$

11.11 Summary: After studying this unit you should be able to define the Bilinear form on a vector space, determine matrix of a bilinear form on V , Congruent matrix, symmetric, skew-symmetric and alternating bilinear forms, quadratic forms, matrix of quadratic form, normal (canonical form) and determine Complex hermitian form and its matrix.

11.12 Terminal Questions

Q.1-Show that the matrix of the bilinear form $f: \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}$ given by

$$f(X, Y) = 2x_1y_2 - x_2y_1 + 2x_2y_3$$

$$X = (x_1 x_2) \text{ \& } Y$$

$= (y_1, y_2, y_3)$. show that the matrix of f relative to the basis of

$$\mathbb{R}^2 \text{ and } \mathbb{R}^3 \quad \text{is} \quad \begin{bmatrix} 0 & 2 & 0 \\ -1 & 0 & 9 \end{bmatrix}.$$

Q.2- Find the matrix of the bilinear form $f(X, Y) = x_1 y_2 + y_1 x_2$ when $X = (x_1 x_2), Y = (y_1 y_2)$ on relation to the ordered basis $\{(1, -1), (0, 1)\}$

Q.3- find the matrix of the following bilinear form $T(x, y) = x_1 y_3 + y_2 x_3$ where

$X = (x_1 x_2 x_3) \text{ \& } Y = (y_1 y_2 y_3)$ w.r.t the base $\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ and

$$\{(1, 0, -1), (0, 1, -1), (1, 1, 1)\}.$$

Q.4- Find the real symmetric matrix A for the quadratic form $q(X) = x_1^2 - x_3 x_4$ in the variables $x_1 x_2 x_3 \text{ \& } x_4$.

Notes

Notes