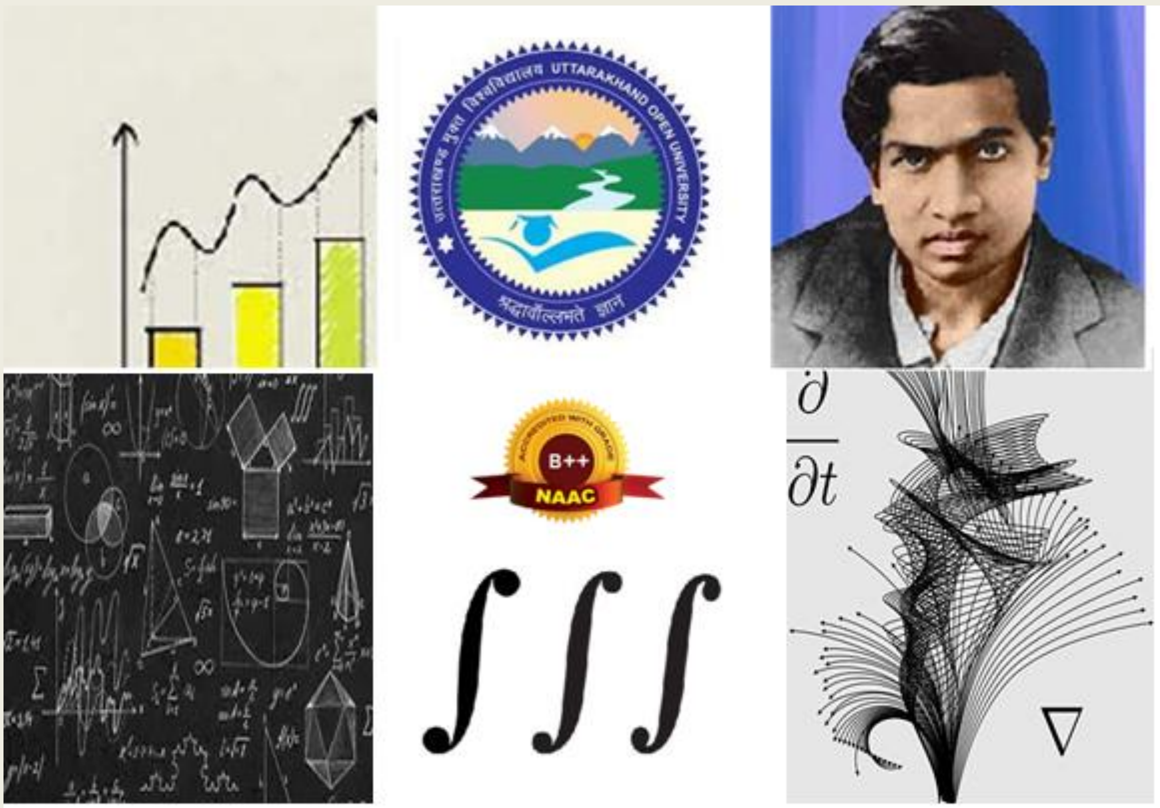
Master of Science
(SECOND SEMESTER)

MAT 505
ADVANCED LINEAR ALGEBRA



DEPARTMENT OF MATHEMATICS
SCHOOL OF SCIENCES
UTTARAKHAND OPEN UNIVERSITY
HALDWANI, UTTARAKHAND
263139

COURSE NAME: ADVANCED LINEAR ALGEBRA

COURSE CODE: MAT-505

Department of Mathematics
School of Science
Uttarakhand Open University
Haldwani, Uttarakhand, India,
263139

BOARD OF STUDIES- 2023

Chairman
Prof. O.P.S. Negi
Honorable Vice Chancellor
Uttarakhand Open University

Prof. P. D. Pant
Director
School of Sciences
Uttarakhand Open University
Haldwani, Uttarakhand

Prof. Harish Chandra
Senior Professor
Department of Mathematics
Institute of Science
Banaras Hindu University
Varanasi

Prof. Manoj Kumar
Professor and Head
Department of Mathematics,
Statistics and Computer Science
G.B. Pant University of Agriculture & Technology,
Pantnagar

Prof. Sanjay Kumar
Professor
Department of Mathematics
Deen Dayal Upadhyaya College
University of Delhi
New Delhi

Dr. Arvind Bhatt
Programme Cordinator
Associate Professor
Department of Mathematics
Uttarakhand Open University
Haldwani, Uttarakhand

Dr. Jyoti Rani
Assistant Professor
Department of Mathematics
Uttarakhand Open University
Haldwani, Uttarakhand

Dr. Kamlesh Bisht
Assistant Professor(AC)
Department of Mathematics
Uttarakhand Open University
Haldwani, Uttarakhand

Dr. Shivangi Upadhyay
Assistant Professor (AC)
Department of Mathematics
Uttarakhand Open University
Haldwani, Uttarakhand

Editor

Dr. Jyoti Rani
Assistant Professor
Department of Mathematics,
School of Sciences,
Uttarakhand Open University

Unit Writer	Blocks	Units
Dr. Shivangi Upadhyay Associate Professor (AC) Department of Mathematics Uttarakhand Open University	I	1 and 2

Dr. Kamlesh Bisht Associate Professor (AC) Department of Mathematics Uttarakhand Open University	I, II, III and IV	3, 4, 5, 6, 7, 8, 13, 14
Dr. Narendra Kumar Singh Assistant Professor Department of Mathematics M.B. Govt. P. G. College, Haldwani	III and IV	9, 10, 12

Block - III	**Units- 11**
Author	
Professor INDRANATH SENGUPTA Department of Mathematics **Indian Institute of Technology Gandhinagar** Gujrat - 382355. https://epgp.inflibnet.ac.in/Home/ViewSubject?catid=ZLCHeZEhCZ8yCri36nSF3A==	
Editing and Typing	
Dr. Kamlesh Bisht Assistant Professor (AC) Department of Mathematics Uttarakhand Open University	
License	
The course on Linear Algebra for **Unit 11** is released under Creative Commons Attribution –Non Commercial- Share Alike -CC-BY-NC-SA license.	

Course Title and Code: Advanced Linear Algebra (MAT-505)

Copyright: Unit 1- Unit 10 and Unit 12 – Unit-14, Uttarakhand Open University

Edition: 2024
Published By: Uttarakhand Open University, Haldwani, Nainital- 263139
ISBN:

***NOTE:** The design and any associated copyright concerns for each unit in this self-learning materials are the sole responsibility of the unit writers.

Contents

Course: Advanced Linear Algebra **Course code: MAT-505**
Credit: 4

COURSE INFORMATION

The present self-learning material **"Advanced Linear Algebra"** has been designed for M.Sc. (Second Semester) learners of Uttarkhand Open University, Haldwani. This self learning material is writing for increase learner access to high-quality learning materials. This course is divided into 14 units of study. The first five units are devoted to vector space & subspace and the application of linear algebra to solve the various types of matrix problem. Unit 6 and Unit 7 are focussed on the topic of quotient space and linear function. The aim of Unit 8, 9 and 10 are to introduce the various application of eigen values, eigen vectors and minimal polynomial to solve the linear equations. Unit 11 explain the Jordan canonical form to understand the application of nilpotent matrix and use of minimal polynomial. Unit 12 and Unit 13 explain the most essential too in linear algebra name as inner product space and operators. Unit 14 will explain the theory of bilinear form. This material also used for competitive examinations. The basic principles and theory have been explained in a simple, concise and lucid manner. Adequate number of illustrative examples and exercises have also been included to enable the leaners to grasp the subject easily.

Course Name: Advanced Linear Algebra **Course Code: MAT 505**

Credit: 4

SYLLABUS

Vector Spaces: Subspaces, Direct Sums, Spanning Sets and Linear Independence, The Dimension of a Vector Space, Ordered Bases and Coordinate Matrices, The Row and Column Spaces of a Matrix.

Linear Transformations, Isomorphisms, The Kernel and Image of a Linear Transformation, Rank-Nullity Theorem, The Matrix of a Linear Transformation, Change of Bases for Linear Transformations, Equivalence of Matrices, Similarity of Matrices, Similarity of Operators, Invariant Subspaces and Reducing Pairs.

The Isomorphism Theorems: Quotient Spaces, The Universal Property of Quotients and the First Isomorphism Theorem, Quotient Spaces, Complements and Co-dimension, Additional Isomorphism Theorems, Linear Functionals, Dual Bases, Reflexivity, Annihilators.

Linear Operator, Characteristic Polynomial and Minimal Polynomial of an Operator, Eigenvalues and Eigenvectors, Geometric and Algebraic Multiplicities, The Jordan Canonical Form, Triangularizability Diagonalizable Operators, Projections, Algebra of Projections, Projections and Invariance.

Real and Complex Inner Product Spaces, Norm and Distance, Isometries, Orthogonality, Orthogonal and Orthonormal Sets, The Projection Theorem and Best Approximations, Orthogonal Direct Sums, The Riesz Representation Theorem. The Adjoint of a Linear Operator, Normal Operators, The Matrix of a Bilinear Form, Quadratic Forms, Orthogonality.

REFRENCE BOOKS

1. Gel′fand I.M. Lectures on linear algebra (1989), Courier Corporation.

2. Kenneth Hoffman & Ray Kunze, Linear Algebra (2nd edition)(2015), Prentice-Hall.

3. David C. Lay, Linear Algebra and its Application (3rd Edition) (2007) Pearson Education Asia, India Reprint.

4. Seymour Lipshutz and Marc Lipson, Schaum’s outlines “Linear Algebra” (3rd Edition)(2012), Mc Graw Hill Education.

SUGGESTED READINGS

1. Minking Eie & Shou-Te Chang (2020), A First Course In Linear Algebra, World Scientific.

2. Axler, Sheldon (2015), Linear algebra done right. Springer.

3. https://nptel.ac.in/courses/111106051

4. https://archive.nptel.ac.in/courses/111/104/111104137

BLOCK- I

VECTOR SPACE AND LINEAR TRANSFORMATION

UNIT–1: VECTOR SPACE AND SUBSPACE

CONTENTS

1.1 *INTRODUCTION*

As we know many physical properties like velocity of a moving object, displacement, force applied on a body etc involve both magnitude and direction and such physical notions which involve both magnitude and direction is called a "vector". A vector is represented by an arrow whose length and direction denotes the magnitude of the vector and the direction of vector respectively.

The idea of a vector space developed from the notion of ordinary two- and three-dimensional spaces as collections of vectors {u, v, w, …} with an associated field of real numbers $\{a,b,c, \dots\}$. Vector spaces as abstract algebraic entities were first defined by the Italian mathematician **Giuseppe Peano in 1888**. Peano called his vector spaces "linear systems" because he correctly saw that one can obtain any vector in the space from a linear combination of finitely many vectors and scalars—$av + bw + \dots + cz$.

Giuseppe Peano

Italian mathematician and glottologist Giuseppe Peano was born on August 27, 1858, and passed away on April 20, 1932. He was the inventor of mathematical logic and set theory, and he wrote more than 200 books and papers. He also contributed a great deal of notation to these fields. The Peano axioms are the basic axiomatization of the natural numbers, named after him.

Reference (https://en.wikipedia.org/wiki/Giuseppe_Peano)

1.2 *OBJECTIVES*

In this unit, we will

- Define vector spaces
- Develop the properties of vectors
- Establish important results apply to all vector spaces
- Understand subspace with examples
- Define basis and dimension of vector space

1.3 *VECTOR SPACE*

The following defines the notion of a vector space V and F is the field of scalars.

Definition- Let V be a nonempty set with two operations

(i) **Vector addition:** If any $u, v \in V$ then $u + v \in V$

(ii) **Scalar Multiplication:** If any $u \in V$ and $k \in F$ then $ku \in V$

Then V is called a vector space (over the field F) if the following axioms hold for any vectors if the following conditions hold

[S_1] $(u+v)+w=u+(v+w)$ for any vectors $u, v, w \in V$

[S_2] there exists a vector denoted by '0' in V, such that, for any $u \in V$,

$$u+0=0+u=u$$

Here '0' is called zero vector

[S_3] for each $u \in V$ there exists a vector denoted by '$-u$' in V such that

$$u+(-u)=0=(-u)+u$$

Here '$-u$' is called additive inverse of vector 'u'

[S_4] $u+v=v+u$ for any vectors $u, v \in V$

[P_1] $k(u+v)=ku+kv$, for any $u \in V$ and for any scalar $k \in F$

[P_2] $(k_1+k_2)u=k_1u+k_2u$, for any $u \in V$ and for any scalar $k_1, k_2 \in F$

[P_3] $(k_1k_2)u=k_1(k_2u)$, for any $u \in V$ and for any scalar $k_1, k_2 \in F$

[P_4] $1.u=u$, for any $u \in V$ and for unit scalar $1 \in F$

The elements of the fiels F are called scalars and the elements of the vector space V are called vectors.

NOTE: (i) The conditions [S_1] $-$ [S_4] concerned with additive structure of V and can be summarized by saying that V is a commutative group under addition.

(ii) The vector space V over the field F is denoted by $V(F)$.

Cancellation Law for vector addition

Theorem 1.1: If u, v and w are vectors in a vector space V such that $u+w=v+w$, then $u=v$.

Proof. There exists a vector w' (additive inverse of w) in V such that

$w+w'=0$ (from [S_3])……………………………….(1.3.1)

Therefore,

$u = u + 0 = u + (w + w')$ (from (1.3.1))

$= (u + w) + w'$ (from [S_1])

It is given that $u + w = v + w$, hence

$u = (v + w) + w'$

$= v + (w + w')$ (from [S_1])

$= v + 0$ (from (1.3.1)

$\Rightarrow u = v$ (from [S_2])

Theorem 1.2: Let V be a vector space over field F. then

(i) For any scalar $k \in F$ and $0 \in V, k0 = 0$
(ii) For $0^* \in F$ and any vector $u \in V, 0^*u = 0$
(iii) For any $k \in F$ and any $u \in V$, $(-k)u = k(-u) = -(ku)$

Proof. (i) Let $k \in F$ and $0 \in V$, then

$k0 + k0 = k(0 + 0)$ **(**from [P_1]

$= k0$

$\Rightarrow k0 + k0 = k0 + 0$ (from [S_2])

Using Cancellation Law for vector addition, we get $k0 = 0$

(ii) Let $0^* \in F$ and $u \in V$, then

$0^*u + 0^*u = (0^* + 0^*)u$ **(**from [P_1])

$= 0^*u$

$\Rightarrow 0^*u + 0^*u = 0^*u + 0$

Using Cancellation Law for vector addition, we get $0^*u = 0$

(iii) Let $k \in F$ and $u \in V$ then $ku \in V$ (by scalar multiplication property)

Hence, there exists a unique element $-(ku) \in V$ such that

$ku + \left(-(ku)\right) = 0$ (from [S_3])

Now, $ku + (-k)u = [k + (-k)]u$ $=0^* u = 0$

i.e., $(-ku)$ is additive inverse of ku.

Hence, $(-ku) = -(ku)$ (as inverse of vector is unique)

Now, $k(-u) = k[(-1)u]$ (as $(-1)u = -u$)

$= [k(-1)]u$ (from [P_3])

$= (-k)u$

Hence, $(-k)u = k(-u) = -(ku)$

1.4 *EXAMPLES OF VECTOR SPACE*

In this section we can learn about some important examples of vector space which will be used throughout the text.

Space F^n

Let F be any arbitrary field. The notion F^n is frequently used to denote the set of all n-tuples of elements in F. Then, F^n is a vector space over F using the following operations:

(i) Vector addition:

$(a_1, a_2, \dots \dots \dots .., a_n) + (b_1, b_2, \dots \dots \dots .., b_n) = (a_1 + b_1, a_2 + b_2, \dots \dots \dots .., a_n + b_n)$

(ii) Scalar Multiplication:

$$k(a_1, a_2, \dots \dots \dots .., a_n) = (ka_1, ka_2, \dots \dots \dots .., ka_n)$$

The **zero vector** in F^n is the n-tuple of zeros i.e. $0 = (0,0, \dots \dots \dots .., 0)$ and

the **additive inverse** of a vector $(a_1, a_2, \dots \dots \dots .., a_n)$ is defined by

$-(a_1, a_2, \dots \dots \dots .., a_n) = (-a_1, -a_2, \dots \dots \dots .., -a_n)$.

Matrix space $M_{m \times n}$

The set of all $m \times n$ matrices with entries from a field F is a vector space, which we denoted by $M_{m\times n}(F)$ with the following operations:

Matrix addition: For $A, B \in M_{m\times n}(F)$

$$(A+B)_{ij} = A_{ij} + B_{ij} \text{ for } 1 \leq i \leq m \text{ and } 1 \leq j \leq n$$

For instance,

$$\begin{bmatrix} 1 & 2 \\ 3 & 0 \\ 3 & 3 \end{bmatrix} + \begin{bmatrix} 4 & 2 \\ 5 & 4 \\ -13 & -1 \end{bmatrix} = \begin{bmatrix} 5 & 4 \\ 8 & 0 \\ -10 & 2 \end{bmatrix} \text{ in } M_{3\times 2}(\mathbb{R})$$

Matrix multiplication: For $A, B \in M_{m\times n}(F)$ and $c \in F$

$$(cA)_{ij} = c(A_{ij}) \text{ for } 1 \leq i \leq m \text{ and } 1 \leq j \leq n$$

For instance,

$$-2\begin{bmatrix} 4 & 1 \\ 3 & -4 \\ -1 & -2 \end{bmatrix} = \begin{bmatrix} -8 & -2 \\ -6 & 8 \\ 2 & 4 \end{bmatrix} \text{ in } M_{3\times 2}(\mathbb{R})$$

Polynomial Space $P(x)$

Let $P(x)$ denote the set of all polynomials of the form $p(x) = c_0 + c_1x + c_2x^2 + \cdots + c_mx^m$ where $m = 1,2,\ldots..$ and $c_0, c_1, \ldots, c_m \in F$ (F is field). Then $P(x)$ is vector space over F with following operations:

(i) Vector Addition: Let $p(x) = c_0 + c_1x + c_2x^2 + \cdots + c_mx^m$ and

$$q(x) = d_0 + d_1x + d_2x^2 + \cdots + d_nx^n$$

be polynomials such that $c_0, c_1, \ldots, c_m, d_0, d_1, \ldots, d_n \in F$.

Suppose $m \leq n$ and we define $c_{m+1} = c_{m+2} = \cdots = c_n = 0$. Then

$$p(x) = c_0 + c_1x + c_2x^2 + \cdots + c_mx^m + c_{m+1}x^{m+1} + c_{m+2}x^{m+2} + \cdots + c_nx^n$$

Then $p(x) + q(x) = (c_0 + d_0) + (c_1 + d_1)x + (c_2 + d_2)x^2 + \cdots + (c_n + d_n)x^n$

(ii) Scalar Multiplication: Let $p(x) = c_0 + c_1x + c_2x^2 + \cdots + c_mx^m$ be a polynomial such that $c_0, c_1, \ldots, c_m \in F$ and for any $k \in F$, then

$kp(x) = kc_0 + kc_1x + kc_2x^2 + \cdots + kc_mx^m$

Polynomial space $P_n(x)$

Let $P_n(x)$ denote the set of all polynomials of the form $P_n(x) = c_0 + c_1x + c_2x^2 + \cdots + c_nx^n$ and $c_0, c_1, \dots, c_n \in F$ (F is field). Then $P_n(x)$ is vector space over F with following operations:

(i) Vector Addition: Let $p(x) = c_0 + c_1x + c_2x^2 + \cdots + c_nx^n$ and

$$q(x) = d_0 + d_1x + d_2x^2 + \cdots + d_nx^n$$

be polynomials such that $c_0, c_1, \dots, c_n, d_0, d_1, \dots, d_n \in F$.Then

$$p(x) + q(x) = (c_0 + d_0) + (c_1 + d_1)x + (c_2 + d_2)x^2 + \cdots + (c_n + d_n)x^n$$

(ii) Scalar Multiplication: Let $p(x) = c_0 + c_1x + c_2x^2 + \cdots + c_nx^n$ be a polynomial such that $c_0, c_1, \dots, c_n \in F$ and for any $k \in F$, then

$$kp(x) = kc_0 + kc_1x + kc_2x^2 + \cdots + kc_nx^n$$

Ex.1.1. Let $S = \{(x_1, x_2): x_1, x_2 \in \mathbb{R}\}$. For $(x_1, x_2), (y_1, y_2) \in S$ and $c \in \mathbb{R}$, define

$(x_1, x_2) + (y_1, y_2) = (x_1 + y_1, x_2 - y_2)$ and $c(x_1, x_2) = (cx_1, cx_2)$. Is S vector space?

Sol. Let $u = (x_1, x_2), v = (y_1, y_2),\ w = (z_1, z_2)$. Now,

$$(u + v) + w = \big((x_1, x_2) + (y_1, y_2)\big) + (z_1, z_2)$$

$$= (x_1 + y_1, x_2 - y_2) + (z_1, z_2)$$

$$= (x_1 + y_1 + z_1, x_2 - y_2 - z_2) \ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(1)$$

Now,

$$u + (v + w) = (x_1, x_2) + \big((y_1, y_2) + (z_1, z_2)\big)$$

$$= (x_1, x_2) + (y_1 - z_1, y_2 - z_2)$$

$$= (x_1 + y_1 - z_1, x_2 - (y_2 - z_2))$$

$$= (x_1 + y_1 - z_1, x_2 - y_2 + z_2) \neq (u + v) + w$$

Since $[S_1]$ fail to holds, S is not a vector space with given operations.

Ex.1.2. Let $S = \{(x_1, x_2): x_1, x_2 \in \mathbb{R}\}$. For $(x_1, x_2), (y_1, y_2) \in S$ and $c \in \mathbb{R}$, define

$(x_1, x_2) + (y_1, y_2) = (x_1 + y_1, 1)$ and $c(x_1, x_2) = (cx_1, 1)$. Is S vector space?

Sol. Let $O = (0,0)$ (zero vector) and $u = (x_1, x_2)$, then

$$O + u = (0,0) + (x_1, x_2) = (0 + x_1, 1) = (x_1, 1) \neq u$$

Since [S_2] fail to holds, S is not a vector space with given operations.

CHECK YOUR PROGRESS

Label the following statements as true or false

1. Every vector space need not contains a zero vector.(F)
2. If V is a vector space then $(a+b)(x+y) = ax+ay+bx+by$ for any $x, y \in V$ and any $a, b \in F$.(T)
3. A vector space has unique zero vector.(T)

1.5 *LINEAR COMBINATIONS*

Definition. Let V be a vector space over a field K. A vector v in V is a linear combination of vectors $(\boldsymbol{x_1}, \boldsymbol{x_2}, \boldsymbol{x_3}, \dots \dots, \boldsymbol{x_n})$ in V if there exist scalars $\boldsymbol{a_1}, \boldsymbol{a_2}, \boldsymbol{a_3}, \dots \dots, \boldsymbol{a_n}$ in K such that

$$v = \boldsymbol{a_1 x_1} + \boldsymbol{a_2 x_2} + \boldsymbol{a_3 x_3} + \dots + \boldsymbol{a_n x_n}$$

Examples

1. Suppose we want to express $u = (3,7,-4)$ in $\mathbb{R}^3$ as a linear combination of the vectors $x_1 = (1,2,3)$, $x_2 = (2,3,7)$, $x_3 = (3,5,6)$

We seek scalars a, b, c such that $u = ax_1 + b\,x_2 + c\,x_3$

i.e. $\begin{bmatrix} 3 \\ 7 \\ -4 \end{bmatrix} = a\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + b\begin{bmatrix} 2 \\ 3 \\ 7 \end{bmatrix} + c\begin{bmatrix} 3 \\ 5 \\ 6 \end{bmatrix}$ or

$a + 2b + 3c = 3$

$2a + 3b + 5c = 7$

$3a + 7b + 6c = -4$

Reducing the system to echelon form yields

$a + 2b + 3c = 3$

$-b - c = 1$

$b - 3c = -13$

implies that $a + 2b + 3c = 3$

$-b - c = 1$

$-4c = -12$

Back-substitution yields the solution $a = 2, y = -4, z = 3$. Thus, $u = 2x_1 - 4\,x_2 + 3\,x_3$.

SPANNING SETS

Let V be a vector space over K. Vectors $x_1, x_2, x_3, \dots\dots, x_n$ in V are said to span V or to form a spanning set of V if every v in V is a linear combination of the vectors $x_1, x_2, x_3, \dots\dots, x_n$, i.e. if there exist scalars $a_1, a_2, a_3, \dots\dots, a_n$ in K such that

$v = a_{\,1}x_1 + a_{\,2}x_2 + a_{\,3}x_3 + \cdots \dots + a_{\,n}x_n$.

Example:

Consider the vector space $V = P_n(t)$ consisting of all polynomials of degree less than equal to n. Clearly every polynomial in $P_n(t)$ can be expressed as a linear combination of the $n+1$ polynomials $1, t, t^2, t^3; \ldots; t^n$. Thus, these powers of t form a spanning set for $P_n(t)$.

1.6 *SUBSPACE*

In this section we can learn about subspace of vector space.

Definition: Suppose that V and S are two vector spaces that have identical definitions of vector addition and scalar multiplication, and that S is a subset of , $S \subseteq V$. Then S is a subspace of V.

Another Definition: A subset S of a vector space V is called a subspace of V if the following two properties are satisfied:

(i) If u, v are in S then $u + v$ is also in S.

(ii) If k is a scalar and u is in S then ku is also in S.

NOTE: Every vector space V has at least two subspaces: V itself and the subspace consisting of the zero vector of V. These are called the trivial subspaces of V.

Theorem 1.3: Show that a subspace of a vector space is itself a vector space.

Sol. All the axioms of a vector space hold for the elements of a subspace.

Theorem 1.4: Show that W is a subspace of V if and only if $ku + v \in W$ for all $u, v \in W$ and $k \in \mathbb{R}$.

Proof. Let W is a subspace of V.

If $u, v \in W$ and $k \in \mathbb{R}$ then $ku \in W$ and therefore $ku + v \in W$.

Conversely, suppose that for all $u, v \in W$ and $k \in \mathbb{R}$ we have $ku + v \in W$.

In particular, if $k = 1$ then $u + v \in W$. If $v = 0$ then $ku + v = ku \in W$.

Hence, W is a subspace of V.

1.6.1 *EXAMPLE OF SUBSPACE*

The following section provides a criterion for deciding whether a subset S of a vector space V is a subspace of V.

Subspace of $\mathbb{R}^3$

We know that $\mathbb{R}^3$ is a vector space. Let W_1 be any plane passing through the origin, as given in Fig. 1.5.1.1.

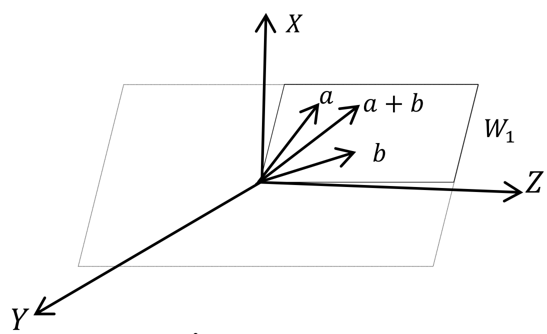


Fig. 1.5.1.1.

Now, we can see that $(0,0,0) \in W_1$ (As we assumed W_1 passing through the origin).

Suppose that vectors $a, b \in W_1$. Then a and b may be viewed as arrows in the plane W_1emanating from origin O, as in given figure. The sum $a + b$ and any multiple ka of a also lie in the plane W_1. Hence, W_1 is a subspace of $\mathbb{R}^3$.

Subspace of $\mathbb{C}^3$

We know that $\mathbb{C}^3$ is a vector space. Now, consider the subset

$$W_2 = \left\{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} : 3x_1 + 5x_2 + 7x_3 = 0 \right\}$$

As we can see that $W_2 \subseteq \mathbb{C}^3$. Now we check the conditions of subspace

(i) Let $x, y \in W_2$ such that $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ and $\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$. Now

$$x + y = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} x_1+y_1 \\ x_2+y_2 \\ x_3+y_3 \end{bmatrix}$$

as follows

$$3(x_1 + y_1) + 5(x_2 + y_2) + 7(x_3 + y_3) = (3x_1 + 5x_2 + 7x_3) + (3y_1 + 5y_2 + 7y_3)$$

$$= 0 + 0$$

$$= 0$$

Hence, $x + y \in W_2$

(ii) Let $x \in W_2$ such that $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ and k is a scalar. Now

$$kx = k\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} kx_1 \\ kx_2 \\ kx_3 \end{bmatrix}$$

as follows

$$3kx_1 + 5kx_2 + 7kx_3 = k(3x_1 + 5x_2 + 7x_3)$$

$$= k.0$$

$$= 0$$

Hence, $\in W_2$, which implies that W_2 satisfies all the conditions of subspace.

Obviously zero vector is in W_2.

Hence W_2 is subspace of $\mathbb{C}^{\mathbf{3}}$**.**

Subspace of Square Matrices:

Consider the vector space of matrices of order $n \times n$.

One possible subspace is the set of **lower triangular matrices**. As $X + Y$ and cX are lower triangular if X and Y are lower triangular and the zero matrix is in given subspace.

Another is the set of **symmetric matrices**. As $X + Y$ and cX are lower triangular if X and Y are lower triangular, and they are symmetric if X and Y are symmetric. Of course, the zero matrix is in given subspaces.

Subspace of the vector space of all functions defined on $[\boldsymbol{a}, \boldsymbol{b}]$.

Let $D([a, b])$ be the collection of all differentiable functions on $[a, b]$.

Let f_1 and f_2 are differential functions on $[a, b]$ and $k \in \mathbb{R}$, then $kf_1 + f_2$ is also differentiable function on $[a, b]$. Hence, $D([a, b])$ is a subspace of all functions defined on $[a, b]$.

Ex.1.3. Consider a Vector Space $\boldsymbol{V}(\mathbb{R})$ as set of all real valued function over $\mathbb{R}$.

$\boldsymbol{V} = \{\boldsymbol{f}: \mathbb{R} \to \mathbb{R}\}$. Then which of the following is\are Subspace of $\boldsymbol{V}(\mathbb{R})$.

(i) $\boldsymbol{W_1} = \{\boldsymbol{f}: \boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{\beta f}(-\boldsymbol{x})\}\ \forall\ \boldsymbol{x} \in \mathbb{R}$ and $\boldsymbol{\beta}$ is given constants over $\mathbb{R}$
(ii) $\boldsymbol{W_2} =$ Set of all integrable functions
(iii) $\boldsymbol{W_3} =$ Set of all non continuous functions

Sol. **(i)** We know that $\boldsymbol{V}(\mathbb{R})$ is a vector space. Now, consider the subset

$W_1 = \{f: f(x) = \beta f(-x)\}\ \forall\ x \in \mathbb{R}$,

As we can see that $W_1 \subseteq \boldsymbol{V}(\mathbb{R})$. Now we check the necessary conditions of subspace.

Let f_1 and $f_2 \in W_1$ and $k \in \mathbb{R}$, then

$(kf_1 + f_2)(x) = kf_1(x) + f_2(x) = k\beta f_1(-x) + \beta f_2(-x) = \beta\big(kf_1(-x) + f_2(-x)\big) = \beta(kf_1 + f_2)(-x) \in W_1$.

Hence, W_1 is a subspace of $V(F)$.

(ii) Let W_2 be the collection of all integrable functions.
Let f_1 and f_2 are integrable functions on $\mathbb{R}$, and $k \in \mathbb{R}$,
then $kf_1 + f_2$ is also integrable functions on $\mathbb{R}$.
Hence, W_2 is a subspace of $V(F)$.

(iii) Let W_3 be the collection of all integrable functions.
Consider non-continuous functions f_1 and f_2 such that

$f_1(x) = \begin{cases} 1, & x \in \mathbb{Q} \\ -1, & x \in \mathbb{R} - \mathbb{Q} \end{cases}$ and $f_2(x) = \begin{cases} -1, & x \in \mathbb{Q} \\ 1, & x \in \mathbb{R} - \mathbb{Q} \end{cases}$
Now, $(f_1 + f_2)(x) = f_1(x) + f_2(x) = 0,$ continuous function.
Therefore, $(f_1 + f_2)(x) \notin W_3$. W_3 is not vector space.

Ex.1.4. Consider a Vector Space $\mathbb{R}^3(\mathbb{R})$. Then which of the following is\are Subspace of $\mathbb{R}^3(\mathbb{R})$.

(i) $\boldsymbol{W_1 = \{(x, y, z): ax + by + cz = 0\},\ a, b, c \in \mathbb{R}}$
(ii) $\boldsymbol{W_2 = \{(x, y, z): x \geq 0\}}$,
(iii) $\boldsymbol{W_3 = \{(x, y, z): x + y + z = 1\}}$**.**

Sol. **(i)** We know that $\mathbb{R}^3(\mathbb{R})$ is a vector space. Now, consider the subset

$W_1 = \{(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}): \boldsymbol{ax} + \boldsymbol{by} + \boldsymbol{cz} = \boldsymbol{0}\},\ \boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c} \in \mathbb{R}$

As we can see that $W_1 \subseteq \mathbb{R}^3(\mathbb{R})$.

Now we check the necessary conditions of subspace.

Let (x_1, y_1, z_1) and $(x_2, y_2, z_2) \in W_1$ and $k \in \mathbb{R}$, then

$$k(x_1, y_1, z_1) + (x_2, y_2, z_2) = (kx_1, ky_1, kz_1) + (x_2, y_2, z_2)$$
$$= (kx_1 + x_2, ky_1 + y_2, kz_1 + z_2).$$

Now, let $a, b, c \in \mathbb{R}$ then

$$k(kx_1 + x_2) + b(ky_1 + y_2) + c(kz_1 + z_2)$$
$$= k(ax_1 + b\,y_1 + cz_1) + (ax_2 + b\,y_2 + cz_2) = k.0 + 0 = 0$$

which implies that $k(x_1, y_1, z_1) + (x_2, y_2, z_2) \in W_1$

Hence, W_1 is a subspace of $\mathbb{R}^3(\mathbb{R})$.

(ii) We know that $\mathbb{R}^3(\mathbb{R})$ is a vector space. Now, consider the subset
$W_2 = \{(x, y, z): x \geq 0\}$
As we can see that $W_2 \subseteq \mathbb{R}^3(\mathbb{R})$. Now we can see that if $k = -1 \in \mathbb{R}$
Let $(x_1, y_1, z_1) \in W_2$ such that $x_1 \geq 0$ then
$(-1)(x_2, y_2, z_2) = (-x_1, y_1, z_1)$
As we know that $x_1 \geq 0 \Rightarrow -x_1 \leq 0$.

Thus, $(-1)(x_1, y_1, z_1) \notin W_2$.
Therefore, W_2 is not a subspace of $\mathbb{R}^3(\mathbb{R})$.

(iii) We know that $\mathbb{R}^3(\mathbb{R})$ is a vector space. Now, consider the subset
$W_3 = \{(x, y, z): x + y + z = 1\}$
As we can see that $W_2 \subseteq \mathbb{R}^3(\mathbb{R})$. Now we can see that $(0,0,0) \notin W_3$.
Therefore, W_3 is not a subspace of $\mathbb{R}^3(\mathbb{R})$.

Ex.1.5. Consider a Vector Space $M_n(\mathbb{R})$. Then which of the following is\are Subspace of $M_n(\mathbb{R})$.

(i) $W_1 = \{A \in M_n(\mathbb{R}): A = bA'\}$, **$b$ is given real number**
(ii) $W_2 = \{A \in M_n(\mathbb{R}): A = A'\}$
(iii) $W_3 = \{A \in M_n(\mathbb{R}): det(A) = 0\}$.
(iv) $W_4 = \left\{A \in M_n(\mathbb{R}): \sum_{i=1}^{n} \sum_{j=1}^{n} K_{ij}\, a_{ij} = 1\right\}, K_{ij} \in \mathbb{R}$.

Sol. **(i)** We know that $M_n(\mathbb{R})$ is a vector space. Now, consider the subset

$$W_1 = \{A \in M_n(\mathbb{R}): A = kA'\},\ k \text{ is given real number}$$

As we can see that $W_1 \subseteq M_n(\mathbb{R})$. Now we check the necessary conditions of subspace.

Let A_1 and $A_2 \in W_1$ and $\alpha, \beta \in \mathbb{R}$, then

$$\alpha A_1 + \beta A_2 = \alpha b A_1' + \beta b A_2' = b(\alpha A_1' + \beta A_2') = b(\alpha A_1 + \beta A_2)' \in W_1.$$

Hence, W_1 is a subspace of $M_n(\mathbb{R})$.

(ii) We know that $M_n(\mathbb{R})$ is a vector space. Now, consider the subset

$$W_2 = \{A \in M_n(\mathbb{R}): A = A'\}$$

As we can see that $W_2 \subseteq M_n(\mathbb{R})$. Now we check the necessary conditions of subspace.

Let A_1 and $A_2 \in W_2$ and $\alpha, \beta \in \mathbb{R}$, then

$$\alpha A_1 + \beta A_2 = \alpha A_1' + \beta A_2' = b(\alpha A_1 + \beta A_2)' \in W_2.$$

Hence, W_2 is a subspace of $M_n(\mathbb{R})$.

(iii) We know that $M_n(\mathbb{R})$ is a vector space. Now, consider the subset

$W_3 = \{A \in M_n(\mathbb{R}): det(A) = 0\}$

As we can see that $W_3 \subseteq M_n(\mathbb{R})$. Now we check the necessary conditions of subspace.

Consider $A_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ and $A_2 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$. We can see that $det(A_1) = 0$ and $det(A_2) = 0$.

Hence, $A_1, A_2 \in W_3$. Now

$A_1 + A_2 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $det(A_1 + A_2) \neq 0$,

which implies that $A_1 + A_2 \notin W_3$

Hence, W_3 is not a subspace of $M_n(\mathbb{R})$.

(iv) We know that $M_n(\mathbb{R})$ is a vector space. Now, consider the subset

$W_4 = \left\{A \in M_n(\mathbb{R}): \sum_{i=1}^{n} \sum_{j=1}^{n} K_{ij}\, a_{ij} = 1\right\}, K_{ij} \in \mathbb{R}$

As we can see that $W_4 \subseteq M_n(\mathbb{R})$.

Now we can see that $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \notin W_4$ as $\sum_{i=1}^{2} \sum_{j=1}^{2} K_{ij}\, a_{ij} = 0 \neq 1$

Therefore, W_4 is not a subspace of $M_n(\mathbb{R})$.

CHECK YOUR PROGRESS

Label the following statements as true or false

1. Every vector space need not contains a zero vector.(F)
2. If V is a vector space then $(a + b)(x + y) = ax + ay + bx + by$ for any $x, y \in V$ and any $a, b \in F$.(T)
3. A vector space has unique zero vector.(T)

1.6.2 *INTERSECTION OF SUBSPACES*

Theorem 1.7: Let W_1 and W_2 be two subspaces of a vector space V then the intersection $W_1 \cap W_2$ is also a subspace of V

Proof. Let W_1 and W_2 be two subspaces of a vector space V.

We show that the intersection $W_1 \cap W_2$ is also a subspace of V.

Clearly, $0 \in W_1$ and $0 \in W_1$ (because W_1 and W_2 be two subspaces). Hence $0 \in W_1 \cap W_2$.

Now suppose w_1 and w_2 belong to the intersection $W_1 \cap W_2$.

Then $w_{1,}\ w_2 \in W_1$ and $w_{1,}\ w_2 \in W_2$. .

For any scalars $, b \in K$, $aw_{1,} + bw_2 \in W_1$ and $aw_{1,} + bw_2 \in W_2$. (because W_1 and W_2 be two subspaces).

Thus, $aw_{1,} + bw_2 \in W_1 \cap W_2$. Therefore, $W_1 \cap W_2$is a subspace of V.

The above result generalizes as follows.

Theorem 1.8: The intersection of any number of subspaces of a vector space V is a subspace of V.

1.7 *LINEAR SPAN*

Let V be a vector space over a field K. A vector v in V is a linear combination of vectors $(x_1, x_2, x_3, \dots \dots, x_n)$ in V if there exist scalars $a_1, a_2, a_3, \dots \dots, a_n$ in K such that $= a_1x_1 + a_2x_2 + a_3x_3 + \cdots + a_nx_n$.

The collection of all such linear combinations, denoted by span$(x_1, x_2, x_3, \dots \dots, x_n)$ or span(x_i) is called the linear span of $x_1, x_2, x_3, \dots \dots, x_n$.

More generally, for any subset S of V, span(x_i) consists of all linear combinations of vectors in S or, when $S = \emptyset$, span$(S) = \{0\}$. Thus, in particular, S is a spanning set of span(S).

Theorem 1.9. Let S be a subset of a vector space V.

(i) Then span(S) is a subspace of V that contains S.
(ii) If W is a subspace of V containing S, then span$(S) \subseteq W$

Proof. (i) Let S be a subset of a vector space V such that $S = \{x_1, x_2, x_3, \dots \dots, x_n\}$

We can see that the zero vector i.e. 0 belongs to span(S), as 0 can be written as

$0 = 0x_1 + 0x_2 + 0x_3 + \cdots + 0x_n$.

Furthermore, let u and u' belong to span(S), i.e.,

$u = a_1x_1 + a_2x_2 + a_3x_3 + \cdots + a_nx_n$.

$u' = b_1x_1 + b_2x_2 + b_3x_3 + \cdots + b_nx_n$

Then, $u + u' = (a_1 + b_1)x_1 + (a_1 + b_2)x_2 + (a_3 + b_3)x_3 + \cdots + (a_n + b_n)x_n$

which implies that $u + u'$ belong to span(S)

and for any scalar $k \in K$,

$ku = ka_1x_1 + ka_2x_2 + ka_3x_3 + \cdots + ka_nx_n$.

which implies that ku belong to span(S).

So, we conclude that span(S) is a subspace of V.

Example

a) Let v_1 be any nonzero vector in $V = \mathbb{R}^3$. Then span(v_1) consists of all scalar multiples of v_1. Geometrically, span(u) is the line through the origin O and the endpoint of u, as shown in Fig. 1.7.1(a).

b) Let u and v be vectors in $V = \mathbb{R}^3$ that are not multiples of each other. Then span(v_1, v_2) is the plane through the origin O and the endpoints of v_1 and v_2 as shown in Fig. 1.7.1(b).

c) Consider the vectors $e_1 = (1,1,1),$, $e_2 = (1,1,0)$, $e_1 = (1,0,0)$ in $V = \mathbb{R}^3$

Row Space of a Matrix

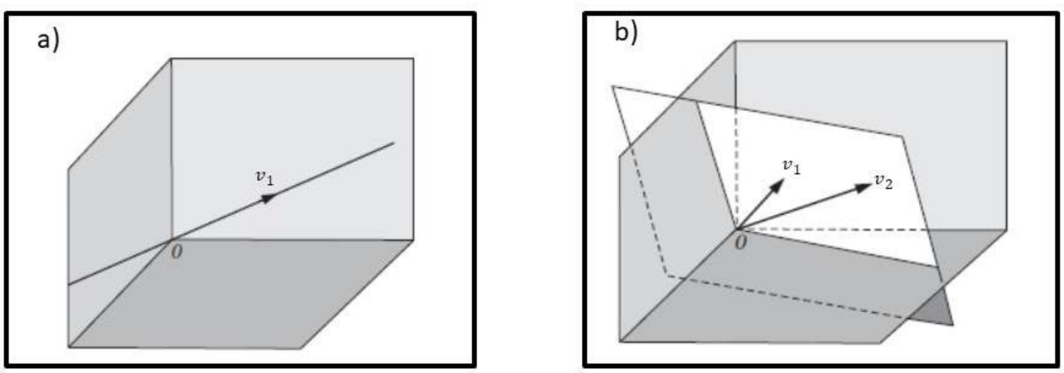


Fig. 1. Linear span

Let $B = [b_{ij}]$ be an arbitrary $m \times n$ matrix over a field K. The rows of A,

$R_1 = (b_{11}, b_{12}, b_{13}, \dots \dots .., b_{1n})$; $R_2 = (b_{21}, b_{22}, b_{23}, \dots \dots .., b_{2n})$, ... $R_m = (b_{m1}, b_{m2}, b_{m3}, \dots \dots .., b_{mn})$;
may be viewed as vectors in K^n; hence, they span a subspace of K^n called the **row space of *B*** and denoted by rowsp(B). That is,
rowsp(B)=span(R_1, R_2; ... , R_m).
Analagously, the columns of B may be viewed as vectors in K^m called the **column space of *B*** and denoted by colsp(B). Observe that colsp(B) = rowsp(A^T).

1.8 *LINEAR INDEPENDENCE*

Let V be a vector space over a field K. The following defines the concept of linear dependence and independence of vectors over K. This notion plays an vital role in the theory of linear algebra and in mathematics in general.

Definition. The vectors $(u_1, u_2, u_3, \dots .., u_n)$ in V are linearly dependent if there exist scalars $(a_1, a_2, a_3, \dots .., a_n)$ in K, not all of them 0, such that $a_1u_1 + a_2u_2 + a_3u_3 + \cdots .. + a_nu_n = 0$ Otherwise, we say that the vectors are linearly independent.

OR

Consider the vector equation

$x_1u_1 + x_2u_2 + x_3u_3 + \cdots .. + x_nu_n = 0$ ....(1.8.1)

where the $x's$ are unknown scalars. This equation always has the zero solution $x_1 = 0$; $x_2 = 0$; ... ; $x_n = 0$. Suppose this is the only solution; that is, suppose we can show: $x_1u_1 + x_2u_2 + x_3u_3 + \cdots .. + x_nu_n = 0$ implies $x_1 = 0, x_2 = 0, x_3 = 0, \dots .., x_n = 0$. Then the vectors $u_1 = 0, u_2 = 0, u_3 = 0, \dots .., u_n = 0$ are linearly independent, On the other hand, suppose the equation (1.8.1) has a nonzero solution; then the vectors are linearly dependent.

NOTE

1. A set $S = (u_1, u_2, u_3, \dots .., u_n)$ of vectors in V is linearly dependent or independent according to whether the vectors $u_1, u_2, u_3, \dots .., u_n$ are linearly dependent or independent.
2. An infinite set S of vectors is linearly dependent or independent according to whether there do or do not exist vectors $u_1, u_2, u_3, \dots .., u_k$ in S that are linearly dependent.
3. Suppose 0 is one of the vectors $u_1, u_2, u_3, \dots .., u_n$ say $u_1 \neq 0$. Then the vectors must be linearly dependent, because we have the following linear combination where the coefficient of $1. u_1 + 0u_2 + 0u_3 + \cdots .. + 0u_n = 0$

4. If a set S of vectors is linearly independent, then any subset of S is linearly independent. Alternatively, if S contains a linearly dependent subset, then S is linearly dependent.

Example

1. Let $u_1 = (1,1,0), u_2 = (1,3,2), u_3 = (4,9,5)$. Then u_1, u_2, u_3 are linearly dependent, because

$$3u_1 + 5u_2 - 2u_3 = 3(1,1,0) + 5(1,3,2) - 2(4,9,5) = (0,0,0) = 0$$

Ex. 1.6. Let V be the vector space of functions from $\mathbb{R}$ into $\mathbb{R}$. Show that the functions $f(x) = e^x$, $g(x) = \sin t$ and $h(x) = x^2$ are linearly independent.

Proof. Let V be the vector space of functions from $\mathbb{R}$ into $\mathbb{R}$. Now we will show that functions $f(x) = e^x$, $g(x) = \sin x$ and $h(x) = x^2$ are linearly independent.

Let a, b and c are unknown scalars such that

$$af + bg + ch = 0 \Rightarrow ae^x + b\sin x + cx^2 = 0, \forall x \in \mathbb{R}.$$

Thus, in this equation, we choose appropriate values of x to easily get $a = 0, b = 0, c = 0$.

For example

i) Substitute $x = 0$ to obtain $ae^0 + b\sin 0 + c0^2 = 0 \Rightarrow a = 0$

ii) Substitute $x = \pi$ and $a = 0$ to obtain $0e^\pi + b\sin\pi + c\pi^2 = 0 \Rightarrow c = 0$

iii) Substitute $x = \frac{\pi}{2}$ and $a = 0$, $c = 0$ to obtain $0e^{\frac{\pi}{2}} + b\sin\frac{\pi}{2} + 0\frac{\pi^2}{2} = 0 \Rightarrow b = 0$

Hence $f(x) = e^x$, $g(x) = \sin t$ and $h(x) = x^2$ are linearly independent.

Ex. 1.7. Let $P_3(\mathbb{R})$ be the vector space of set of polynomials of degree less than equal to 3 defined on $\mathbb{R}$. Show that set $S = \{1 + x + x^2, 7 + x^3, 11 + x + x^2 + x^3, 13 + 4x\}$ are linearly independent.

Proof. Let $P_3(\mathbb{R})$ be the vector space of set of polynomials of degree less than equal to 3 defined on $\mathbb{R}$ and let $S = \{1 + x + x^2, 7 + x^3, 11 + x + x^2 + x^3, 13 + 4x\}$. Now we will show that set S is linearly independent.

Let a, b, c and d are unknown scalars such that

$$a(1 + x + x^2) + b(7 + x^3) + c(11 + x + x^2 + x^3) + d(13 + 4x) = 0$$

$\Rightarrow (a + 7b + 11c + 13d) + (a + c + 4d)x + (a + c)x^2 + (b + c)x^3 = 0$

$\Rightarrow (a + 7b + 11c + 13d) = 0;$ ………………..(1)

$(a + c + 4d) = 0;$ …………………………..(2)

$(a + c) = 0;$ ………………………………….(3)

$(b + c) = 0$ ……………………………………(4)

Using (3) in (2), we get $d = 0$

From (3) and (4), we get $a = b = -c$…………….(5)

Using (5) in (1), we get $a = b = c = 0$.

Hence, set $S = \{1 + x + x^2, 7 + x^3, 11 + x + x^2 + x^3, 13 + 4x\}$ are linearly independent.

LEMMA-Suppose two or more nonzero vectors $v_1, v_2, v_3, \dots\dots, v_m$ are linearly dependent. Then one of the vectors is a linear combination of the preceding vectors; that is, there exists $k > 1$ such that

$v_k = c_1v_1 + c_2v_2 + \cdots .. + c_{k-1}v_{k-1}$

1.9 *SUMMARY*

We discussed about vector space and subspace with the help of illustrative examples.

1.10 *GLOSSARY*

Set: is the mathematical model for a collection of different things

Scalar: is an element of a field which is used to define a vector space

Vector: a term that refers to elements of some vector spaces.

1.11 *REFERENCES*

- David C. Lay, Linear Algebra and its Application, 3rd Edition, Pearson Education Asia, India Reprint, 2007.

- Seymour Lipshutz and Marc Lipson, Schaum's outlines "linear Algebra", 3rd Edition, Mc Graw Hill Education, 2012.
- Gilbert Strang, Introduction to Linear Algebra (Gilbert Strang), 5th Edition

1.12 *SUGGESTED READINGS*

- P.B. Bhattacharya, S.K. Jain and S.R Nagpaul, First Course in Linear Algebra, Wiley Eastern Ltd., 1991.
- S.H Friedberg, A.J. Insel and L.E.Spence, Linear Algebra, 4th Edition, Prentice-Hall of India, New Delhi, 2004.
- R.K. Gupta, Linear Programming, Krishna Prakasan, 2010
- K. Hoffman and R. Kunze, Linear Algebra, 2nd Edition, Prentice-Hall of India, New Delhi, 2000.
- D. Kalman, A singularly valuable decomposition; the SVD of a matrix, The College Math. Journal, Vol .27, No.1, (1996).
- S. Kumaresan, Linear Algebra-A Geometric approach, Prentice-Hall of India, New Delhi, 2001.

1.13 *TERMINAL QUESTIONS*

(TQ-1) Define Vector space
(TQ-2) Define Subspace
(TQ-3) Give Example of Vector Space.

Choose one of correct Choioce:

(TQ-4) If A and B are square matrices of the same order, then tr(AB) =

(a) $tr(A + B)$
(b) $tr(A)tr(B)$
(c) $tr(BA)$
(d) $tr(A) + tr(B)$

(TQ-5) If A and B are square matrices of the same order, then $(AB)^T =$

(a) $A^T B^T$

(b) $B^T . A^T$

(c) $A^T + B^T$

(d) $(BA)^T$

(TQ-6) Let V be the vector space of functions from $\mathbb{R}$ into $\mathbb{R}$. Show that the functions $f(x) = e^x$, $g(x) = \sin t$ and $h(x) = x^2$ are ________________.

1.14 *ANSWERS*

(TQ-4) (c) **(TQ-5)** (b) **(TQ-6)** linearly independent

UNIT–2 BASIS AND DIMENSION

CONTENTS

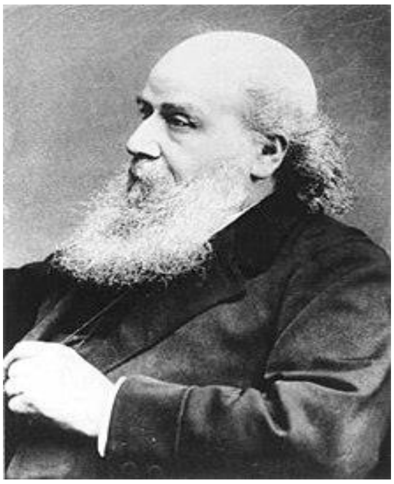
2.1 *INTRODUCTION*

In previous unit we studied about vector space, In this unit we will try to understand basis and dimensions.

We turn now to the task of assigning a dimension to certain vector spaces. Although we usually associate 'dimension' with something geometrical, we must find a suitable algebraic definition of the dimension of a vector space. This will be done through the concept of a basis for the space. One of the useful features of a basis B in an n −dimensional space V is

that it essentially enables one to introduce coordinates in V analogous to the 'natural coordinates' X_i of a vector $x = (x_1, \ldots, x_n)$ in the space F^n. In this scheme, the coordinates of a vector a in V relative to the basis B will be the scalars which serve to express a as a linear combination of the vectors in the basis.

Many mathematical terms, including "matrix" (in 1850), "graph" (in the sense of a network), "discriminant," and "totient" (for Euler's totient function φ(n), were created by Sylvester. He is also credited with solving Sylvester's problem and a result on the orchard problem in discrete geometry, and discovering Sylvester's determinant identity in matrix theory, which generalizes the Desnanot–Jacobi identity. His body of scientific writings fills four volumes. The Royal Society of London awarded Sylvester the Copley Medal, its highest honor for scientific achievement, in 1880, and in 1901 it instituted the Sylvester Medal in his memory, to promote mathematical research following his passing in Oxford.



James Joseph Sylvester
(3 September 1814 – 15 March 1897)
(reference: https://en.wikipedia.org/wiki/James_Joseph_Sylvester

2.2 *OBJECTIVES*

In this unit, we will,

- Define basis with examples
- Understand dimension of vector space

2.3 *BASIS*

A set $S = \{u_1, u_2, u_3, \ldots\ldots, u_m\}$ of vectors is a basis of V if it has the following two properties:

(1) S is linearly independent.
(2) S spans V.

OR

A set $S = \{u_1, u_2, u_3, \dots \dots, u_m\}$ of vectors is a basis of V if every $v \in V$ can be written uniquely as a linear combination of the basis vectors.

Example: (1) Standard basis for $\mathbb{R}^n$ is

$e_1 = (1,0,0, \dots .,0,0), e_2 = (0,1,0, \dots .,0,0), \dots ., e_n = (0,0,0, \dots .,0,1)$

(2) Standard basis for Matrices $M_{2\times 2}$ is

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

(3) The infinite set $\{1, x, x^2, \dots \dots \dots, x^{n-1}, \dots .\}$ form basis for P, the space of all polynomial.

Theorem 1.1. Let V denote a vector space and $S = \{u_1, u_2, u_3, \dots \dots, u_m\}$ a basis of V.

a) Any subset of V containing more than n vectors must be dependent.
b) Any subset of V containing less than n vectors cannot span V.

Proof. (a) Let $S_1 = \{v_1, v_2, v_3, \dots \dots, v_n\}$ a subset of V where $n > m$.

Now we will prove that W is dependent.

Since S is a basis, we can write each v_i in term of elements in S.

Now, there exists constants c_{ij} with $1 \leq i \leq n$ and $1 \leq j \leq m$ such that

$v_i = c_{i1}u_1 + c_{i2}u_2 + \cdots + c_{im}u_m$. Consider the linear combination

$\sum_{j=1}^{n} d_j v_j = \sum_{j=1}^{n} d_j (c_{j1}u_1 + c_{j2}u_2 + \cdots + c_{jm}u_m) = 0$

So we solve

$$\begin{cases} d_1c_{11}u_1 + d_1c_{12}u_2 + \cdots + d_1c_{1m}u_m = 0 \\ d_2c_{21}u_1 + d_2c_{22}u_2 + \cdots + d_2c_{1m}u_m = 0 \\ \quad\quad . \\ \quad\quad . \\ \quad\quad . \\ d_nc_{n1}u_1 + d_nc_{n2}u_2 + \cdots + d_nc_{nm}u_m = 0 \end{cases}$$ where $d_1, d_2, \dots, d_n$ are unknowns

Here we can easily observe that the number of unknowns is less than number of equation.

Hence given Homogeneous equation system will have a nontrivial solution.

Hence S_1 is dependent.

(b) Let $S_1 = \{v_1, v_2, v_3, \dots \dots, v_n\}$ a subset of V where $n < m$.

Now we will prove that S_1 does not span V.

Let we assume that it does span V and show this would imply that S is dependent.

Now, there exists constants c_{ij} with $1 \leq i \leq m$ and $1 \leq j \leq n$ such that

$u_i = c_{i1}v_1 + c_{i2}v_2 + \cdots + c_{in}v_n$. Consider the linear combination

$\sum_{j=1}^{m} d_j u_j = \sum_{j=1}^{n} d_j (c_{j1}v_1 + c_{j2}v_2 + \cdots + c_{jn}v_n) = 0$

So we solve

$$\begin{cases} d_1c_{11} + d_1c_{12} + \cdots + d_1c_{1n} = 0 \\ d_2c_{21} + d_2c_{22} + \cdots + d_2c_{2n} = 0 \\ \quad\quad\quad\quad . \\ \quad\quad\quad\quad . \\ \quad\quad\quad\quad . \\ d_nc_{n1} + d_nc_{n2} + \cdots + d_nc_{mn} = 0 \end{cases}$$ where $d_1, d_2, \ldots, d_n$ are unknowns

Here we can easily observe that the number of unknowns is more than number of equation.

Hence given Homogeneous equation system will have a nontrivial solution.

Hence S is dependent, but it can’t be possible since it is a basis.

Thus our assumption is wrong, S_1 does not span V.

Theorem 1.2 Let V be a vector space such that one basis has m elements and another basis has n elements. Then $m = n$.

Proof. Assume that S is a basis of V with n elements and S^* is another basis with m elements. We need to show that $m = n$.

Since S is a basis, S^* being also a basis implies that $m \geq n$.

If we had m > n, by the theorem, S^* would be dependent, hence not a basis.

Similarly, since S^* is a basis, S being also a basis implies that $n \geq m$. The only way we can have $m \geq n$ and $n \geq m$ is if $m = n$.

CHECK YOUR PROGRESS: 1

1: Prove that every basis of a vector space V has the same number of elements.

2: Any subset of V containing more than n vectors must be ________________.

2.4 *DIMENSIONS*

Let V denote a vector space. Consider a basis of V has m vectors (therefore all bases will have m vectors), m is called the **dimension of $\boldsymbol{V}$**. We can write $dim(V) = m$.

A vector space V is said to be finite-dimensional if there exists a finite subset of V which is a basis of V. If no such finite subset exists, then V is said to be infinite-dimensional.

NOTE:

1: If V is just the vector space consisting of $\{0\}$, then we say that $dim(V) = 0$.

Examples:

1. $\mathbb{R}^n$, the set of all ordered pairs (x, y) where x and y are in $\mathbb{R}$. We have already seen that the standard basis for $\mathbb{R}^2$ is $\{(1, 0,0, \dots ,0), (0, 1,0, \dots ,0), (0, 0,1, \dots ,0), \dots , (0, 0,0, \dots ,1)\}$. This basis has n elements, therefore, $\dim(\mathbb{R}^n) = n$.
2. P_n, the set of polynomials of degree less than or equal to n. Similarly, the standard basis for P_n is $\{1, x, x^2, \dots \dots , x^n\}$. This basis has $n + 1$ elements, therefore $dim(P_n) = 1$.
3. M_{32}, the set of 3×2 matrices. A basis for M_{32} is 6.

Ex.1.8. Find a basis and the dimension of subspace $\boldsymbol{W} = \left\{\begin{bmatrix} a+b+c \\ 2a+b+3c+d \\ b+c+d \\ 2a+2c+d \end{bmatrix} : a, b, c, d \text{ are real}\right\}$

Proof. It is given that

$$W = \left\{\begin{bmatrix} a+b+c \\ 2a+b+3c+d \\ b+c+d \\ 2a+2c+d \end{bmatrix} : a, b, c, d \ are\ real\right\}$$

Now

$$\begin{bmatrix} a+b+c \\ 2a+b+3c+d \\ b+c+d \\ 2a+2c+d \end{bmatrix} = a\begin{bmatrix} 1 \\ 2 \\ 0 \\ 2 \end{bmatrix} + b\begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} + c\begin{bmatrix} 1 \\ 3 \\ 1 \\ 2 \end{bmatrix} + d\begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

Here we can see that $W = span\{v_1, v_2, v_3, v_4\}$ where

$$v_1 = \begin{bmatrix} 1 \\ 2 \\ 0 \\ 2 \end{bmatrix}, v_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix}, v_3 = \begin{bmatrix} 1 \\ 3 \\ 1 \\ 2 \end{bmatrix}, v_4 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

v_1, v_2, v_3, v_4 are linearly independent.
Hence basis of $W = \{v_1, v_2, v_3, v_4\}$ and dimension is 4.

NOTE:

- If V is spanned by a finite set, then V is said to be finite-dimensional, and the dimension of V, written as dim V, is the number of vectors in a basis for V.
- The dimension of the zero vector space $\{0\}$ is defined to be zero.
- If V is not spanned by a finite set, then V is said to be infinite-dimensional.

Theorem 1.4 Let H be a subspace of a finite-dimensional vector space V. Any linearly independent set in H can be expanded, if necessary, to a basis for H. Also, H is finite-dimensional and $dim\, H \leq dim\, V$.

Proof. Let $H = \{0\}$, then clearly $dim\, H = 0 \leq dim\, V$.

Let $H \neq \{0\}$ and $S = \{x_1, x_2, \dots, x_m\}$ be any linearly independent set in H.

If S spans H implies S is a basis of H.

otherwise there exist some x_{m+1} in H which is not in S.

Then $\{x_1, x_2, \dots, x_m, x_{m+1}\}$ will be linearly independent as no vector in the set can be a linear combination of vectors that precede it.

We can keep expanding S to a larger linearly independent set in H as long as the new set does not span H.

However, the number of vectors in an expansion of S that is linearly independent can never be greater than the dimension of V.

Hence the expansion of S will span H and therefore will be a basis for H, and $dim\, H \leq dim\, V$.

NOTE:

1: Above theorem is also natural counterpart to the spanning set theorem.

Theorem 1.5. Let V be a p-dimensional vector space, $p \geq 1$. Any linearly independent set of exactly p elements in V is automatically a basis for V. Any set of exactly p elements that spans V is automatically a basis for V.

Proof. From Above theorem we conclude that, a linearly independent set S of p elements can be extended to a basis for V.

But that basis must contain exactly p elements, since $dim\, V = p$.

So S must already be a basis for V.

Let we assume that S has p elements and spans V.

Since V is nonzero, a subset S' of S is a basis of V (by the Spanning Set Theorem).

Because$dim\ V = p$, $S\ '$ must contain p vectors. Therefore $= S'$.

2.5 *APPLICATION TO MATRICES, RANK OF A MATRIX*

Suppose A be any m_n matrix over a field K. As we know that that the rows of A may be viewed as vectors in K_n and that the row space of A, written $rowsp(A)$, is the subspace of K_n spanned by the rows of A.

Rank of matrix A: The rank of a matrix A, written $rank(A)$, is equal to the maximum number of linearly independent rows of A or, equivalently, the dimension of the row space of A.

The Dimensions of $Nul\ A$ and $Col\ A$:

The dimension of $Nul\ A$ is the number of free variables in the equation $Ax = 0$, and the dimension of $Col\ A$ is the number of pivot columns in A.

As we know vectors in K^n and that the column space of A, written $colsp(A)$, is the subspace of K^n spanned by the columns of A. Although m may not be equal to n—that is, the rows and columns of A may belong to different vector spaces—we have the following fundamental result.

Theorem 1.6: The maximum number of linearly independent rows of any matrix A is equal to the maximum number of linearly independent columns of A. Hence, the dimension of the row space of A is equal to the dimension of the column space of A.

Ex.1.9. Find the dimensions of the null space and the column space of

$$\begin{bmatrix} \mathbf{-3} & \mathbf{6} & \mathbf{-1} & \mathbf{1} & \mathbf{-7} \\ \mathbf{1} & \mathbf{-2} & \mathbf{2} & \mathbf{3} & \mathbf{-1} \\ \mathbf{2} & \mathbf{-4} & \mathbf{5} & \mathbf{8} & \mathbf{-4} \end{bmatrix}$$

Proof. Let $A = \begin{bmatrix} -3 & 6 & -1 & 1 & -7 \\ 1 & -2 & 2 & 3 & -1 \\ 2 & -4 & 5 & 8 & -4 \end{bmatrix}$

Now we reduce above matrix in echelon form

$$\begin{bmatrix} -3 & 6 & -1 & 1 & 0 \\ 1 & -2 & 2 & 3 & 0 \\ 2 & -4 & 5 & 8 & 0 \end{bmatrix} (by\ C_5 \rightarrow C_2 + C_4 + c_5)$$

$$\sim \begin{bmatrix} -3 & 6 & -1 & 1 & 0 \\ 1 & -2 & 2 & 3 & 0 \\ 2 & -4 & 5 & 8 & 0 \end{bmatrix} \; (by\ R_3 \to R_1 + R_2 + R_3)$$

$$\sim \begin{bmatrix} -3 & 6 & -1 & 1 & 0 \\ 1 & -2 & 2 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \; (by\ R_3 \to R_1 + R_2 + R_3)$$

$$\sim \begin{bmatrix} -3 & 6 & -1 & 1 & 0 \\ 0 & 0 & 5 & 10 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \; (by\ R_3 \to R_1 + R_2 + R_3)$$

$$\sim \begin{bmatrix} -3 & 6 & -1 & 1 & 0 \\ 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \; \left(by\ R_2 \to \frac{R_2}{2}\right)$$

There are three free variables—x_2, x_4, and x_5. Hence the dimension of Nul A is 3.

Also, $dim\ Col\ A =\ 2$ because A has two pivot columns.

2.5.1 *BASIS FINDING PROBLEMS*

An echelon form of any matrix A gives us the solution to certain problems

about A itself.

Let $A = \begin{bmatrix} 1 & 2 & 1 & 3 \\ 2 & 5 & 5 & 6 \\ 3 & 7 & 6 & 11 \\ 4 & 8 & 4 & 12 \end{bmatrix}$

$$\begin{bmatrix} 1 & 2 & 1 & 3 \\ 2 & 5 & 5 & 6 \\ 3 & 7 & 6 & 11 \\ 4 & 8 & 4 & 12 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 2 & 1 & 3 \\ 0 & 1 & 3 & 0 \\ 0 & 1 & 3 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix} \; (by\ R_2 \to R_2 - 2R_1,\ R_3 \to R_3 - 3R_1\ and\ R_4 \to R_4 - 3R_1\ )$$

$$\sim \begin{bmatrix} 1 & 2 & 1 & 3 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix} (by\ R_2 \to R_2 - R_3) \sim \begin{bmatrix} \circled{1} & 2 & 1 & 3 \\ 0 & \circled{1} & 3 & 0 \\ 0 & 0 & 0 & \circled{1} \end{bmatrix} (by\ R_2 \to \tfrac{1}{2} R_3)$$ (pivots are circled)

Let $B = \begin{bmatrix} 1 & 2 & 1 & 3 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$ then B is echelon form of matrix A.

We solve the following four problems about the matrix A, where $C_1;\ C_2;\ C_3$and C_4 denote its columns:

(a) Find a basis of the row space of A.

(b) Find each column C_k of A that is a linear combination of preceding columns of A.

(c) Find a basis of the column space of A.

(d) Find the rank of A.

Answer:

(a) We can see that A and B are row equivalent, so they have the same row space. Also, B is in echelon form, hence its nonzero rows are linearly independent and therefore form a basis of the row space of B. Thus, they also form a basis of the row space of A. i.e.

basis of row space of A. i.e. $basis\ of\ row\ sp(A)$: $(1,2,1,3), (0,1,3,0), (0,0,0,1)$

(b) Let $M_k = [C_1, C_2 \dots \dots C_k]$., the submatrix of A consisting of the first k columns of A.

Then $M_{k-1}$and M_k are, respectively, the coefficient matrix and augmented matrix of the vector equation

$x_1C_1 + x_2C_2 + \cdots + x_{k-1}C_{k-1} = C_k$

As we know that the system has a solution, or, equivalently, C_k is a linear combination of the preceding columns of A if and only if $rank\ (M_k) = rank(M_{k-1})$where $rank\ (M_k)$ means the number of pivots in an echelon form of M_k.

Now the first k column of the echelon matrix B is also an echelon form of Mk.

Hence, $rank\ (M_3) = rank(M_2) = 2$ and $rank\ (M_4) = 3$

Thus,C_3 is a linear combination of the preceding columns of A.

(c) The fact that the remaining columns C_1, C_2, C_4 are not linear combinations of their respective preceding columns also tells us that they are linearly independent. Thus, they form a basis of the column space of A. That is,

$basis\ of\ colsp(A)$: $\begin{bmatrix}1\\2\\3\\4\end{bmatrix}, \begin{bmatrix}2\\5\\7\\8\end{bmatrix}, \begin{bmatrix}3\\6\\11\\12\end{bmatrix}$

Observe that C_1, C_2, C_4 C4 may also be characterized as those columns of A that contain the pivots in any echelon form of A.

(d) Here we see that three possible definitions of the rank of A yield the same value.

(i) There are three pivots in B, which is an echelon form of A.

(ii) The three pivots in B correspond to the nonzero rows of B, which form a basis of the row space of A.

(iii) The three pivots in B correspond to the columns of A, which form a basis of the column space of A.

Thus, $rank\ (A) = 3$.

2.5.2 *APPLICATION TO HOMOGENEOUS SYSTEMS OF LINEAR EQUATIONS*

Consider a homogeneous system $AX = 0$ of linear equations over K with n unknowns.

As we know that the solution set W of a homogeneous system $AX = 0$ in n unknowns is a subspace of K^n, hence W has a dimension.

Theorem 1.7: The dimension of the solution space W of a homogeneous system $AX = 0$ is n_r , where n is the number of unknowns and r is the rank of the coefficient matrix A.

Proof. In the case where the system $AX = 0$ is in echelon form, it has precisely n _ r free variables, say x_{i_1} ; x_{i_2} ; …; $x_{i_{n-r}}$.

Let v_j be the solution obtained by setting $x_{i_j} = 1$ (or any nonzero constant) and the remaining free variables equal to 0.

As we clearly see that the solutions $v_1;\ v_2;\ \ldots;\ v_{n_r}$ are linearly independent.

Hence, they form a basis of the solution space W.

CHECK YOURB PROGRESS 2

1: Find a basis and the dimension of subspace $W = \left\{ \begin{bmatrix} a+b+2c \\ 2a+b+3c+d \\ b+c+3d \\ a+c+d \end{bmatrix} : a, b, c, d \text{ are real} \right\}$

2: Find the dimensions of the null space and the column space of $\begin{bmatrix} -1 & 6 & -1 & 1 & -7 \\ 1 & -1 & 2 & 2 & -1 \\ 2 & -4 & 4 & 7 & -4 \end{bmatrix}$

2.6 *SUM AND DIRECT SUMS*

Let U and W be subsets of a vector space V. The sum of U and W, written $U + W$, consists of all sums $u + w$ where $u \epsilon U$ and $w \epsilon W$. i.e.,

$U + W = \{v: v = u + w, where\ \ u\epsilon\ U$ and $w\epsilon\ W$

Now suppose U and W are subspaces of V.

Then one can easily show that $U \cap W$ is a subspace of V.

As we know that $U \cap W$ is also a subspace of V.

The following theorem relates the dimensions of these subspaces.

Theorem 1.8: If W_1 and W_2 are finite-dimensional subspaces of a vector space V, then W_1+W_2 is finite-dimensional and

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

Proof. As we know if W is a subspace of a finite-dirnensional vector space V,

every linearly independent subset of W is finite and is part of a (finite) basis

for W and $\dim W < \dim V$.

Hence $W_1 \cap W_2$ has a finite basis $\{a_1, \dots, a_k\}$ which is part of a basis

$\{a_1, \dots, a_k, b_1, \dots, b_m\}$ for W_1

and part of a basis

$\{a_1, \dots, a_k, c_1, \dots, c_n\}$ for W_2.

The subspace $W_1 + W_2$ is spanned by the vectors

$a_1, \ldots, a_k, b_1, \ldots, b_m, c_1, \ldots, c_n$

and these vectors form an independent set.

Let

$\sum x_i a_i + \sum y_j b_j + \sum z_r c_r = 0$

which implies

$-\sum z_r c_r = \sum x_i a_i + \sum y_j b_j$

Hence $\sum z_r c_r$ belong to W_1.

As $\sum z_r c_r$ also belongs to W_2 it follows that

$\sum z_r c_r = \sum d_i a_i$

for certain scalars $d_1, \ldots, d_k$.

As the set is independent, each of the scalars $z_r = 0$.

Therefore

$\sum x_i a_i + \sum y_j b_j = 0$ and because $\{a_1, \ldots, a_k, b_1, \ldots, b_m\}$ the set is also an independent set, each $x_i = 0$ and each $y_i = 0$.

Hence $\{a_1, \ldots, a_k, b_1, \ldots, b_m, c_1, \ldots, c_n\}$ is also a basis for $W_1 + W_2$.

Hence

$$\dim W_1 + \dim W_2 = (k + m) + (k + n)$$
$$= k + (m + k + n)$$

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

Direct Sums: The vector space V is said to be the direct sum of its subspaces U and W, denoted by $V = U \oplus W$ if every $v\ in\ V$ can be written in one and only one way as $v = u + w$ where $u \in U$ and $w \in W$.

General Direct Sums: The notion of a direct sum is extended to more than one factor in the obvious way. That is, V is the direct sum of subspaces $W_1; W_2; \ldots; W_r$, written

$V = W_1 \oplus W_2 \oplus \ldots \ldots \oplus Wr$

if every vector $v \in V$ can be written in one and only one way as $v = w_1 + w_2 + \cdots + w_r$

where $w_1 \in W_1; w_2 \in W_2; \ldots; w_r \in W_r$.

2.7 *COORDINATES*

Let V be an n −dimensional vector space over K with basis $S = \{u_1;\ u_2;\ \ldots;\ u_n\}$. Then any vector $v\ in\ V$ can be expressed uniquely as a linear combination of the basis vectors in S, say

$v = a_1u_1 + a_2u_2 + \_\_ + a_nu_n$

These n scalars $a_1;\ a_2;\ \ldots;\ a_n$ are called the coordinates of v relative to the basis S, and they form a vector $[a_1;\ a_2;\ \ldots;\ a_n]$ in K^n called the coordinate vector of v relative to S.

We denote this vector by $[v]_s$, or simply $[v]$ when S is understood.

Therefore,

$[v_s] = [a_1, a_2, \ldots, a_n]$

Ex. Consider the vector space $P_2(t)$ of polynomials of degree ≤ 2. The polynomials

$p_1 = t + 1$, $p_2 = t - 1$, $p_3 = (t - 1)^2 = t^2 - 2t + 1$ form a basis S of $P_2(t)$. Find the coordinates.

Proof. The coordinate vector $[v] of\ v = 2t^2 - 5t + 9$ relative to S is obtained as follows.

Set $v = xp_1 + yp_2 + zp_3$ using unknown scalars x, y, z, and simplify:

$$2t^2 - 5t + 9 = x(t + 1) + y(t - 1) + z(t^2 - 2t + 1)$$

$$= xt + x + yt - y + zt^2 - 2zt + z$$

$$= zt^2 + (x + y - 2z)t + (x - y + z)$$

Then set the coefficients of the same powers of t equal to each other to obtain the system

$z = 2$, $x + y - 2z = -5$, $x - y + z = 9$

The solution of the system is $x = 3, y = -4, z = 2$. Therefore,

$v = 3p_1 - 4p_2 + 2p_3$ and hence; $[v] = [3, -4, 2]$.

NOTE:

There is a geometrical interpretation of the coordinates of a vector v relative to a basis S for the real space R^n, which we illustrate using the basis S of R^3 in above example. First consider the space

R^3 with the usual x, y, z axes. Then the basis vectors determine a new coordinate system of R^3, say with x_0, y_0, z_0 axes, as shown in Fig.2. i.e.,

(1) The x_0-axis is in the direction of u_1 with unit length $||u_1||$.

(2) The y_0-axis is in the direction of u_2 with unit length $||u_2||$.

(3) The z_0-axis is in the direction of u_3 with unit length $||u_3||$.

Then each vector $v = (a, b, c)$ or, equivalently, the point $P(a, b, c)$ in R^3 will have new coordinates with respect to the new x_0, y_0, z_0 axes. These new coordinates are precisely $[v]_S$, the coordinates of v with respect to the basis S.

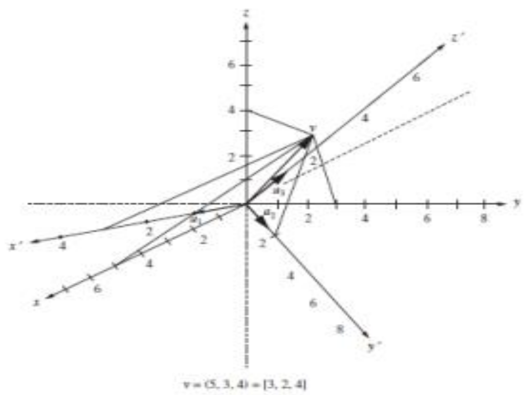


Fig. 2.1

2.8 *SUMMARY*

We discussed about basis and dimension of the vector space and also with the help of them solve number of illustrative examples.

2.9 *GLOSSARY*

- Basis
- Dimension
- Coordinates
- Direct sum

2.10 *REFERENCES*

- David C. Lay, Linear Algebra and its Application, 3rd Edition, Pearson Education Asia, India Reprint, 2007.
- Seymour Lipshutz and Marc Lipson, Schaum's outlines "linear Algebra", 3rd Edition, Mc Graw Hill Education, 2012.
- Gilbert Strang, Introduction to Linear Algebra (Gilbert Strang), 5th Edition

2.11 *SUGGESTED READINGS*

- P.B. Bhattacharya, S.K. Jain and S.R Nagpaul, First Course in Linear Algebra, Wiley Eastern Ltd., 1991.
- S.H Friedberg, A.J. Insel and L.E.Spence, Linear Algebra, 4th Edition, Prentice-Hall of India, New Delhi, 2004.
- R.K. Gupta, Linear Programming, Krishna Prakasan, 2010
- K. Hoffman and R. Kunze, Linear Algebra, 2nd Edition, Prentice-Hall of India, New Delhi, 2000.
- D. Kalman, A singularly valuable decomposition; the SVD of a matrix, The College Math. Journal, Vol .27, No.1, (1996).
- S. Kumaresan, Linear Algebra-A Geometric approach, Prentice-Hall of India, New Delhi, 2001.

2.12 *TERMINAL QUESTIONS*

(TQ-1) Define basis.
(TQ-2) Define Dimensions.
(TQ-3) Give example of basis.

Choose one of correct Choioce:

(TQ-4) Let U and W be subspaces of a vector space then

a) U+V is subspace of V
b) U and W are contained in U+W
c) $W + W = W$
d) All of the above

(TQ-5) The coordinate vector of $v = (a, b, c)$ in R^3 relative to (a) the usual basis $E = \{(1,0,0), (0,1,0), (0,0,1)\}$ is

a) $[a, b, -c]$

b) $[a/2, b/2, c/2]$

c) $[2a, 2b, 2c]$

d) $[a, b, c]$

(TQ-6) Does the vectors $v_1 = (-3, 7)$ and $v_2 = (5, 5)$ form a basis for R^2.

a). Data not complete

b). No

c). Yes

d). Not in R^2

(TQ-7) Are the vectors $v_1 = (2, 0, -1), v_2 = (4, 0, 7)$, and $v_3 = (-1, 1, 4)$ linearly independent inR^3?

a) linearly dependent

b) linearly independent

c) Data not complete

d) none of the above

2.13 *ANSWERS*

Answer of check your progress 1:

2: Linearly dependent

Answer of terminal question

(TQ-4) (d)

(TQ-5) (d)

(TQ-6) (c)

(TQ-7) (a)

UNIT-3: LINEAR TRANSFORMATION

CONTENTS:

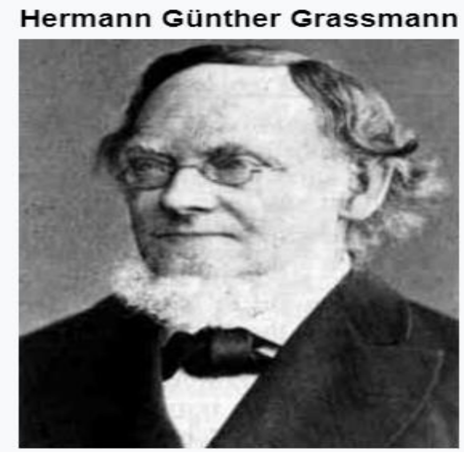
3.1 *INTRODUCTION*

By much, the idea of a matrix did not come before the idea of a linear transformation. Sylvester only used the name "matrix" to refer to an array of integers in 1850, despite the fact that matrices are implicitly mentioned in Cramer's work on determinants (1750), Euler's (1760), and Cauchy's (1829) work on quadratic forms. Though he did not work with them much, Cayley began to construct a theory about them in 1857, when he found that every matrix satisfies an equation of its own order and defined "characteristic values".

Around the same period, the idea of linear transformations is mentioned implicitly in Grassman's Ausdehnungslehre (1844), and particularly in Hamilton's work on quaternions (1845–1849), which heavily relied on quaternions' capacity to describe rotations in space. Motivated by the idea of describing forces in statics, Darboux presented the first axiomatization of vector spaces in

1875. This one looked very different from the current one. Furthermore, Peano provided an essentially modern axiomatization in 1888, but like Sylvester, he did not do much with it and few people took notice of it.

A key idea in mathematics is linear transformation, especially when it comes to linear algebra. It is a mapping that maintains the scalar multiplication and vector addition operations between two vector spaces. A function that takes a vector and converts it into another vector in a fashion that is consistent with the vector space's structure is known as a linear transformation.



15 April 1809- 26 September 1877

https://en.wikipedia.org/wiki/Hermann_Grassmann

German polymath Hermann Günther Grassmann (15 April 1809 – 26 September 1877) was renowned both as a mathematician and linguist in his day. In addition, he was a publisher, general scholar, and physicist. Not much was known about his mathematical efforts until he was in his sixties. His approach was both ahead of and better than the idea that is currently understood as a vector space. He presented the Grassmannian, a space that parameterizes every linear subspace of k dimensions in an n-dimensional vector space V.

3.2 *OBJECTIVE*

After reading this unit learners will be able to

- Understand the basic concept of linear transformation.
- Visualized the concept of homomorphism and isomorphism in vector space.
- Implement the important theorem of linear transformation.

3.3 *HOMOMORPHISM OF VECTOR SPACE OR LINEAR TRANSFORMATION*

Definition: Let $U(F)$ and $V(F)$ be two vector spaces. Then the mapping $f : U \to V$ is called a homomorphism or a linear transformation of U into V if they satisfy the following properties,

(*i*) $f(\alpha+\beta) = f(\alpha) + f(\beta), \forall \alpha, \beta \in U$

(*ii*) $f(a\alpha) = af(\alpha) \forall \alpha \in U$

The conditions (i) and (ii) can also be combined into the single condition i.e.,

$f(a\alpha + b\beta) = af(\alpha) + bf(\beta), \forall a,b \in F \text{ and } \forall \alpha, \beta \in U$

If f is a homomorphism of U into V, then V is called a homomorphic image of U.

Theorem 1: If f be a homomorphism of $U(F)$ into $V(F)$, then

(*i*) $f(0) = 0'$ where 0 and $0'$ are the zero's of vector U and V respectively.

(*ii*) $f(-\alpha) = -f(\alpha) \forall \alpha \in U$

Proof (*i*): Let $\alpha \in U$. Then $f(\alpha) \in V$. Since $0'$ is the zero vector of V, therefore

$f(\alpha) + 0' = f(\alpha) = f(\alpha + 0) = f(\alpha) + f(0).$

Now V is an abelian group with respect to addition of vectors.

$\therefore f(\alpha) + 0' = f(\alpha) + f(0)$

$\Rightarrow 0' = f(0)$ [By left cancellation rule]

(*ii*) If $\alpha \in U$, then $\alpha \in U$. Also we have

$0' = f(0) = f[\alpha + (-\alpha)] = f(\alpha) + f(-\alpha)$.

Now $f(\alpha) + f(-\alpha) = 0' \Rightarrow f(-\alpha) =$ additive inverse of $f(\alpha)$

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

Another definition of linear transformation:

Definition: Let $U(F)$ and $V(F)$ be two vector spaces over the same field F. A linear transformation from U into V is a function T from U into V such that

$T(a\alpha + b\beta) = aT(\alpha) + bT(\beta)$ …. (1)

for all α, β in U and for all $a, b \in F$.

The condition (1) is also called linearity property.

Linear operator: Let $V(F)$ be a vector space. A linear operator on V is a function T from V into V such that

$T(a\alpha + b\beta) = aT(\alpha) + bT(\beta)$, for all $\alpha, \beta \in V$ *and* $a, b \in F$

Thus T is a linear operator on V if T is linear transformation from V into V itself.

Example 1: The function $T : V_3(R) \to V_2(R)$

Defined by $T(a,b,c) = (a,b) \forall a,b \in R$ is a linear transformation from $V_3(R)$ into $V_2(R)$.

Let $\alpha = (a_1, b_1, c_1), \beta = (a_2, b_2, c_2) \in V_3(R)$

If $a, b \in R$, then

$$
\begin{aligned}
T(a\alpha + b\beta) &= T[a(a_1, b_1, c_1) + b(a_2, b_2, c_2)] \\
&= T(aa_1 + ba_2, ab_1 + bb_2, cc_1 + bc_2) \\
&= (aa_1 + ba_2, ab_1 + bb_2) \qquad \text{[by def. of } T\text{]} \\
&= (aa_1, ab_1) + (ba_2, bb_2) \\
&= a(a_1, b_1) + b(a_2, b_2) \\
&= a(a_1, b_1, c_1) + b(a_2, b_2, c_2) \\
&= aT(\alpha) + bT(\beta).
\end{aligned}
$$

Hence T is a linear transformation from $V_3(R)$ into $V_2(R)$

Example 2: Let $V(F)$ be the vector space of all $m \times n$ matrices over the field F. Let P be a fixed $m \times m$ matrix over F, and let Q be a fixed $n \times n$ matrix over F. The correspondence T from V into V defined by

$$T(A) = PAQ \ \forall \ A \in V$$

is a linear operator on V.

If A is an $m \times n$ matrix over the field F, then PAQ is also an $m \times n$ matrix over the field F. Therefore T is a function from V into V. Now let $A, B \in V$ and $a, b \in F$. Then

$T(aA+bB) = P(aA+bB)Q$ [By definition of T]

$= (aPA+bPB)Q = aPAQ + bPBQ = aT(A) + bT(B)$

So, T is a linear transformation from V into V. Thus T is a linear operator on V.

3.4 *SOME SPECIAL LINEAR TRANSFORMATION*

Some important linear transformation:

1. Zero transformation: Let $U(F)$ and $V(F)$ be two vector spaces. The function T, from U into V defined by, $T(\alpha) = 0$ (from zero vector of V)$\forall \alpha \in U$, is a linear transformation from U into V. Let $\alpha, \beta \in U$ and $a, b \in F$. Then $a\alpha + b\beta \in U$.

We have $T(a\alpha + b\beta) = 0$

$$a0 + b0 = aT(\alpha) + bT(\beta).$$

$\therefore T$ is a linear transformation and we will denote it by $\hat{0}$.

2. Identity transformation: Let $V(F)$ be a vector space. The function I from V into V defined by $I(\alpha) = \alpha \ \forall \ \alpha \in V$ is a linear transformation from V into V.

If $\alpha, \beta \in V$ and $a, b \in F$, then $a\alpha + b\beta \in V$ and we have

$I(a\alpha + b\beta) = aI(\alpha) + bI(\beta)$

$\therefore I$ is a linear transformation from V into V. This transformation is called as identity operator on V and denoted by I.

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

Since $T(\alpha) \in V \Rightarrow -[T(\alpha)] \in V$, therefore $-T$ is a function from U into V.

Let $\alpha, \beta \in U$ and $a, b \in F$. Then $a\alpha + b\beta \in U$ and we have

$(-T)(a\alpha + b\beta) = -[T(a\alpha + b\beta)]$ [By definition]

$= -[aT(\alpha) + bT(\beta)]$ [$\therefore T$ is a linear transformation]

$= a[-T(\alpha)] + b[-T(\beta)] = a[(-T)\alpha] + b[(-T)\beta].$

$\therefore T$ is a linear transformation from U into V. The linear transformation $-T$ is called the negative of the linear transformation T.

Some properties of linear transformation:

Theorem 2: Let T be a linear transformation from a vector space $U(F)$ into a vector space $V(F)$. Then

(i) $T(0) = 0^{'}$ where $0^{'}$ on the left hand side is zero vector of U and 0 on the right hand side is zero vector of V.

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

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

(iv) $T(a_1\alpha_1 + a_2\alpha_2 + ... + a_n\alpha_n) = a_1T(\alpha_1) + a_2T(\alpha_2) + ... + a_nT(\alpha_n)$

Where $\alpha_1, \alpha_2, ..., \alpha_n \in U$ and $a_1, a_2, ..., a_n \in F$

Proof (i): Let $\alpha \in U$. Then $T(\alpha) \in V$. We have

$T(\alpha) + 0 = T(\alpha)$ [$\therefore 0$ is zero vector space of V and $T(\alpha) \in V$]

$= T(\alpha + 0)$ [$\therefore 0$ is zero vector space of U]

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

Now in the vector space V, we have

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

$\Rightarrow 0 = T(0),$ by left cancellation law for addition in V.

Note: When we write $T(0) = 0$, there should be no confusion about the vector 0. Here T is a function from U into V. Therefore if $0 \in U,$ then its image under T i.e., $T(0) \in V$. Thus in $T(0) = 0,$ the zero on the right hand side is zero vector of V.

(ii) We have $T[\alpha + (-\alpha)] = T(\alpha) + T(-\alpha)$ [$\therefore T$ is a linear transformation]

But $T[\alpha + (-\alpha)] = T(0) = 0 \in V$ [By (i)]

Thus in $V,$ we have

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

$\Rightarrow T(\alpha) = -T(-\alpha)$

(iii) $T(\alpha - \beta) = T[\alpha + (-\beta)]$

$= T(\alpha) + T(-\beta)$ [$\because T$ is a linear transformation]

$= T(\alpha) + T(-\beta)$

$= T(\alpha) - T(\beta)$

(iv) We shall prove this result using induction method on n, the number of vectors in the linear combination $a_1\alpha_1 + a_2\alpha_2 + ... + a_n\alpha_n$.

Suppose $T(a_1\alpha_1 + a_2\alpha_2 + ... + a_{n-1}\alpha_{n-1}) = a_1T(\alpha_1) + a_2T(\alpha_2) + ... + a_{n-1}T(\alpha_{n-1})$ … (1)

Then, $T(a_1\alpha_1 + a_2\alpha_2 + ... + a_{n-1}\alpha_{n-1})$

$= T[(a_1\alpha_1 + a_2\alpha_2 + ... + a_{n-1}\alpha_{n-1}) + a_n\alpha_n]$

$= [a_1T(\alpha_1) + a_2T(\alpha_2) + ... + a_{n-1}T(\alpha_{n-1})] + a_nT(\alpha_n)$ by (1)

$= a_1T(\alpha_1) + a_2T(\alpha_2) + ... + a_{n-1}T(\alpha_{n-1}) + a_nT(\alpha_n)$

Now the proof is complete by induction method. Since the result is true when the number of vectors in the linear combination is 1.

Example 1: Show that the mapping $T: V_3(R) \to V_2(R)$ defined as

$T(a_1, a_2, a_3) = (3a_1 - 2a_2 + a_3, a_1 - 3a_2 - 2a_3)$ is a linear transformation from $V_3(R)$ to $V_2(R)$

Proof: Let $\alpha = (a_1, a_2, a_3), \beta = (b_1, b_2, b_3) \in V_3(R)$.

Then $T(\alpha) = T(a_1, a_2, a_3) = (3a_1 - 2a_2 + a_3, a_1 - 3a_2 - 2a_3)$

And $T(\beta) = T(b_1, b_2, b_3) = (3b_1 - 2b_2 + b_3, b_1 - 3b_2 - 2b_3)$.

Let $a, b \in R$. Then $a\alpha + b\beta \in V_3(R)$. We have

$T(a\alpha + b\beta) = T[a(a_1, a_2, a_3) + b(b_1, b_2, b_3)]$

$= T(aa_1 + bb_1, aa_2 + bb_2, aa_3 + bb_3)$

$= (3(3aa_1 + bb_1) - 2(aa_2 + bb_2) + aa_3 + bb_3, aa_1 + bb_1 - 3(aa_2 + bb_2) - 2(aa_3 + bb_3))$

$= (a(3a_1 - 2a_2 + a_3) + b(3b_1 - 2b_2 + b_3), a(a_1 - 3a_2 - 2a_3) + b(b_1 - 3b_2 - 2b_3)$

$= a(3a_1 - 2a_2 + a_3, a_1 - 3a_2 - 2a_3) + b(3b_1 - 2b_2 + b_3, b_1 - 3b_2 - 2b_3)$

$= aT(\alpha) + bT(\beta)$

Example 2: Show that the mapping $T: V_2(R) \to V_3(R)$ defined as

$$T(a, b) = (a + b, a - b, b)$$

is a linear transformation from $V_2(R)$ into $V_3(R)$.

Solution: Let the vectors $\alpha = (a_1, b_1), \beta = (a_2, b_2) \in V_2(R)$.

Then $T(\alpha) = T(a_1, b_1) = (a_1 + b_1, a_1 - b_1, b_1)$ and $T(\beta) = (a_2 + b_2, a_2 - b_2, b_2)$.

Also let $a, b \in R$. Then $a\alpha + b\beta \in V_2(R)$ and

$T(a\alpha + b\beta) = T[a(a_1, b_1) + b(a_2, b_2)]$

$= T(aa_1 + ba_2, ab_1 + bb_2)$

$= (aa_1 + ba_2 + ab_1 + bb_2, aa_1 + ba_2 - ab_1 - bb_2, ab_1 + bb_2)$

$= a(a_1 + b_1, a_1 - b_1, b_1) + b(a_2 + b_2, a_2 - b_2, b_2)$

$= aT(\alpha) + bT(\beta)$

$\therefore T$ is a linear transformation from $V_2(R)$ into $V_3(R)$.

3.5 *ISOMORPHISM OF VECTOR SPACE*

Definition: Let $U(F)$ and $V(F)$ be two vector spaces. Then a mapping $f: U \to V$ is called an isomorphism of U onto V if

(i) f is one-one

(ii) f is onto

(iii) $f(a\alpha + b\beta) = af(\alpha) + bf(\beta) \forall a, b \in F, \alpha, \beta \in U$

Also then the two vector spaces U and V are said to be isomorphic and symbolically we write $U(F) \cong V(F)$.

The vector space $V(F)$ is also called the isomorphic image of the vector space $U(F)$. If f is homomorphism of $U(F)$ into $V(F)$, then f will becomes an isomorphism of U into V if f is one-one. Also in addition if f is onto V, then f will become an isomorphism of U onto V.

Isomorphism of finite dimensional vector space:

Theorem 1: Two finite dimensional vector spaces over the same field are isomorphic if and only if they are of the same dimension.

Proof: First suppose that $U(F)$ and $V(F)$ are two finite dimensional vector spaces each of dimension n. Then to prove that $U(F) \cong V(F)$.

Let the sets of vectors $\{\alpha_1, \alpha_2, \alpha_3, ..., \alpha_n\}$ and $\{\beta_1, \beta_2, \beta_3, ..., \beta_n\}$ are the bases of U and V respectively.

Any vector $\alpha \in U$ can be uniquely expressed as

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

Let $f : U \to V$ be defined by

$f(\alpha) = a_1\beta_1 + a_2\beta_2 + ... + a_n\beta_n$.

Since in the expression of α as a linear combination of $\{\alpha_1, \alpha_2, \alpha_3, ..., \alpha_n\}$ the scalars $a_1, a_2, ..., a_n$ are unique, therefore the mapping f is well defined.

i.e., $f(\alpha)$ is a unique element of V.

f **is one-one:** We have

$f(a_1\alpha_1 + a_2\alpha_2 + ... + a_n\alpha_n) = f(b_1\alpha_1 + b_2\alpha_2 + ... + b_n\alpha_n)$

$\Rightarrow a_1\beta_1 + a_2\beta_2 + ... + a_n\beta_n = b_1\beta_1 + b_2\beta_2 + ... + b_n\beta_n$

$\Rightarrow (a_1 - b_1)\beta_1 + (a_2 - b_2)\beta_2 + ... + (a_n - b_n)\beta_n = 0'$ [zero vector of V]

$\Rightarrow a_1 - b_1 = 0, a_2 - b_2 = 0, ..., a_n - b_n = 0$ because

$\beta_1, \beta_2, ..., \beta_n$ are linearly independent

$\Rightarrow a_1 = b_1, a_2 = b_2, ..., a_n = b_n$

$\Rightarrow a_1\alpha_1 + a_2\alpha_2 + ... + a_n\alpha_n = b_1\alpha_1 + b_2\alpha_2 + ... + b_n\alpha_n$

$\therefore f$ is one-one.

f **is linear transformation:** We have

$f[a(a_1\alpha_1 + a_2\alpha_2 + ... + a_n\alpha_n) + b(b_1\alpha_1 + b_2\alpha_2 + ... + b_n\alpha_n)]$

$= f[(aa_1 + bb_1)\alpha_1 + (aa_2 + bb_2)\alpha_2 + ... + (aa_n + bb_n)\alpha_n)]$

$= (aa_1 + bb_1)\beta_1 + (aa_2 + bb_2)\beta_2 + ... + (aa_n + bb_n)\beta_n)]$

$= a(a_1\beta_1 + a_2\beta_2 + ... + a_n\beta_n) + b(b_1\beta_1 + b_2\beta_2 + ... + b_n\beta_n)$

$= af(a_1\alpha_1 + a_2\alpha_2 + ... + a_n\alpha_n) + bf(b_1\alpha_1 + b_2\alpha_2 + ... + b_n\alpha_n)$

$\therefore f$ is linear transformation.

Hence f is an isomorphism of U into V.

Thus $U \cong V$

Conversely, Let $U(F)$ and $V(F)$ be two isomorphic finite dimensional vector spaces. Now we have to prove that $\dim U = \dim V$.

Let $\dim U = n$. Let $S = \{\alpha_1, \alpha_2, ..., \alpha_n\}$ be a basis of U. If f is an isomorphism of U onto V, we shall show that $S^{'} = \{f(\alpha_1), f(\alpha_2), ..., f(\alpha_n)\}$ is a basis of V. Then V will also be a finite dimensional n. First we will show that $S^{'}$ is linearly independent.

Let $a_1 f(\alpha_1) + a_2 f(\alpha_2) + ... a_n f(\alpha_n) = 0^{'}$ (Zero vector of V)

$\Rightarrow f(a_1\alpha_1 + a_2\alpha_2 + ... a_n\alpha_n) = 0^{'}$ [$\because f$ is a linear transformation]

$\Rightarrow a_1\alpha_1 + a_2\alpha_2 + ... a_n\alpha_n = 0$ [$\because f$ is one-one and $f(0) = 0^{'}$, where $0 \in U$]

$\Rightarrow a_1 = 0, a_2 = 0, ... a_n = 0$ since $\alpha_1, \alpha_2, ..., \alpha_n$ are linearly independent. Hence $S^{'}$ is linearly independent.

Now we have only to prove that $L(S^{'}) = V$. For it let any vector $\beta \in V$ can be expressed as a linear combination of the vectors of the set $S^{'}$. Since f is onto V, therefore $\beta \in V \Rightarrow$ there exists $\alpha \in U$ such that $f(\alpha) = \beta$.

Let $\alpha = c_1\alpha_1 + c_2\alpha_2 + ... + c_n\alpha_n$.

Then $\beta = f(\alpha) = f(c_1\alpha_1 + c_2\alpha_2 + ... + c_n\alpha_n)$

$= c_1 f(\alpha_1) + c_2 f(\alpha_2) + ... + c_n f(\alpha_n)$

Thus β is a linear transformation of the vector of $S^{'}$.

Hence $V = L(S^{'})$.

$\therefore S^{'}$ is a basis of V. Since $S^{'}$ contains n vectors, therefore $\dim V = n$

Note: While proving the converse, we have proved that if f is an isomorphism of U onto V, then f maps a basis of U onto a basis of V.

Theorem 2: Every $n-$dimensional vector space $V(F)$ is isomorphic to $V_n(F)$.

Proof: Let $\{\alpha_1, \alpha_2, ..., \alpha_n\}$ be any basis of $V(F)$. Then every vector $\alpha \in V$ can be uniquely expressed as

$\alpha = a_1\alpha_1 + a_2\alpha_2 + ... a_n\alpha_n, a_i \in F$

The ordered $n-$tuple $(a_1, a_2, ..., a_n) \in V_n(F)$.

Let $f: V(F) \to V_n(F)$ be defined by $f(\alpha) = (a_1, a_2, ..., a_n)$.

Since in the expression of α as a linear combination of $\alpha_1, \alpha_2, ..., \alpha_n$ the scalars $a_1, a_2, ..., a_n$ are unique, therefore $f(\alpha)$ is a unique element of $V_n(F)$ and thus the mapping f is well defined.

f **is one-one:** Let $\alpha = a_1\alpha_1 + a_2\alpha_2 + ... + a_n\alpha_n$ and $\beta = b_1\alpha_1 + b_2\alpha_2 + ... + b_n\alpha_n$ be any two elements of V. We have $f(\alpha) = f(\beta)$

$\Rightarrow f(a_1\alpha_1 + a_2\alpha_2 + ... + a_n\alpha_n) = f(b_1\alpha_1 + b_2\alpha_2 + ... + b_n\alpha_n)$

$\Rightarrow (a_1, a_2, ..., a_n) = (b_1, b_2, ..., b_n)$

$\Rightarrow a_1 = b_1, a_2 = b_2, ..., a_n = b_n$

$\Rightarrow \alpha = \beta$

Hence, f is one-one.

f **is onto** $V_n(F)$**:** Let $(a_1, a_2, ..., a_n)$ be any element of $V_n(F)$. Then there exists an element $a_1\alpha_1 + a_2\alpha_2 + ... + a_n\alpha_n \in V(F)$ such that $f(a_1\alpha_1 + a_2\alpha_2 + ... + a_n\alpha_n) = (a_1, a_2, ..., a_n)$.

$\therefore f$ is onto $V_n(F)$.

f **is linear transformation:** If $a, b \in F$ and $\alpha, \beta \in V(F)$ we have

$f(a\alpha + b\beta) = f[a(a_1\alpha_1 + a_2\alpha_2 + ... + a_n\alpha_n) + b(b_1\alpha_1 + b_2\alpha_2 + ... + b_n\alpha_n)]$

$= f[(aa_1 + bb_1)\alpha_1 + (aa_2 + bb_2)\alpha_2 + ... + (aa_n + bb_n)\alpha_n]$

$= (aa_1 + bb_1, aa_2 + bb_2, ..., aa_n + bb_n)$

$= (aa_1, aa_2, ...aa_n) + (bb_1, bb_2, ..., bb_n)$

$= a(a_1, a_2, ...a_n) + b(b_1, b_2, ..., b_n)$

$= af(a_1\alpha_1 + a_2\alpha_2 + ... + a_n\alpha_n) + bf(b_1\alpha_1 + b_2\alpha_2 + ... + b_n\alpha_n)$

$= af(\alpha) + bf(\beta)$

$\therefore f$ is a linear transformation.

$\therefore f$ is an isomorphism of $V(F)$ onto $V_n(F)$.

Hence $V(F) \cong V_n(F)$.

Solved Example

Example 1: Show that the mapping $f: V_3(F) \to V_2(F)$ defined by

$f(a_1, a_2, a_3) = (a_1, a_2)$

Is a homomorphism of $V_3(F)$ onto $V_2(F)$.

Solution: Let $\alpha = (a_1, a_2, a_3)$ and $\beta = (b_1, b_2, b_3)$ be any two elements of $V_3(F)$. Also let a, b be any two elements of F. We have

$f(a\alpha + b\beta) = f[a(a_1, a_2, a_3) + b(b_1, b_2, b_3)]$

$= a(a_1, a_2) + b(b_1, b_2) = af(a_1, a_2, a_3) + bf(b_1, b_2, b_3)$

$= af(\alpha) + bf(\beta)$

$\therefore\ f$ is a linear transformation.

To show that f is onto $V_2(F)$. Let (a_1, a_2) be any element of $V_2(F)$. Then $(a_1, a_2, 0) \in V_3(F)$ and we have $f(a_1, a_2, 0) = (a_1, a_2)$. Therefore f is onto $V_2(F)$.

Example 2: Let $V(R)$ be the vector space of all complex numbers $a + ib$ over the field of reals R and let T be a mapping from $V(R)$ to $V_2(R)$ defined as $T(a+ib) = (a,b)$. Show that T is an isomorphism.

Solution: T **is one-one:** Let $\alpha = a+ib, \beta = c+id$ be any two members of $V(R)$. Then $a,b,c,d \in R$.

We have

$T(\alpha) = T(\beta) \Rightarrow (a,b) = (c,d)$

$\Rightarrow a = c, b = d \Rightarrow a+ib = c+id$

$\Rightarrow \alpha = \beta$

$\therefore T$ is one-one.

T **is on-to:** Let (a,b) be an arbitrary member of $V_2(R)$. Then there exist a vector $a+ib \in V(R)$ such that $T(a+ib) = (a,b)$. Hence T is onto.

T **is linear transformation:** Let $\alpha = a+ib, \beta = c+id$ be any two members of $V(R)$ and k_1, k_2 be any two elements of field $R.$ Then

$k_1\alpha + k_2\beta = k_1(a+ib) + k_2(c+id) = (k_1a + k_2c) + i(k_1b + k_2d)$

We have

$T(k_1\alpha + k_2\beta) = (k_1a + k_2c) + i(k_1b + k_2d)$, by definition of T

$= (k_1a, k_2b) + (k_2c, k_2d) = k_1(a,b) + k_2(c,d)$

$= k_1T(a+ib) + k_2T(c+id)$

$= k_1T(\alpha) + k_2T(\beta)$

Hence T is a linear transformation.

Thus T is an isomorphism.

Example 3: If V is a finite dimensional vector space and f is an isomorphism of V into V, prove that f must map V onto V.

Proof: Let $V(F)$ be a n dimensional vector space. Let f be an isomorphism of V into V i.e., f is a linear transformation and f is one-one. Now we have to prove that f is onto V.

Let $S = \{\alpha_1, \alpha_2, ..., \alpha_n\}$ be a basis of V. First we will prove that

$S' = \{f(\alpha_1), f(\alpha_2), ..., f(\alpha_n)\}$ is also a basis of V. We claim that S' is linearly independent. For it let, $a_1 f(\alpha_1) + a_2 f(\alpha_2) + ... + a_n f(\alpha_n) = 0$ (zero vector of V)

$\Rightarrow f(a_1\alpha_1 + a_2\alpha_2 + ... + a_n\alpha_n) = 0$ [$\because f$ is linear transformation]

$\Rightarrow a_1\alpha_1 + a_2\alpha_2 + ... + a_n\alpha_n = 0$ [$\because f$ is one-one and $f(0) = 0$]

$\Rightarrow a_1 = 0, a_2 = 0, ..., a_n = 0$ since $\alpha_1, \alpha_2, ..., \alpha_n$ are linearly independent.

$\therefore S'$ is linearly independent.

Now V is of dimension n and S' is linearly independent subset of V containing n vectors. Therefore S' must be a basis of V. Therefore each vector in V can be expressed as a linear combination of the vectors belonging to S'.

Now we shall show that f is onto V. Let α be any element of V. Then there exist scalars $c_1, c_2, ..., c_n$ such that

$\alpha = c_1 f(\alpha_1) + c_2 f(\alpha_2) + ... + c_n f(\alpha_n)$

$= f(c_1\alpha_1 + c_2\alpha_2 + ... + c_n\alpha_n)$

Now $S = \{\alpha_1, \alpha_2, ..., \alpha_n\}$ and $f -$image of this element is α. Therefore f is onto V. Hence f is an isomorphism of V onto V.

Example 4: If V is a finite dimensional and f is a homomorphism of V onto V prove that f must be one-one and so, an isomorphism.

Solution: Let $V(F)$ be a finite dimensional vector space of dimension n. Let f be a homomorphism of V onto V i.e., f is a linear transformation and f is onto V. To prove that f is one-one.

Let $S = \{\alpha_1, \alpha_2, ..., \alpha_n\}$ be a basis of V. We shall first prove that $S' = \{f(\alpha_1), f(\alpha_2), ..., f(\alpha_n)\}$ is also a basis of V. We claim that $L(S') = V$. The proof is as follows:

Let α be any element of V. We shall show that α can be expressed as a linear combination of $f(\alpha_1), f(\alpha_2), ..., f(\alpha_n)$. Since f is onto V, therefore $\alpha \in V$ implies that there exist $\beta \in V$ such that $f(\beta) = \alpha$. Now β can be expressed as a linear combination of $\alpha_1, \alpha_2, ..., \alpha_n$. Let

$\beta = a_1\alpha_1 + a_2\alpha_2 + ... + a_n\alpha_n$

Then, $\alpha = f(\beta) = f(a_1\alpha_1 + a_2\alpha_2 + ... + a_n\alpha_n)$

$= a_1 f(\alpha_1) + a_2 f(\alpha_2) + ... + a_n f(\alpha_n)$

Thus α has been expressed as a linear combination of $f(\alpha_1), f(\alpha_2), ..., f(\alpha_n)$.

Therefore $L(S') = V$.

Since V is of finite dimension n and S' is a subset of V containing n vectors and $L(S') = V$, therefore S' must be a basis of V. Therefore each vector in V can be expressed as a linear combination of vectors belonging to S' and S' is linearly independent. Now we shall show that f is one-one. Let γ and δ be any two elements of V such that

$\gamma = c_1\alpha_1 + c_2\alpha_2 + ... + c_n\alpha_n,\ \delta = d_1\alpha_1 + d_2\alpha_2 + ... + d_n\alpha_n$

We have $f(\gamma) = f(\delta)$

$\Rightarrow f(c_1\alpha_1 + c_2\alpha_2 + ... + c_n\alpha_n) = f(d_1\alpha_1 + d_2\alpha_2 + ... + d_n\alpha_n)$

$\Rightarrow c_1 f(\alpha_1) + c_2 f(\alpha_2) + ... + c_n f(\alpha_n) = d_1 f(\alpha_1) + d_2 f(\alpha_2) + ... + d_n f(\alpha_n)$

$\Rightarrow (c_1 - d_1) f(\alpha_1) + (c_2 - d_2) f(\alpha_2) + ... + (c_n - d_n) f(\alpha_n) = 0$

$\Rightarrow c_1 - d_1 = 0, c_2 - d_2 = 0, ..., c_n - d_n = 0$

Since $f(\alpha_1), f(\alpha_2), ..., f(\alpha_n)$ are linearly independent

$\Rightarrow c_1 = d_1, c_2 = d_2, ..., c_n = d_n$

$\Rightarrow \gamma = \delta$

$\therefore f$ is one-one.

$\therefore f$ is an isomorphism of V onto V.

Example 5: If V is finite dimensional vector space and f is a homomorphism of V into itself which is not onto prove that there is some $\alpha \neq 0$ in V such that $f(\alpha) = 0$.

Solution: If f is a homomorphism of V into itself, then $f(0) = 0$. Suppose there is no non-zero vector α in V such that $f(\alpha) = 0$. Then f is one-one. Because

$f(\beta) = f(\gamma)$

$\Rightarrow f(\beta) - f(\gamma) = 0$

$\Rightarrow f(\beta - \gamma) = 0$

$\Rightarrow \beta - \gamma = 0 \Rightarrow \beta = \gamma$

Now V is finite dimensional and f is a linear transformation of V into itself. Since f is one-one, therefore f must be onto V. But it is given that f is not onto. Therefore our assumption is wrong. Hence there will be a non-zero vector α in V such that $\Rightarrow f(\alpha) = 0$.

Example 6: Define linear transformation of a vector space $V(F)$ into a vector space $W(F)$. Show that the mapping $T : (a, b) \to (a + 2, b + 3)$

of $V_2(R)$ into itself is not a linear transformation.

Solution: We have to prove that the mapping

$T : (a, b) \to (a + 2, b + 3)$

Of $V_2(R)$ into itself is not a linear transformation.

Take $\alpha = (1, 2)$ and $\beta = (1, 3)$ as two vectors of $V_2(R)$ and $a = 1, b = 1$ as two elements of the field R.

Then $a\alpha + b\beta = 1(1, 2) + 1(1, 3) = (1, 2) + (1, 3) = (2, 5)$

By the definition of the mapping T, we have

$T(a\alpha + b\beta) = T(2, 5) = (2 + 2, 5 + 3) = (4, 8)$ … (1)

Also $T(\alpha) = T(1, 2) = (1 + 2, 2 + 3) = (3, 5)$

And $T(\beta)=T(1,3)=(1+2,3+3)=(3,6)$.

$\therefore aT(\alpha)+bT(\beta)=1(3,5)+1(3,6)=(3,5)+(3,6)=(6,11)$ … (2)

From equation (1) and (2), we see that

$T(a\alpha+b\beta)\neq aT(\alpha)+bT(\beta)$

Hence T is not a linear transformation.

Example 7: Let f be a linear transformation from a vector space U into a vector space V. If S is a subspace of U, prove that $f(S)$ will be a subspace of V.

Solution: Since $U(F)$ and $V(F)$ are two vector space over the same field F. The mapping f is linear transformation of U into V i.e.,

$f:U\to V$ such that

$f(a\alpha+b\beta)=af(\alpha)+bf(\beta)\ \forall a,b\in F$ and $\alpha,\beta\in U$

Let S be a subspace of U. Then to prove that $f(S)$ is a subspace of V. Let $a,b\in F$ and $f(\alpha),f(\beta)\in f(S)$ *where* $\alpha,\beta\in S$.

Since S is a subspace of U, therefore $a,b\in F$ and $\alpha,\beta\in S\Rightarrow a\alpha+b\beta\in S$

$\Rightarrow f(a\alpha+b\beta)\in f(S)$

$\Rightarrow af(\alpha)+bf(\beta)\in f(S)$ [$\because f(a\alpha+b\beta)=af(\alpha)+bf(\beta)$]

Thus $a,b\in F$ and $f(\alpha),f(\beta)\in f(S)$

$\Rightarrow af(\alpha)+bf(\beta)\in f(S)$

Hence $f(S)$ is a subspace of V.

Example 8: If $f:U\to V$ is an isomorphism of the vector space U into the vector space V, then a set of vectors $\{f(\alpha_1),f(\alpha_2),...,f(\alpha_r)\}$ is linearly independent if and only if the set $\{\alpha_1,\alpha_2,...,\alpha_r\}$ is linearly independent.

Solution: $U(F)$ and $V(F)$ are two vector spaces over the same field F and f is an isomorphism of U into V i.e.,

$f:U\to V$ such that

f is $1-1$ and $f(a\alpha+b\beta)=af(\alpha)+bf(\beta)\ \forall\ a,b\in F$ and $\forall\ \alpha,\beta\in U$

Let $\{\alpha_1,\alpha_2,...,\alpha_r\}$ be a subset of U. First suppose that the vector $\alpha_1,\alpha_2,...,\alpha_r$ are linearly independent. Then to show that the vectors $f(\alpha_1),f(\alpha_2),...,f(\alpha_r)$ are also linearly independent.

We have

$a_1f(\alpha_1)+a_2f(\alpha_2)+...+a_rf(\alpha_r)=0$ where $a_1,a_2,...,a_r\in F$

$\Rightarrow f(a_1\alpha_1+a_2\alpha_2+...+a_r\alpha_r)=0$ [$\because f$ is linear information]

$\Rightarrow f(a_1\alpha_1+a_2\alpha_2+...+a_r\alpha_r)=f(0)$ [$\because f(0)=0$]

$\Rightarrow a_1\alpha_1+a_2\alpha_2+...+a_r\alpha_r=0$ [$\because f$ is 1-1]

$\Rightarrow a_1=0,a_2=0,...,a_r=0$ since the vectors $\alpha_1,\alpha_2,...,\alpha_r$ are linearly independent.

Hence the vector $f(\alpha_1), f(\alpha_2),..., f(\alpha_r)$ are also linearly independent.

Conversely suppose that the vectors $f(\alpha_1), f(\alpha_2),..., f(\alpha_r)$ are linearly independent. Then show that the vectors $\alpha_1, \alpha_2,..., \alpha_r$ are also linearly independent.

We have

$a_1\alpha_1 + a_2\alpha_2 + ... + a_r\alpha_r = 0$ where $a_1, a_2,..., a_r \in F$

$\Rightarrow f(a_1\alpha_1 + a_2\alpha_2 + ... + a_r\alpha_r) = f(0)$

$\Rightarrow a_1 f(\alpha_1) + a_2 f(\alpha_2) + ... + a_r f(\alpha_r) = 0$ [$\because f$ is linear information]

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

Since the vectors $f(\alpha_1), f(\alpha_2),..., f(\alpha_r)$ are linearly independent. Hence the vectors $\alpha_1, \alpha_2,..., \alpha_r$ are also linearly independent.

Check your progress

Problem 1: Verify that the mapping $T: F^3 \to F^3$ defined by, $T(x, y, z) = (x - y + z, 2x + y - z, -x - 2y)$ is a linear transformation.

Problem 2: Verify that the mapping $T: V_3(R) \to V_2(R)$ defined by $T(a,b,c) = (a-b, a-c)$ is a linear transformation.

Problem 3: Show that the mapping $T: R^2 \to R^3$ defined as $T(a,b) = (a-b, b-a, -a)$ is a linear transformation from R^2 into R^3.

3.6 *SUMMARY*

In this unit, we have learned about the important concept of linear algebra like, linear transformation, homomorphism and isomorphism in vector space. Given that they maintain a vector space's structure; linear transformations are advantageous. Therefore, under certain circumstances, a lot of qualitative evaluations of a vector space that is the domain of a linear transformation may automatically hold in the image of the linear transformation. These essential tools are very important to solve many matrices related problems. The overall summarization of this units are as follows:

- Two finite dimensional vector spaces over the same field are isomorphic if and only if they are of the same dimension.
- Every $n-$dimensional vector space $V(F)$ is isomorphic to $V_n(F)$.

3.7 *GLOSSARY*

- Linear transformation
- Homomorphism
- Isomorphism

3.8 *REFERENCES*

- Gel′fand I.M. Lectures on linear algebra (1989), Courier Corporation.
- Kenneth Hoffman & Ray Kunze, Linear Algebra (2nd edition)(2015), Prentice-Hall.
- David C. Lay, Linear Algebra and its Application (3rd Edition) (2007) Pearson Education Asia, India Reprint.
- Seymour Lipshutz and Marc Lipson, Schaum's outlines "Linear Algebra" (3rd Edition)(2012), Mc Graw Hill Education.
- J. N. Sharma and A. R. Vasistha, Linear Algebra (29th Edition) (1999), Krishna Prakashan.

3.9 *SUGGESTED READING*

- Minking Eie & Shou-Te Chang (2020), **A First Course In Linear Algebra, World Scientific.**
- Axler, Sheldon (2015), Linear algebra done right. Springer.
- https://nptel.ac.in/courses/111106051
- https://archive.nptel.ac.in/courses/111/104/111104137
- https://epgp.inflibnet.ac.in/
- https://old.sggu.ac.in/wp-content/uploads/2020/07/Linear-AlgebraTY_506.pdf

3.10 *TERMINAL QUESTION*

Long Answer Type Question:

1. Let $T: V_2(R) \to V_2(R)$ be defined as $T(a_1, b_1) = (b_1, a_1)$, show that T is an isomorphism.

2. If f is an isomorphism of a vector space V onto a vector space W, prove that f maps a basis of V onto a basis of W.

3. If $f: U \to V$ is an isomorphism of the vector space U into the vector space V, then a set of vectors $f(\alpha_1), f(\alpha_2), \ldots, f(\alpha_r)$ is linearly dependent in V if and only if the set $\alpha_1, \alpha_2, \ldots, \alpha_r$ is linearly dependent in U.

4. Prove that a finite dimensional vector space $V(R)$ with dimension $V = n$ is isomorphic to R^n.

5. Let V be a finite dimensional vector space. If $f: V \to V$ is a one-one linear transformation, show that f is an isomorphism of V onto itself.

6. If T is a linear operator on a finite dimensional vector space V, show that T is one-one if and only if T is onto.

7. Define the following.

(*i*) Linear transformation

(*ii*) Homomorphism

(*iii*) Isomorphism

Short answer type question:

1. Show that the mapping $T: F^3 \to F^3$ defined by,

$T(x, y, z) = (x - y + 2z, 2x + y, -x - 2y + 2z)$ is a linear transformation.

2. Which of the following functions $T: R^2 \to R^2$ are linear transformation

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

b. $T(a,b) = (b, a)$

c. $T(a,b) = (a+b, a)$

3. Show that the $T: R^3 \to R^3$ is a liner transformation defined by,

$$T(x, y, z) = (x + 2y - z, y + z, x + y - 2z)$$

4. Show that the mapping $T: V_4(R) \to V_3(R)$ defined by

$T(x, y, z, w) = (x - y + z + w, x + 2z - w, x + y + 3z - 3w)$ is a linear transformation.

5. Show that the mapping $T : V_2(R) \to V_2(R)$ defined by

$T(a,b) = (b,0) \forall a, b \in R$ is a linear transformation.

6. Show that the mapping $T : V_3(R) \to V_3(R)$ defined by,

$T(x, y, z) = (3x, x - y, 2x + +y + z) \forall (x, y, z) \in V_3(R)$ is a linear transformation.

7. Show that the mapping $T : R^3 \to R^3$ defined by $T(a,b,c) = (0, a, b) \forall a, b \in R$ is a linear transformation.

Fill in the blanks:

1: Zero transformation is a …………………….

2: Negative of a linear transformation is ………………

3: Identity transformation is a ……..………..

4: If T is a linear transformation then $T(\alpha - \beta) =$ ……………

5: Two finite dimensional vector spaces over the same field are isomorphic if and only if they are of the same ……………..

6: Every $n-$dimensional vector space $V(F)$ is isomorphic to ……………

3.11 *ANSWERS*

Answer of short answer type question

2. **(a)** T is a linear transformation. **(b)** T is a linear transformation

(c) T is a linear transformation.

Answer of fill in the blank question

1: Linear transformation

2: Linear transformation

3: Linear transformation

4: $T(\alpha) - T(\beta)$

5: Dimension

6: $V_n(F)$

UNIT-4: RANK NULLITY THEOREM

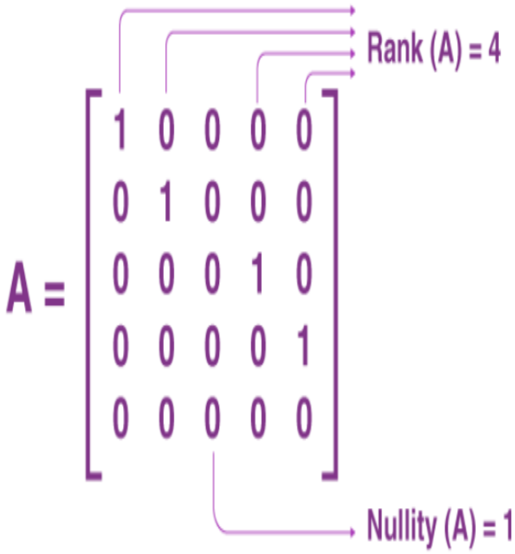
CONTENTS:

4.1 *INTRODUCTION*

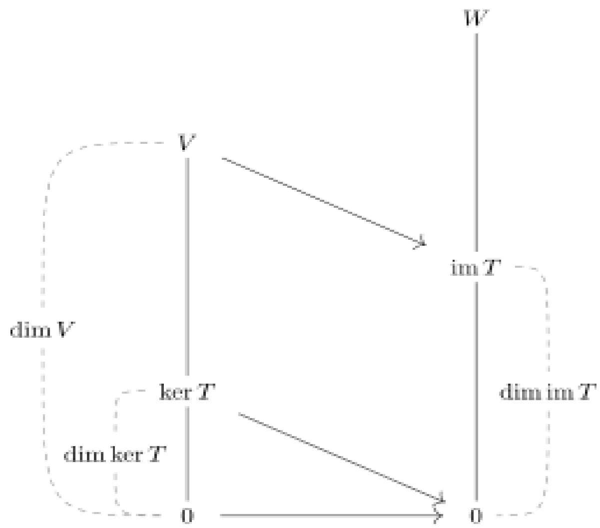
In 1878, Frobenius established a matrix's rank, and in 1884, Sylvester established a matrix's nullity. The rank–nullity theorem is a linear algebraic theorem that states:

- The dimension of the domain of a linear transformation f is the sum of the rank of f (the dimension of the image of f) and the nullity of f (the dimension of the kernel of f).
- The number of columns of a matrix M is the sum of the rank of M and the nullity of M.

It follows that either surjectivity or injectivity implies bijectivity for linear transformations of vector spaces of equal finite dimension.



https://byjus.com/maths/rank-and-nullity/



Rank–nullity theorem

https://en.wikipedia.org/wiki/Rank%E2%80%93nullity_theorem

4.2 *OBJECTIVES*

After the completion of this unit learners will be able to:

- Understand the concept of range and null space of a linear transformation
- Visualized the concept of rank and nullity

4.3 *RANGE OF A LINEAR TRANSFORMATION*

Definition: Let $U(F)$ and $V(F)$ be two vector spaces and let T be a linear transformation from U into V. Then the range of T written as $R(T)$ is the set of all vectors of β in V such that $\beta = T(\alpha)$ for some α in U.

Thus the range of T is the image set of U under T i.e.,

$$Range(T) = \{T(\alpha) \in V : \alpha \in U\}$$

Theorem 1: If $U(F)$ and $V(F)$ are two vector spaces and T is a linear transformation from U into V, then the range of T is a subspace of V.

Proof: Obviously $R(T)$ is a non-empty subset of V.

Let $\beta_1, \beta_2 \in R(T)$. Then there exist vectors α_1, α_2 in U such that $T(\alpha_1) = \beta_1, T(\alpha_2) = \beta_2$.

Let a, b be any elements of the field F. We have

$a\beta_1 + b\beta_2 = aT(\alpha_1) + bT(\alpha_2) = T(a\alpha_1 + b\alpha_2)$ [$\because T$ is linear transformation]

Now U is a vector space. Therefore $\alpha_1, \alpha_2 \in U$ and $a, b \in F \Rightarrow a\alpha_1 + b\alpha_2 \in U$

Consequently $T(a\alpha_1 + b\alpha_2) = a\beta_1 + a\beta_2 \in R(T)$.

Thus $a, b \in F$ and $\beta_1, \beta_2 \in R(T) \Rightarrow a\beta_1 + b\beta_2 \in R(T)$.

Therefore $R(T)$ is a subspace of V.

4.4 *NULL SPACE OF A LINEAR TRANSFORMATION*

Definition: Let $U(F)$ and $V(F)$ be two vector space and T is a linear transformation form U into V. Then the null space of T written as $N(T)$ is the set of all vectors α in U such that $T(\alpha) = 0$ (zero vector of V). Thus

$N(T) = \{\alpha \in U : T(\alpha) = 0 \in V\}$.

If we regard the linear transformation T from U into V as a vector space homomorphism of U into V, then the null space T is called the **kernel of** T.

Theorem 2: If $U(F)$ and $V(F)$ are two vector space and T is a linear transformation from U into V, then the kernel of T or the null space of T is a subspace of U.

Proof: Let $N(T) = \{\alpha \in U : T(\alpha) = 0 \in V\}$.

Since $T(0) = 0 \in V$, therefore at least $0 \in N(T)$. Thus $N(T)$ is non-empty subset of U.

Let $\alpha_1, \alpha_2 \in N(T)$. Then $T(\alpha_1) = 0$ and $T(\alpha_2) = 0$.

Let $a, b \in F$. Then $a\alpha_1 + b\alpha_2 \in U$ and

$T(a\alpha_1 + b\alpha_2) = aT(\alpha_1) + bT(\alpha_2)$ [$\because T$ is a linear transformation]

$= a0 + b0 = 0 + 0 = 0 \in V$

$\therefore a\alpha_1 + b\alpha_2 \in N(T)$

Thus $a, b \in F$ and $\alpha_1, \alpha_2 \in N(T) \Rightarrow a\alpha_1 + b\alpha_2 \in N(T)$. Therefore $N(T)$ is a subspace of U.

Theorem 1: Let T be a linear transformation from a vector space $U(F)$ into a vector space $V(F)$. If U is a finite dimensional, then the range of T is a finite dimensional subspace of V.

Proof: Since U is finite dimensional, therefore there exist a finite subset of U, say $\{\alpha_1, \alpha_2, ..., \alpha_n\}$ which spans U.

Let $\beta \in$Range of T. Then there exist α in U such that

$T(\alpha) = \beta$.

Now $\alpha \in U \Rightarrow \exists a_1, a_2, ..., a_n \in F$ such that

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

$\Rightarrow T(\alpha) = T(a_1\alpha_1 + a_2\alpha_2 + ... + a_n\alpha_n)$

$\Rightarrow \beta = a_1T(\alpha_1) + a_2T(\alpha_2) + ... + a_nT(\alpha_n)$ … (1)

Now the vectors $T(\alpha_1), T(\alpha_2), ..., T(\alpha_n)$ are in range of T. If β is any vector in the range of T, then from (1), we see that β can be expressed as linear combination of $T(\alpha_1), T(\alpha_2), ..., T(\alpha_n)$.

Therefore the range of T is spanned by the vectors $T(\alpha_1), T(\alpha_2), ..., T(\alpha_n)$.

Hence the range of T is finite dimensional.

4.5 *RANK AND NULITY OF A LINEAR TRANSFORMATION*

Definition: Let T be a linear transformation from a vector space $U(F)$ into a vector space $V(F)$ with U as finite dimensional. The rank of T denoted by $\rho(T)$ is the dimension of the range of T i.e.,

$$\rho(T) = \dim R(T)$$

The nullity of T denoted by $\nu(T)$ is the dimension of the null space of T i.e.,

$$\nu(T) = \dim R(T)$$

Theorem 2: Let U and V be vector space over the field F and let T be a linear transformation from U into V. Suppose that U is finite dimensional. Then

$$rank(T) + nulity(T) = \dim U$$

Proof: Let N be the null space of T. Then N is a subspace of U. Since U is finite dimensional, therefore N is finite dimensional. Let $\dim N = nulity(T) = k$ and let $\{\alpha_1, \alpha_2, ..., \alpha_k\}$ be a basis for N.

Since $\{\alpha_1, \alpha_2, ..., \alpha_k\}$ is linearly independent subset of U, therefore we can extend it to form a basis of U. Let $\dim U = n$ and let $\{\alpha_1, \alpha_2, ..., \alpha_k, \alpha_{k+1}, ..., \alpha_n\}$ be a basis for U.

The vectors $T(\alpha_1), T(\alpha_2), ..., T(\alpha_k), T(\alpha_{k+1}), ..., T(\alpha_n)$ are in range of T. We claim that $\{T(\alpha_{k+1}), T(\alpha_{k+2}), ..., T(\alpha_n)\}$ is a basis for the range of T.

(i) First we shall prove that the vectors
$\{T(\alpha_{k+1}), T(\alpha_{k+2}), ..., T(\alpha_n)\}$ span the range of T.
Let $\beta \in$ range of T. Then their exists $\alpha \in U$ such that $T(\alpha) = \beta$.
Now $\alpha \in U \Rightarrow \exists a_1, a_2, ..., a_n \in F$ such that
$\alpha = a_1\alpha_1 + a_2\alpha_2 + ... + a_n\alpha_n \in F$
$\Rightarrow T(\alpha) = T(a_1\alpha_1 + a_2\alpha_2 + ... + a_n\alpha_n)$
$\Rightarrow \beta = a_1T(\alpha_1) + a_2T(\alpha_2) + ... + a_kT(\alpha_k) + a_{k+1}T(\alpha_{k+1}) + ... + a_nT(\alpha_n)$
[
$\because \alpha_1, \alpha_2, ..., \alpha_k \in N \Rightarrow T(\alpha_1) = 0, ..., T(\alpha_k) = 0$]
$\therefore$ the vectors $T(\alpha_{k+1}), ..., T(\alpha_n)$ span the range of T.
(ii) Now we shall show that the vectors $T(\alpha_{k+1}), ..., T(\alpha_n)$ are linearly independent.
Let $c_{k+1}, ..., c_n \in F$ such that
$c_{k+1}T(\alpha_{k+1}) + ... + c_nT(\alpha_n) = 0$
$\Rightarrow T(c_{k+1}\alpha_{k+1}) + ... + c_n\alpha_n) = 0$
$\Rightarrow c_{k+1}\alpha_{k+1} + ... + c_n\alpha_n \in$ null space of T i.e., N
$\Rightarrow c_{k+1}\alpha_{k+1} + ... + c_n\alpha_n = b_1\alpha_1 + b_2\alpha_2 + ... + b_k\alpha_k$ for some $b_1, b_2, ..., b_k \in F$
[$\because$ each vector in N can be expressed as a linear combination of vectors $\alpha_1, \alpha_2, ..., \alpha_k$ forming a basis of N]

$\Rightarrow b_1\alpha_1 + b_2\alpha_2 + ... + b_k\alpha_k - c_{k+1}\alpha_{k+1} - ... - c_n\alpha_n = 0$

$\Rightarrow b_1 = b_2 = ... = b_k = c_{k+1} = ...c_n = 0$

[$\because \alpha_1, \alpha_2, ..., \alpha_k, \alpha_{k+1}, ..., \alpha_n$ are linearly independent being basis for U]

$\Rightarrow$the vector $T(\alpha_{k+1}) + ... + T(\alpha_n) = 0$ are linearly independent.

$\therefore$ The vector $T(\alpha_{k+1}), ..., T(\alpha_n) = 0$ form a basis of range of T .

$\therefore$ rank $T =$ dim of range of $T = n - k$

$\therefore$ rank $(T) +$ nullity $(T) = (n-k) + k = n = \dim U$.

Note: If in place of the vector space V, we take the vector space U i.e., if T is a linear transformation on an n dimensional vector space U , even then as a special case of the above theorem,

$$\rho(T) + \nu(T) = n.$$

Example 1: Find the range, rank, null-space and nullity of the linear transformation $T : V_2(R) \to V_3(R)$, defined by $T(a,b) = (a+b, a-b, b)$.

Solution: Since we have given that T is linear transformation from $V_2(R)$ to $V_3(R)$. Since $\{(1,0),(0,1)\}$ is a basis for $V_2(R)$.

We have $T(1,0) = (1+0, 1-0, 0) = (1,1,0)$

and $T(0,1) = (0+1, 0-1, 0) = (1,-1,1)$.

The vector $T(1,0), T(0,1)$ span the range of T . Thus the range of T is the subspace of $V_3(R)$ spanned by the vectors $(1,1,0),(1,-1,1)$.

Now the vectors $(1,1,0),(1,-1,1) \in V_3(R)$ are linearly independent because if $x, y \in R$, then

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

$\Rightarrow x(1,1,0) + y(1,-1,1) = (0,0,0)$

$\Rightarrow (x+y, x-y, y) = (0,0,0)$

$\Rightarrow (x+y, x-y, y) = (0,0,0)$

$\Rightarrow x+y=0, x-y=0, y=0 \Rightarrow x=0, y=0$

$\therefore$ the vectors $(1,1,0),(1,-1,1)$ form a basis for range of T. Hence rank $T =$ dim of range of $T = 2$

Nullity of $T =$ dim of $V_2(R) -$ rank $T = 2-2=0$

$\therefore$ null space of T must be the zero subspace of $V_2(R)$.

Otherwise: $(a,b) \in$ null space of T

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

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

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

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

$\Rightarrow a=0, b=0$

$\therefore (0,0)$ is the only element of $V_2(R)$ which belongs to null space of T.

$\therefore$ null space of T is the zero subspace of $V_2(R)$.

Example 2: Let T be the linear transformation from F^3 into F^3 defined by $T(x_1,x_2,x_3) = (x_1 - x_2 + 2x_3, 2x_1 + x_2 - x_3, -x_1 - 2x_2)$. Describe the null space of T.

Solution: Let $\alpha = (x_1,x_2,x_3), \beta = (y_1,y_2,y_3) \in F^3$. Then

$T(\alpha) = T(x_1,x_2,x_3), T(\beta) = T(y_1,y_2,y_3)$

$T(\alpha) = (x_1 - x_2 + 2x_3, 2x_1 + x_2 - x_3, -x_1 - 2x_2)$ and

$T(\beta) = (y_1 - y_2 + 2y_3, 2y_1 + y_2 - y_3, -y_1 - 2y_2)$

Also let $a,b \in F$ Then $a\alpha + b\beta \in F^3$ and

$a\alpha + b\beta = a(x_1, x_2, x_3) + b(y_1, y_2, y_3)$

$= (ax_1 + by_1, ax_2 + by_2, ax_3 + by_3)$

Now by definition of T , we have

$T(a\alpha + b\beta) = ([ax_1 + by_1] - [ax_2 + by_2] + 2[ax_3 + by_3], 2[ax_1 + by_1] + ax_2 + by_2 - [ax_3 + by_3],$

$-[ax_1 + by_1] - 2[ax_2 + by_2])$

$= (a[x_1 - x_2 + 2x_3] + b[y_1 - y_2 + 2y_3], a[2x_1 + x_2 - x_3] + b[2y_1 + y_2 - y_3], a[-x_1 - 2x_2] + b[-y_1 - 2y_2])$

$= (a(x_1 - x_2 + 2x_3, 2x_1 + x_2 - x_3, -x_1 - 2x_2) + b(y_1 - y_2 + 2y_3, 2y_1 + y_2 - y_3, -y_1 - 2y_2)$

$= aT(\alpha) + bT(\beta)$

$\therefore T$ is a linear transformation from F^3 into F^3 .

Now $(x_1, x_2, x_3) \in$ null space of T

$\Leftrightarrow T(x_1, x_2, x_3) = (0,0,0)$

$\Leftrightarrow (x_1 - x_2 + 2x_3, 2x_1 + x_2 - x_3, -x_1 - 2x_2) = (0,0,0)$

$$\Leftrightarrow \begin{aligned} x_1 - x_2 + 2x_3 &= 0, \\ 2x_1 + x_2 - x_3 &= 0, \\ -x_1 - 2x_2 + 0x_3 &= 0, \end{aligned} \qquad \ldots (1)$$

$\therefore$ the null space of T is the solution space of the system of linear homogeneous equation (1). Let A be the coefficient matrix of the equation (1). Then

$$A = \begin{bmatrix} 1 & -1 & 2 \\ 2 & 1 & -1 \\ -1 & -2 & 0 \end{bmatrix}$$ [Performing the elementary row operation $R_2 \to R_2 - 2R_1, R_3 \to R_3 + R_1$]

This last matrix is in the Echelon form. Its rank $A = 3 =$ the number of unknowns in the equations (1). Hence the equation (1). Hence the equations (1) have no linearly independent solutions. Therefore $x_1 = 0, x_2 = 0, x_3 = 0$ is the only solution of the equations (1). Thus $(0,0,0)$ is the only vector which belongs to the null space of T. Hence the null space of T is the zero subspace of F^3.

Example 3: Let V be the vector space of all $n \times n$ matrices over the field F, and let B be a fixed $n \times n$ matrices over the field F, and let B be a fixed $n \times n$ matrix if

$T(A) = AB - BA \forall A \in V$

Verify that T is a linear transformation from V into V.

Solution: If $A \in V$, then $T(A) = AB - BA \in V$ because $AB - BA$ is also an $n \times n$ matrix over the field F. Thus T is a function from V into V.

Let $A_1, A_2 \in V$ and $a, b \in F$. Then $aA_1 + bA_2 \in V$ and

$T(aA_1 + bA_2) = (aA_1 + bA_2)B - B(aA_1 + bA_2)$

$= aA_1B + bA_2B - aBA_1 - bBA_2 = a(A_1B - BA_1) + b(A_2B - BA_2)$

$= aT(A_1) + bT(A_2)$

$\therefore T$ is a linear transformation from V into V.

Example 4: Let V be an n-dimensional vector space over the field F and let T be a linear transformation from V into V such that the range and null space of T are identical. Prove that n is even. Also give an example of such a linear transformation.

Solution: Let N be the null space of T. Then N be the null space of T. Then N is also the range of T.

Now $\rho(T) + \upsilon(T) = \dim V$

i.e. Dimension of range of T + Dimension of null space of $T = \dim V = n$

i.e. $2 \dim N = n$ $[\because range\, of\ T = null\ space\, of\ T = N]$

i.e., n is even.

Example of such a transformation:

Let $T : V_2(R) \to V_2(R)$ be defined by

$T(a,b)=(b,0)\,\forall\, a,b\in R$.

Let $\alpha=(a_1,b_1),\beta=(a_2,b_2)\in V_2(R)$ and let $x,y\in R$

Then $T(x\alpha+y\beta)=T[x(a_1,b_1)+y(a_2,b_2)]$

$=T(xa_1+ya_2,xb_1+yb_2)=(xb_1+yb_2,0)$

$=(xb_1,0)+(yb_2,0)=x(b_1,0)+(yb_2,0)=x(b_1,0)+y(b_2,0)$

$=xT(a_1,b_1)+yT(a_2,b_2)=xT(\alpha)+yT(\beta)$

$\therefore T$ is a linear transformation from $V_2(R)$ into $V_2(R)$.

Now {(1,0), (0,1)} is a basis of $V_2(R)$.

We have $T(1,0)=(0,0)$ and $T(0,1)=(1,0)$

Thus the range of T is the subspace of $V_2(R)$ spanned by the vectors $(0,0)$ and $(1,0)$. The vector $(0,0)$ can be omitted from this spinning set because it is zero vector. Therefore the range of T is the subspace of $V_2(R)$ spanned by the vector $(1,0)$. Thus,

Range of $T=\{a(1,0):a\in R\}=\{(a,0):a\in R\}$.

Now let $(a,b)\in N$ (The null space of T).

Then $(a,b)\in N\Rightarrow T(a,b)=(0,0)\Rightarrow(b,0)=(0,0)\Rightarrow b=0$.

$\therefore$ null space of $T=\{(a,0):a\in R\}$.

Thus range of $T=$ null space of T.

Also we observe that $\dim V_2(R)=2$, which is even.

Example 5: Let V be a vector space and T is a linear transformation from V into V. Prove that the following two statements about T are equivalent.

(i) The intersection of the range of T and the null space of T is the zero subspace of V i.e., $R(T)\cap N(T)=\{0\}$.

(ii) $T[T(\alpha)]=0\Rightarrow T(\alpha)=0$.

Solution: First we shall show that (i) $\Rightarrow$(ii)

We have $T[T(\alpha)] = 0 \Rightarrow T(\alpha) \in N(T)$

$\Rightarrow T(\alpha) \in R(T) \cap N(T)$ $[\because \alpha \in V \Rightarrow T(\alpha) \in R(T)]$

$\Rightarrow T(\alpha) = 0$ because $R(T) \cap N(T) = \{0\}$.

Now we will show that (ii) $\Rightarrow$(i).

Let $\alpha \neq 0$ and $\alpha \in R(T) \cap N(T)$.

Then $\alpha \in R(T)$ and $\alpha \in N(T)$.

Since $\alpha \in N(T)$, therefore $T(\alpha) = 0$. … (1)

Also $\alpha \in R(T) \Rightarrow \exists \beta \in V$ such that $T(\beta) = \alpha$.

Now, $T(\beta) = \alpha$

$\Rightarrow T[T(\beta)] = T(\alpha) = 0$ [From (1)]

Thus $\exists \beta \in V$ such that $T[T(\beta)] = 0$ but $T(\beta) = \alpha \neq 0$.

This contradict the fact that the given hypothesis (ii).

Therefore there exist no $\alpha \in R(T) \cap N(T)$ such that $\alpha \neq 0$.

Hence $R(T) \cap N(T) = \{0\}$.

Check your progress

Problem 1: Check the rank and nullity of the matrix

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

Solution: Given

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

Let us reduce this in row reduced echelon form

Applying $R_2 \to R_2 + (-3)R_1$ and $R_3 \to R_3 + R_1$

$$A \sim \begin{bmatrix} 1 & 1 & 2 & 3 \\ 0 & 1 & -7 & -7 \\ 0 & -1 & 7 & 7 \end{bmatrix}$$

Applying $R_3 \to R_3 + R_2$

$$A \sim \begin{bmatrix} 1 & 1 & 2 & 3 \\ 0 & 1 & -7 & -7 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Applying $C_2 \to C_2 + C_1$, $C_3 \to C_3 + (-2)C_1$ and $C_4 \to C_4 + (-3)C_1$

$$A \sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -7 & -7 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Applying $C_3 \to C_3 + 7C_2$ and $C_4 \to C_4 + 7C_2$

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

Clearly, rank(A) = 2 and nullity(A) = 2

Therefore, rank(A) + nullity(2) = 2 + 2 = 4 = Number of columns.

Problem 2: Check the nullity of the matrix

$$A = \begin{bmatrix} 1 & 3 & 4 & 3 \\ 3 & 9 & 12 & 9 \\ 1 & 3 & 4 & 1 \end{bmatrix}$$

Solution: Given matrix

$$A = \begin{bmatrix} 1 & 3 & 4 & 3 \\ 3 & 9 & 12 & 9 \\ 1 & 3 & 4 & 1 \end{bmatrix}$$

Applying elementary operations, $R_{21}(-3)$ and $R_{31}(-1)$

$$A \sim \begin{bmatrix} 1 & 3 & 4 & 3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -2 \end{bmatrix}$$

Applying elementary operations, $R_3(-½)$

$$A \sim \begin{bmatrix} 1 & 3 & 4 & 3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Applying elementary operations, C_{24}, we get

$$A \sim \begin{bmatrix} 1 & 3 & 4 & 3 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Applying elementary operations, $C_{21}(-3)$, $C_{31}(-4)$ and $C_{41}(-3)$ we get

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

Thus, Rank of matrix A = 2 and Nullity = Number of columns − rank = 4 − 2 = 2.

4.6 *SUMMARY*

In this unit, we have learned about the one of the important concept in linear algebra name as rank and nullity theorem. After the completion of this unit these important about the rank and nullity:

- An invertible matrix has a rank equal to its order, and its nullity is equal to zero.
- In the row-reduced echelon form of the given matrix, rank is the number of leading columns or non-zero row vectors; nullity is the number of zero columns.
- The dimension of A's null space, also known as the kernel of A, determines a matrix's nullity.
- Assuming A is an invertible matrix, null space (A) has the value $\{0\}$.
- The number of non-zero eigenvalues in a matrix represents its rank, while the number of zero eigenvalues establishes the matrix's nullity.

4.7 *GLOSSARY*

- Range space
- Null space
- Rank
- Nullity

4.8 *REFERENCES*

- Gel′fand I.M. Lectures on linear algebra (1989), Courier Corporation.
- Kenneth Hoffman & Ray Kunze, Linear Algebra (2nd edition)(2015), Prentice-Hall.
- David C. Lay, Linear Algebra and its Application (3rd Edition) (2007) Pearson Education Asia, India Reprint.
- Seymour Lipshutz and Marc Lipson, Schaum's outlines "Linear Algebra" (3rd Edition)(2012), Mc Graw Hill Education.
- J. N. Sharma and A. R. Vasistha, Linear Algebra (29th Edition) (1999), Krishna Prakashan.

4.9 *SUGGESTED READING*

- Minking Eie & Shou-Te Chang (2020), **A First Course In Linear Algebra, World Scientific.**
- Axler, Sheldon (2015), Linear algebra done right. Springer.
- https://nptel.ac.in/courses/111106051
- https://archive.nptel.ac.in/courses/111/104/111104137
- https://byjus.com/maths/rank-and-nullity/
- https://old.sggu.ac.in/wp-content/uploads/2020/07/Linear-AlgebraTY_506.pdf

4.10 *TERMINAL QUESTION*

Long Answer Type Question:

1: If T is a linear transformation from U into V, then prove that the range of T is a subspace of V.

2: If T is a linear transformation from U into V, then prove that the null of T is a subspace of U.

3: Define the following.

(*i*) Range of linear transformation

(*ii*) Null space of linear transformation

(*iii*) Kernel of a linear transformation

4: State and prove the rank and nullity of a linear transformation.

5: Let V be a vector space and T is a linear transformation from V into V. Prove that the following two statements about T are equivalent.

(i) The intersection of the range of T and the null space of T is the zero subspace of V i.e., $R(T) \cap N(T) = \{0\}$.

(ii) $T[T(\alpha)] = 0 \Rightarrow T(\alpha) = 0$.

Short answer type question:

1: If T is a linear transformation from U into V. If U is a finite dimensional, then the range of T is a finite dimensional subspace of V.

2: Let T be the linear transformation from $V_3(F)$ into $V_3(F)$ defined by $T(x, y, z) = (x - y + 2z, 2x + y - z, -x - 2y)$. Describe the null space of T.

Fill in the blanks:

1: The number of linearly independent row or column vectors of a matrix is the …… of the matrix

2: The dimension of the null space or kernel of the given matrix is the ……….. of the matrix

3: For any matrix *A* of order *m* by *n*, rank(*A*) + ……………… = number of columns in *A*

4: The nullity of an invertible matrix is …………

4.11 *ANSWERS*

Answer of short answer type question:

2: null space of T is the zero subspace of $V_3(F)$.

Answer of fill in the blank question:

1: Rank

2: Nulity

3: nullity(A)

4: Zero

UNIT-5: CHANGE OF BASES

CONTENTS

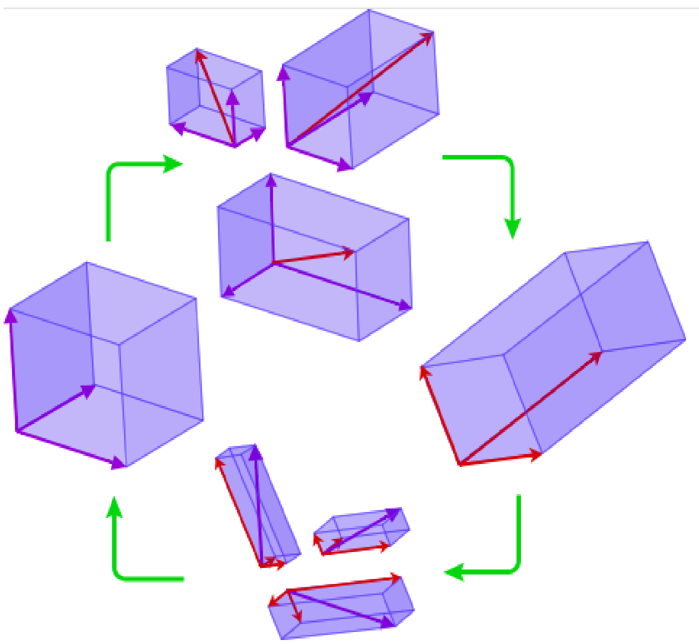
5.1 *INTRODUCTION*

A coordinate vector, which is a series of n scalars, can uniquely represent any element of a vector space using an ordered basis of a vector space of finite dimension n in mathematics. The coordinate vector representing a vector V on one basis differs, in general, from the coordinate vector representing V on the other basis when two separate bases are taken into account. Every assertion expressed in terms of coordinates with respect to one basis must be changed into an assertion expressed in terms of coordinates with respect to the other basis. This is known as a change of basis.

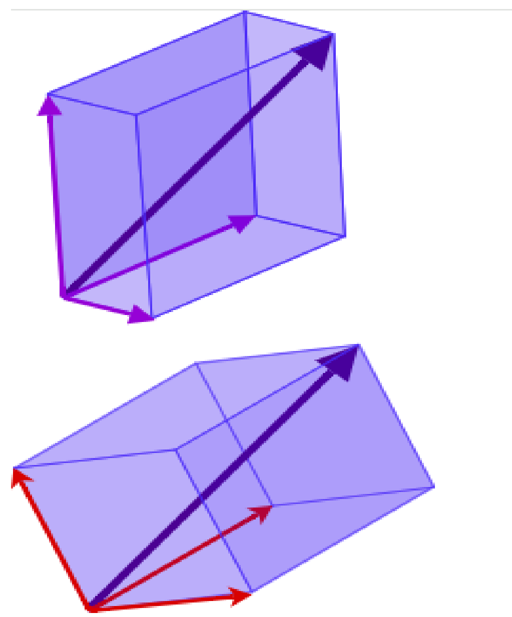
The change-of-basis formula, which describes the coordinates relative to one basis in terms of coordinates relating to the other basis, leads to this kind of conversion. This formula can be stated using matrices.

$$X_{old} = AX_{new}$$

where A is the change-of-basis matrix (also known as the transition matrix), which is the matrix whose columns are the coordinate vectors of the new basis vectors on the old basis; "old" and "new" refer to the firstly defined basis and the other basis, respectively; and X_{old} and X_{new} are the column vectors of the coordinates of the same vector on the two bases.



New vectors (red) are obtained by a linear combination of one basis of vectors (purple). Should they exhibit linear independence, these establish a novel basis. The linear transformation known as the change of basis is the result of the linear combinations connecting the first basis to the other.



A vector with two distinct bases (red and purple arrows).

https://en.wikipedia.org/wiki/Change_of_basis

5.2 *OBJECTIVES*

After reading this unit learners will be able to

- Understand the concept of change of basis.
- Implement the application of theorems related to change of basis.
- Understand the concept of similarity of matrices.
- Trace of matrices and determinant of linear transformation in finite dimensional vector space

5.3 *CHANGE OF BASIS*

Suppose V is an $n-$dimensional vector space over any field F. Let B and $B^{'}$ be two ordered basis for V. If α is any vector in V, then we are now interested to know its relation between the coordinates with respect to B and its coordinates with respect to $B^{'}$.

Theorem 1: Let $V(F)$ be an $n-$dimensional vector space and let B and $B^{'}$ be two ordered bases for V. Then there exist a unique $n \times n$ invertible matrix A having entries from F such that

(1) $[\alpha]_B = A[\alpha]_{B^{'}}$

(2) $[\alpha]_{B^{'}} = A^{-1}[\alpha]_B$

for every vector α in V.

Solution: Let $B = \{\alpha_1, \alpha_2, ..., \alpha_n\}$ and $B^{'} = \{\beta_1, \beta_2, ..., \beta_n\}$. Then there exists a unique linear transformation T from V into V such that

$$T(\alpha_j) = \beta_j,\ j = 1, 2, ..., n \qquad \dots (1)$$

Since T maps a basis B onto a basis $B^{'}$, therefore T is necessarily invertible. The matrix of T relative to B *i.e.*, $[T]_B$ will be a unique $n \times n$ matrix with element in F. Also this matrix will be invertible because T is invertible.

Let $[T]_B = A = [a_{ij}]_{n\times n}$. Then,

$$T(\alpha_j) = \sum_{i=1}^{n} a_{ij}\alpha_j,\ j = 1, 2, ..., n \qquad \dots (2)$$

Let $x_1, x_2, ..., x_n$ be the coordinates of α with respect to B and $y_1, y_2, ..., y_n$ be the coordinates of α with respect to $B^{'}$. Then

$$\alpha = y_1\beta_1 + y_2\beta_2 + ... + y_n\beta_n = \sum_{j=1}^{n} y_j\beta_j$$

$= \sum_{j=1}^{n} y_j T(\alpha_j)$ [From (1)]

$= \sum_{j=1}^{n} y_j \sum_{i=1}^{n} a_{ij}\alpha_i$ [From (2)]

$= \sum_{i=1}^{n} y_j \left(\sum_{j=1}^{n} a_{i\,j} y_j \right) \alpha_i$

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

$\therefore\ x_i = \sum_{j=1}^{n} a_{ij} y_j$ because the expression for α is a linear combination of elements of B is unique.

Now $[\alpha]_B$ is a column matrix of the type $n \times 1$. Also $[\alpha]_{B'}$ is a column matrix of the type $n \times 1$. The product matrix $A[\alpha]_{B'}$ will also be of the type $n \times 1$.

The i^{th} entry of $[\alpha]_B = x_i = \sum_{j=1}^{n} a_{i\,j} y_j$

$= i^{th}$ entry of $A[\alpha]_{B'}$.

$\therefore [\alpha]_B = A[\alpha]_{B'}$

$\Rightarrow A^{-1}[\alpha]_B = A^{-1}A[\alpha]_{B'}$

$\Rightarrow A^{-1}[\alpha]_B = I[\alpha]_{B'}$

$\Rightarrow A^{-1}[\alpha]_B = [\alpha]_{B'}$.

Note: The matrix $A = [T]_B$ is called the transition matrix from B to B'. It express the coordinates of each vector in V relative to B in terms of its coordinates relative to B'.

Working rule to write the transition matrix from one basis to another:

Let $B = \{\alpha_1, \alpha_2, \alpha_3, ..., \alpha_n\}$ and $B' = \{\beta_1, \beta_2, \beta_3, ..., \beta_n\}$ be two ordered bases for the *n*-dimensional vector space $V(F)$. Let A be the transition matrix from the basis B to the basis B'. Let we consider T be the linear transformation from V into V which maps the basis B onto the basis B'. Then A is the matrix of T relative to B i.e., $A = [T]_B$. So, in order to find the matrix A, we should first express each vector in the basis B' as a linear combination over F of the vectors in B. Thus we write the relations

$\beta_1 = a_{11}\alpha_1 + a_{12}\alpha_2 + ... + a_{n1}\alpha_n$

$\beta_2 = a_{12}\alpha_1 + a_{22}\alpha_2 + ... + a_{n2}\alpha_n$

...

...

$\beta_n = a_{1n}\alpha_1 + a_{2n}\alpha_2 + ... + a_{nn}\alpha_n$

Then the matrix $A=[a_{ij}]_{n\times n}$ i.e., A is the transpose of the matrix of coefficients in the above relations. Thus,

$$A=\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$$

Now suppose α is any vector in V. If $[\alpha]_B$ is the coordinate matrix of α relative to the basis B and $[\alpha]_{B'}$ its coordinate matrix relative to the basis B' then,

$[\alpha]_B = A[\alpha]_{B'}$

and $[\alpha]_{B'} = A^{-1}[\alpha]_B$.

Theorem 2: Let $B=\{\alpha_1,\alpha_2,\alpha_3,...,\alpha_n\}$ and $B'=\{\beta_1,\beta_2,\beta_3,...,\beta_n\}$ be two ordered bases for an $n-$dimensional vector space $V(F)$. If $(x_1,x_2,...,x_n)$ is an ordered set of n scalars, let

$\alpha=\sum_{i=1}^{n} x_i\alpha_i$ and $\beta=\sum_{i=1}^{n} x_i\beta_i$. Then show that, $T(\alpha)=\beta$,

Where, T is the linear operator on V defined by

$T(\alpha_i)=\beta_i, i=1,2,...,n$.

Proof: We have $T(\alpha)=T\left(\sum_{i=1}^{n} x_i\alpha_i\right)$

$$=\sum_{i=1}^{n} x_i T(\alpha_i) \qquad [\because T \text{ is linear}]$$

$$=\sum_{i=1}^{n} x_i\beta_i = \beta$$

5.4 *SIMILARITY OF MATRICES*

Definition: Let A and B be square matrices of order n over the field F. Then B is said to be similar to A if there exist an $n\times n$ invertible square matrix C with elements in F such that

$B=C^{-1}AC$

Theorem 2: The relation of similarity is an equivalence relation in the set of all $n\times n$ matrices over the field F.

Proof: If A and B are two $n\times n$ matrices over the field F, then B is said to be similar to A if there exists an $n\times n$ invertible matrix C over F such that,

$B=C^{-1}AC$.

Reflexive: Let A be any $n\times n$ matrix over F. We can write $A = I^{-1}AI$, where I is $n\times n$ unit matrix over F.

$\therefore$ A is similar to A because I is definitely invertible.

Symmetry: Let A be similar B. Then there exists an $n\times n$ invertible matrix P over F such that

$$A = P^{-1}BP$$

$\Rightarrow PAP^{-1} = P(P^{-1}BP)P^{-1}$

$\Rightarrow PAP^{-1} = B$

$\Rightarrow B = PAP^{-1}$

$\Rightarrow B = (P^{-1})^{-1}AP^{-1}$

[$\because P$ is invertible means P^{-1} is invertible and $(P^{-1})^{-1} = P$]

$\Rightarrow B$ is similar to A.

Transitive: Let A be similar to B and B be similar to C. Then

$A = P^{-1}BP$

and $B = Q^{-1}BQ$,

where P and Q are invertible $n\times n$ matrices over F.

We have $A = P^{-1}BP = P^{-1}(Q^{-1}CQ)P$

$A = (P^{-1}Q^{-1})C(QP)$

$A = (QP)^{-1}C(QP)$

[$\therefore$ P and Q are invertible means QP is invertible and $(QP)^{-1} = P^{-1}Q^{-1}$]

$\therefore A$ is similar to C.

Hence similarity is an equivalence relation on the set of $n\times n$ matrices over the field F.

Theorem 3: Similar matrices have the same determinant.

Proof: Let us consider the matrix B is similar to the matrix A. It means there exist an invertible matrix C such that

$B = C^{-1}AC$

$\Rightarrow \det B = \det(C^{-1}AC) \Rightarrow (\det C^{-1})(\det A)(\det C)$

$\Rightarrow \det B = (\det C^{-1})(\det C)(\det A) \Rightarrow \det B = (\det C^{-1}C)(\det A)$

$\Rightarrow \det B = (\det I)(\det A) \Rightarrow \det B = 1(\det A) \Rightarrow \det B = \det A$.

5.4.1 *SIMILARITY OF LINEAR TRANSFORMATION*

Definition: Let A and B be linear transformation on a vector space $V(F)$. Then B is said to be similar to A if there exist an invertible linear transformation C on V such that

$$B = CAC^{-1}$$

Theorem 4: The relation of similarity is an equivalence relation in the set of all linear transformations on a vector space $V(F)$.

Proof: If A and B are two linear transformation on the vector space $V(F)$, then B is said to be similar to A if there exists an invertible linear transformation C on V such that

$$B = CAC^{-1}$$

Reflexive: Let A be any linear transformation on V such that we rewrite,

$A = IAI^{-1}$, where I denote the identity transformation on V.

$\therefore A$ is similar to A because I is definitely invertible.

Symmetry: Let A is similar to B. Then there exist an invertible linear transformation P on V such that

$A = PBP^{-1}$

$\Rightarrow P^{-1}AP = P^{-1}(PBP^{-1})P$

$\Rightarrow P^{-1}AP = B \Rightarrow B = P^{-1}AP$

$\Rightarrow B = P^{-1}A(P^{-1})^{-1} \Rightarrow B$ is similar to A.

Transition: Let A be similar to B and also B is similar to C.

Then, $A = PBP^{-1}$

and $B = QCQ^{-1}$

where P and Q are invertible linear transformation on V.

We have $B = CAC^{-1} = P(QCQ^{-1})P^{-1}$

$= (PQ)C(Q^{-1}P^{-1}) = (PQ)C(PQ)^{-1}$

$\therefore A$ be similar to C.

Theorem 5: Let T be a linear transformation on an $n-$dimensional vector space $V(F)$ and let B and B' be two ordered basis for V. Then the matrix of T relative to B' is similar to the matrix of T relative to $B.$

Proof: Let $B = \{\alpha_1, \alpha_2, ..., \alpha_n\}$ and $B' = \{\beta_1, \beta_2, ..., \beta_n\}$.

Let $A = [a_{ij}]_{n\times n}$ be the matrix of T relative to B

and $C = [c_{ij}]_{n\times n}$ be the matrix of T relative to B'. Then

$$T(\alpha_j) = \sum_{i=1}^{n} a_{ij}\alpha_i,\ j = 1, 2, ..., n \qquad \dots (1)$$

and $$T(\beta_j) = \sum_{i=1}^{n} c_{ij}\beta_i,\ j = 1, 2, ..., n \qquad \dots (2)$$

Let S be the linear operator on V defined by

$$S(\alpha_j) = \beta_j,\ j = 1, 2, ..., n \qquad \dots (3)$$

Since S maps a basis B onto a basis B', therefore S is necessarily invertible. Let P be the matrix of S relative to B. Then P is also an invertible matrix.

If, $P = [p_{ij}]_{n\times n}$, then

$$S(\alpha_j) = \sum_{i=1}^{n} p_{ij}\alpha_i,\ j = 1, 2, ..., n \qquad \ldots (4)$$

We have,

$T(\beta_j) = T[S(\alpha_j)]$ [From (3)]

$= T\left(\sum_{k=1}^{n} p_{kj}\alpha_k\right)$ [From (4), on replacing i by k which is immaterial]

$= \sum_{k=1}^{n} p_{kj} T(\alpha_k)$ [$\because T$ is linear]

$= \sum_{k=1}^{n} p_{kj} \sum_{i=1}^{n} a_{ik}\alpha_i$ [From (1), on replacing j by k]

$= \sum_{i=1}^{n}\left(\sum_{k=1}^{n} a_{ik} p_{kj}\right)\alpha_i. \qquad \ldots (5)$

Also, $T(\beta_j) = \sum_{k=1}^{n} c_{kj}\beta_k$ [From (2), on replacing i by k]

$= \sum_{k=1}^{n} c_{kj} S(\alpha_k)$ [From (3)]

$= \sum_{k=1}^{n} c_{kj} \sum_{i=1}^{n} p_{ik}\alpha_i$ [From (4), on replacing j by k]

$= \sum_{i=1}^{n}\left(\sum_{k=1}^{n} p_{ik} c_{kj}\right)\alpha_i \qquad \ldots (6)$

From (5) and (6), we have

$$\sum_{i=1}^{n}\left(\sum_{k=1}^{n} a_{ik} p_{kj}\right)\alpha_i = \sum_{i=1}^{n}\left(\sum_{k=1}^{n} p_{ik} c_{kj}\right)\alpha_i$$

$\Rightarrow \sum_{k=1}^{n} a_{ik} p_{kj} = \sum_{k=1}^{n} p_{ik} c_{kj}$

$\Rightarrow [a_{ik}]_{n\times n}[p_{kj}]_{n\times n} = [p_{ik}]_{n\times n}[c_{kj}]_{n\times n}$ [By def. of matrix multiplication]

$\Rightarrow AP = PC$

$\Rightarrow P^{-1}AP = P^{-1}PC$ [$\because P^{-1}$ exists]

$\Rightarrow P^{-1}AP = IC = P^{-1}AP = C$

$\Rightarrow C$ is similar to A

Note: Suppose B and $B^{'}$ are two ordered basis for an $n-$dimensional vector space $V(F)$. Let T be a linear operator on V. Suppose A is the matrix of T relative to B and C is the matrix of T relatives to $B^{'}$. If P is the transition matrix from the basis B to the basis $B^{'}$, then $C = P^{-1}AP$. When we already know the matrix of T with respect to basis B, this solution will allow us to find the matrix of T with respect to basis $B^{'}$.

Theorem 6: Let V be an n-dimensional vector space over the field F and T_1, T_2 be two linear operator on V. If there exist two ordered B and $B^{'}$ for V such that $[T_1]_B = [T_1]_{B^{'}}$, then show that T_2 is similar to T_1.

Proof: Let $B = \{\alpha_1, \alpha_2, ..., \alpha_n\}$ and $B^{'} = \{\beta_1, \beta_2, ..., \beta_n\}$.

Let $[T_1]_B = [T_1]_{B^{'}} = A = [a_{ij}]_{n\times n}$. Then

$$T_1(\alpha_j) = \sum_{i=1}^{n} a_{ij}\alpha_i,\ j = 1, 2, ..., n \qquad \dots (1)$$

and $$T_1(\beta_j) = \sum_{i=1}^{n} a_{ij}\beta_i,\ j = 1, 2, ..., n \qquad \dots (2)$$

Let S be the linear operator on V defines by

$$S(\alpha_j) = \beta_j,\ j = 1, 2, ..., n \qquad \dots (3)$$

Since S maps a basis of V onto a basis of V, therefore S is invertible.

We have $T_2(\beta_j) = T_2[S(\alpha_j)]$ [From (3)]

$= (T_2S)(\alpha_j)$ … (4)

Also, $T_2(\beta_j) = \sum_{i=1}^{n} a_{ij}\beta_i$ [From (2)]

$= \sum_{i=1}^{n} a_{ij}S(\alpha_i)$ [From (3)]

$= S\left(\sum_{i=1}^{n} a_{ij}\alpha_i\right)$ [$\because S$ is linear]

$= S[T_1(\alpha_j)]$ [From (1)]

$= (ST_1)(\alpha_j)$ … (5)

From (4) and (5), we have

$(T_2S)(\alpha_j) = (ST_1)(\alpha_j).\ j = 1, 2, ..., n$

Since T_2S, ST_1 agree on a basis for V, therefore we have

$T_2S = ST_1$

$\Rightarrow T_2SS^{-1} = ST_1S^{-1} \Rightarrow T_2I = ST_1S^{-1}$

$\Rightarrow T_2 = ST_1S^{-1} \Rightarrow T_2$ is similar to T_1

5.5 *DETERMINANT OF LINEAR TRANSFORMATION IN FINITE DIMENSIONAL VECTOR SPACE*

Let we assume that T be a linear operator in $n-$dimensional vector space $V(F)$ and B, B' are two ordered basis for V, then $[T]_B, [T]_{B'}$ are two similar matrices. As we know that similar matrices have same determinant. This allows us to define in the manner that follows:

Definition: Let T be a linear operator on $n-$dimensional vector space $V(F)$. Then, with respect to any ordered basis for V, the determinant of T equals the determinant of the matrix of T. Following the discussion above, our definition of T's determinant is reasonable since it is a unique element of F.

Definition (Scalar transformation): In a given vector space $V(F)$, a linear transformation T on V is referred to be a scalar transformation T on V if, $T(\alpha) = c\alpha \; \forall \; \alpha \in V$, where c is fixed scalar in F.

If the linear transformation T is equal to the scalar c, then we rewrite $T = cI$, where I is the identity matrix.

5.6 *TRACE OF A MATRIX*

Definition: Let A be n order square matrix over a field F. The trace of A is the total sum of the elements of A that lie along the principal diagonal. Mathematically we define trace of the matrix by

$$tr(A) = \sum_{i=1}^{n} a_{ii} = a_{11} + a_{22} + \ldots + a_{nn}$$

Some fundamental properties/theorems of the trace of a matrix are as follows:

Theorem 7: Let A and B be two square matrices of order n over a field F and $\lambda \in F$. Then

(1) $tr(\lambda A) = \lambda tr(A)$

(2) $tr(A+B) = tr(A) + tr(B)$

(3) $tr(AB) = tr(BA)$

Proof: Let $A = [a_{ij}]_{n\times n}$ and $B = [b_{ij}]_{n\times n}$.

(1) We have $\lambda A = [\lambda a_{ij}]_{n\times n}$ by def. of multiplication of a matrix by a scalar.

$$\therefore tr(\lambda A) = \sum_{i=1}^{n} \lambda a_{ii} = \lambda \sum_{i=1}^{n} a_{ii} = \lambda tr(A)$$

(2) We have $A + B = [a_{ij} + b_{ij}]_{n\times n}$

$$\therefore tr(A+B) = \sum_{i=1}^{n} (a_{ii} + b_{ii}) = \sum_{i=1}^{n} a_{ii} + \sum_{i=1}^{n} b_{ii} = tr(A) + tr(B)$$

(3) We have $AB = [c_{ij}]_{n\times n}$ where $c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$.

Also $BA = [d_{ij}]_{n\times n}$, where $d_{ij} = \sum_{k=1}^{n} b_{ik}a_{kj}$

Now $tr(AB) = \sum_{i=1}^{n} c_{ii} = \sum_{i=1}^{n}\left(\sum_{k=1}^{n} a_{ik}b_{ki}\right)$

$$= \sum_{k=1}^{n}\sum_{i=1}^{n} a_{ik}b_{ki} \quad \text{[interchanging the order of summation in the last sum]}$$

$$= \sum_{k=1}^{n}\left(\sum_{i=1}^{n} b_{ki}a_{ik}\right) = \sum_{k=1}^{n} d_{kk}$$

$$= d_{11} + d_{22} + ... + d_{nn} = tr(BA)$$

Theorem 8: Trace of the similar matrices are same.

Proof: Let us consider that T be the linear operator in a $n-$dimensional vector space $V(F)$. If B and B' are two ordered basis for V then, $[T]_B$, $[T]_{B'}$ are the similar matrices. Also similar matrices have the same trace. Also we know that similar matrices have the same trace. This allows us to define in the manner that follows:

Definition (Trace of linear transformation): Let T be a linear operator in vector space $V(F)$ of dimension n. In that case, the trace of T is the matrix of T with respect to any ordered basis for V.

Based on the previous explanation, our definition of the trace of T is reasonable since it is a distinct element of F.

Solved Examples

Example 1: Find the matrix of the linear transformation T on $V_3(R)$ defined as

$$T(a,b,c) = (2b+c, a-4b, 3a)$$

With respect to the ordered basis B and also with respect to the basis B' where,

(i) $B = \{(1,0,0),(0,1,0),(0,0,1)\}$

(ii) $B' = \{(1,1,1),(1,1,0),(1,0,0)\}$

Answer (*i*): We have,

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

$T(0,1,0) = (2,-4,0) = 2(1,0,0) - 4(0,1,0) + 0(0,0,1)$

and $T(0,0,1) = (1,0,0) = 1(1,0,0) + 0(0,1,0) + 0(0,0,1)$.

Thus, by the definition of T with respect to B, we have

$$[T]_B = \begin{bmatrix} 0 & 2 & 1 \\ 1 & -4 & 0 \\ 3 & 0 & 0 \end{bmatrix}$$

Note: In order to find the matrix of T relative to the standard basis B, it is sufficient to compute $T(1,0,0)$, $T(0,1,0)$ and $T(0,0,1)$. There is no need of further expressing these vectors as linear combinations of $(1,0,0)$, $(0,1,0)$ and $(0,0,1)$. Obviously the coordinates of the vectors $T(1,0,0)$, $T(0,1,0)$ and $T(0,0,1)$ respectively constitutes the first, second and third columns of the matrix $[T]_B$.

(ii) We have $T(1,1,1) = (3,-3,3)$

Now our aim to express $(3,-3,3)$ as a linear combination of vectors in B'. Let

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

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

Then, $x+y+z = a, x+y = b, x = c$

i.e., $x = c, y = b-c, z = a-b$ … (1)

putting $a = 3, b = -3$ and $c = 3$ in (1), we get

$x = 3, y = -6$ and $z = 6$.

$\therefore T(1,1,1) = (3,-3,3) = 3(1,1,1) - 6(1,1,0) + 6(1,0,0)$

Also, $T(1,1,0) = (2,-3,3)$.

Putting $a = 2, b = -3$ and $c = 3$ in (1), we get

$T(1,1,0) = (2,-3,3) = 3(1,1,1) - 6(1,1,0) + 6(1,0,0)$

Finally, $T(1,0,0) = (0,1,3)$.

Putting $a = 0, b = 1$ and $c = 3$ in (1), we get

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

$$\therefore [T]_{B'} = \begin{bmatrix} 3 & 3 & 3 \\ -6 & -6 & -2 \\ 6 & 5 & -1 \end{bmatrix}.$$

Example 2: Let T be the linear operator on R^3 defined by

$T(x_1, x_2, x_3) = (3x_1 + x_3, -2x_1 + x_2, -x_1 + 2x_2 + 4x_3)$. What is the matrix of T in the ordered basis $\{\alpha_1, \alpha_2, \alpha_3\}$ where $\alpha_1 = (1,0,1), \alpha_2 = (-1,2,1)$ and $\alpha_3 = (2,1,1)$?

Solution: By definition of T , we have

$T(\alpha_1) = T(1,0,1) = (4,-2,3)$.

Now our aim is to express $(4,-2,3)$ as a linear combination of the vectors in the basis $B = \{\alpha_1, \alpha_2, \alpha_3\}$. Let

$$(a,b,c) = x\alpha_1 + y\alpha_2 + z\alpha_3$$

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

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

Then, $x - y + 2z = a,\ 2y + z = b, x + y + z = c$

Solving these equations, we get

$$x = \frac{-a - 3b + 5c}{4}, y = \frac{b + c - a}{4}, z = \frac{b - c + a}{2} \qquad \ldots (1)$$

Putting $a = 4, b = -2, c = 3$ in (1), we get

$$x = \frac{17}{4}, y = -\frac{3}{4},\ z = -\frac{1}{2}.$$

$$\therefore T(\alpha_1) = \frac{17}{4}\alpha_1 - \frac{3}{4}\alpha_2 - \frac{1}{4}\alpha_3$$

Also $T(\alpha_2) = T(-1,2,1) = (-2,4,9)$. Putting

$a=-2, b=4, c=9$ in (1), we get $x=\frac{35}{4}, y=\frac{15}{4}, z=-\frac{7}{2}$

$$T(\alpha_2)=\frac{35}{4}\alpha_1+\frac{15}{4}\alpha_2-\frac{7}{2}\alpha_3$$

Finally $T(\alpha_3)=T(2,1,1)=(7,-3,4)$. Putting,

$a=7, b=-3, c=4$ in (1), we get $x=\frac{11}{2}, y=-\frac{3}{2}, z=0$

$$\therefore T(\alpha_3)=\frac{11}{2}\alpha_1-\frac{3}{2}\alpha_2+0\alpha_3$$

$$\therefore [T]_B=\begin{bmatrix} \frac{17}{4} & \frac{35}{4} & \frac{11}{2} \\ -\frac{3}{4} & \frac{15}{4} & -\frac{3}{2} \\ \frac{1}{2} & -\frac{7}{2} & 0 \end{bmatrix}$$

Example 3: Let T be a linear operator on R^3 defined by $T(x_1,x_2,x_3)=(3x_1+x_3,-2x_1+x_2,-x_1+2x_2+4x_3)$. Prove that T is invertible and find the a formula for T^{-1} .

Solution: Suppose B is the standard ordered basis for R^3 . Then $B=\{(1,0,0),(0,1,0),(0,0,1)\}$. Let $A=[T]_B$ i.e. let A be the matrix of T with respect to B. First we shall compute A.

We have

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

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

$$T(0,0,1)=(1,0,4)$$

And

$$\therefore A = [T]_B = \begin{bmatrix} 3 & 0 & 1 \\ -2 & 1 & 0 \\ -1 & 2 & 4 \end{bmatrix}.$$

Now T will be invertible If the matrix $[T]_B$ is invertible.

We have $\det A = |A| = \begin{vmatrix} 3 & 0 & 1 \\ -2 & 1 & 0 \\ -1 & 2 & 4 \end{vmatrix} = 3(4-0)+(-4+1) = 9$

Since $\det A \neq 0,$ therefore the matrix A is invertible and consequently T is invertible.

Now we shall compute the matrix A^{-1}. For this let us first find $adj\ A$.

The cofactors of the elements of the first row of A are

$$\begin{vmatrix} 1 & 0 \\ 2 & 4 \end{vmatrix}, \begin{vmatrix} -2 & 0 \\ -1 & 4 \end{vmatrix}, \begin{vmatrix} -2 & 1 \\ -1 & 2 \end{vmatrix} \quad \text{i.e, } 4,\ 8, -3$$

The cofactors of the elements of the first row of A are

$$-\begin{vmatrix} 0 & 1 \\ 2 & 4 \end{vmatrix}, \begin{vmatrix} 3 & 1 \\ -1 & 4 \end{vmatrix}, -\begin{vmatrix} 3 & 0 \\ -1 & 2 \end{vmatrix} \quad \text{i.e, } 2,\ 13, -6$$

The cofactors of the elements of the first row of A are

$$\begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix}, -\begin{vmatrix} 3 & 1 \\ -2 & 0 \end{vmatrix}, \begin{vmatrix} 3 & 0 \\ -2 & 1 \end{vmatrix} \quad \text{i.e, } -1,\ -2, 3$$

$$\therefore Adj\ A = \text{transpose of the matrix } \begin{bmatrix} 4 & 8 & -3 \\ 2 & 13 & -6 \\ -1 & -2 & 3 \end{bmatrix}$$

$$= \begin{bmatrix} 4 & 2 & -1 \\ 8 & 13 & -2 \\ -3 & -6 & 3 \end{bmatrix}$$

$$\therefore A^{-1} = \frac{1}{\det A} Adj\, A = \frac{1}{9}\begin{bmatrix} 4 & 2 & -1 \\ 8 & 13 & -2 \\ -3 & -6 & 3 \end{bmatrix}$$

Now $[T^{-1}]_B = ([T]_B)^{-1} = A^{-1}$.

We shall now find a formula for T^{-1}. Let $\alpha = (a,b,c)$ be any vector belonging to R^3. Then

$$[T^{-1}(\alpha)]_B = [T^{-1}]_B[\alpha]_B$$

$$= \frac{1}{9}\begin{bmatrix} 4 & 2 & -1 \\ 8 & 13 & -2 \\ -3 & -6 & 3 \end{bmatrix}\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \frac{1}{9}\begin{bmatrix} 4a+2b-c \\ 8a+13b-2c \\ -3a-6b+3c \end{bmatrix}$$

Since B is the standard ordered basis for R^3,

$$\therefore T^{-1}(\alpha) = T^{-1}(a,b,c) = \frac{1}{9}(4a+2b-c, 8a+13b-2c, -3a-6b+3c).$$

Example 4: Let T be the linear operator on R^3 defined by

$$T(x_1,x_2,x_3) = (3x_1 + x_3, -2x_1 + x_2, -x_1 + 2x_2 + 4x_3)$$

- **(i)** Find the matrix of T in the standard ordered basis B for R^3.
- **(ii)** Find the transition matrix P from the ordered basis B to the ordered basis $B' = \{\alpha_1, \alpha_2, \alpha_3\}$ where $\alpha_1 = (1,0,1), \alpha_2 = (-1,2,1)$ and $\alpha_3 = (2,1,1)$. Hence find the matrix of T relative to the ordered basis B'.

Solution (i): Let $A = [T]_B$. Then

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

(ii) Since B is the standard ordered basis, so, the transition matrix P from B to B' can written as

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

Now $[T]_{B'} = P^{-1}[T]_B P$.

Now we compute matrix P^{-1}, then we find that $\det P = -A$.

Therefore $P^{-1} = \dfrac{1}{\det P} Adj\, P = -\dfrac{1}{4}\begin{bmatrix} 1 & 3 & -5 \\ 1 & -1 & -1 \\ -2 & -2 & 2 \end{bmatrix}$

$$\therefore [T]_{B'} = -\frac{1}{4}\begin{bmatrix} 1 & 3 & -5 \\ 1 & -1 & -1 \\ -2 & -2 & 2 \end{bmatrix}\begin{bmatrix} 3 & 0 & 1 \\ -2 & 1 & 0 \\ -1 & 2 & 4 \end{bmatrix}\begin{bmatrix} 1 & -1 & 2 \\ 0 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

$$= -\frac{1}{4}\begin{bmatrix} 2 & -7 & -19 \\ 6 & -3 & -3 \\ -4 & 2 & 6 \end{bmatrix}\begin{bmatrix} 1 & -1 & 2 \\ 0 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

$$= -\frac{1}{4}\begin{bmatrix} -17 & -35 & -22 \\ 3 & -15 & 6 \\ 2 & 14 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} \frac{17}{4} & \frac{35}{4} & \frac{11}{2} \\ -\frac{3}{4} & \frac{15}{4} & -\frac{3}{2} \\ -\frac{1}{2} & -\frac{7}{2} & 0 \end{bmatrix}$$

Example 5: Let T be the linear operator on R^2 defined by

$T(x, y) = (4x - 2y, 2x + y)$

Compute the matrix of T relative to the basis $\{\alpha_1, \alpha_2\}$ where $\alpha_1 = (1,1), \alpha_2 = (-1, 0)$.

Solution: By def. of T, we have

$T(\alpha_1) = T(1,1) = (2,3)$

Now, we express the vector $(2,3)$ as the linear combination to the basis $\{\alpha_1, \alpha_2\}$.

Let $(a,b) = x\alpha_1 + y\alpha_2 = x(1,1) + y(-1,0) = (x-y, x)$.

Then $x - y = a, x = b$

Solving these equations, we get

$x = b, y = b - a$ … (1)

Putting $a = 2, b = 3$ in (1), we get $x = 3, y = 1$

$\therefore T(\alpha_1) = 3\alpha_1 + 1\alpha_2$ … (2)

Again $T(\alpha_2) = T(-1,0) = (-4,-2)$. Putting $a = -4, b = -2$ in (1), we get $x = -2, y = 2$.

$\therefore T(\alpha_2) = -2\alpha_1 + 2\alpha_2$ … (3)

From the relation (2) and (3), we see that the matrix of T relative to the basis $\{\alpha_1, \alpha_2\}$ is $= \begin{bmatrix} 3 & -2 \\ 1 & 2 \end{bmatrix}$.

Example 6: Let T be a linear operator on R^2 defined by:

$T(x,y) = (2y, 3x - y)$

Find the matrix representation of T relative to the basis $\{(1,3),(2,5)\}$.

Solution: Let $\alpha_1 = (1,3)$ and $\alpha_2 = (2,5)$. By def. of T, we have

$T(\alpha_1) = T(1,3) = (2.3, 3.1 - 3) = (6,0)$

And $T(\alpha_2) = T(2,5) = (2.5, 3.2 - 5) = (10,1)$.

Now our aim is to express the vectors $T(\alpha_1)$ and $T(\alpha_2)$ as linear combinations of the vectors in the basis $\{\alpha_1, \alpha_2\}$.

Let $(a,b) = p\alpha_1 + q\alpha_2 = p(1,3) + q(2,5) = (p+2q, 3p+5q)$.

Then $p+2q = a, 3p+5q = b$

Solving these equations, we get

$p = -5a + 2b, q = 3a - b$ …(1)

Putting $a = 6, b = 0$ in (1), we get $p = -30, q = 18$.

$\therefore T(\alpha_1) = (6,0) = -30\alpha_1 + 18\alpha_2$ … (2)

Again putting $a = 10, b = 1$ in (1), we get

$p = -48, q = 29$

$\therefore T(\alpha_2) = (10,1) = -48\alpha_1 + 29\alpha_2$ … (3)

From the relations (2) and (3), we see that the matrix of T relative to the basis $\{\alpha_1, \alpha_2\}$ is

$$\begin{bmatrix} -30 & -48 \\ 18 & 29 \end{bmatrix}$$

Example 7: Show that the vectors $\alpha_1 = (1,0,-1), \alpha_2 = (1,2,1), \alpha_3 = (0,-3,2)$ form a basis for R^3. Express the each standard basis vector in the linear combination of the vectors of $\alpha_1, \alpha_2, \alpha_3$.

Solution: Let a,b,c be scalars such that,

$a\alpha_1 + b\alpha_2 + c\alpha_3 = 0$

i.e., $a(1,0,-1) + b(1,2,1) + c(0,-3,2) = (0,0,0)$

i.e., $(a+b+0c, 0a+2b-3c, -a+b+2c) = (0,0,0)$

i.e., $a+b+0c = 0$

$0a + 2b - 3c = 0$

$-a + b + 2c = 0$

The coefficient matrix A of these equation is,

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & -3 \\ -1 & 1 & 2 \end{bmatrix}.$$

We have $\det A = |A| = \begin{vmatrix} 1 & 1 & 0 \\ 0 & 2 & -3 \\ -1 & 1 & 2 \end{vmatrix} = 1(4+3) - 1(0-3) = 7 + 3 = 10$

Since $\det A \neq 0$, therefore the matrix A is non-singular and rank $A = 3$ i.e., equal to the number of unknowns $a, b, c.$ Hence $a = 0, b = 0, c = 0$ is the only solution of the equation (1). Therefore, the vectors $\alpha_1, \alpha_2, \alpha_3$ are linearly independent over R. Since $\dim R^3 = 3$, therefore the set $\{\alpha_1, \alpha_2, \alpha_3\}$ containing three linearly independent vectors form a basis for R^3.

Now let $B = \{e_1, e_2, e_3\}$ be the ordered standard basis for R^3. Then $e_1 = (1, 0, 0), e_2 = (0, 1, 0), e_3 = (0, 0, 1)$. Let $B' = \{\alpha_1, \alpha_2, \alpha_3\}$. We have

$\alpha_1 = (1, 0, -1) = 1e_1 + 0e_2 - 1e_3$

$\alpha_2 = (1, 2, 1) = 1e_1 + 2e_2 + 1e_3$

$\alpha_3 = (0, -3, 2) = 0e_1 - 3e_2 + 2e_3$

If P is the transition matrix from the basis B to the basis B', then

$$P = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & -3 \\ -1 & 1 & 2 \end{bmatrix}$$

Let us find the matrix P^{-1}. For this let us first find $Adj\, P$. The cofactors of the elements of the first row of P are

$$\begin{vmatrix} 2 & -3 \\ 1 & 2 \end{vmatrix}, -\begin{vmatrix} 0 & -3 \\ -1 & 2 \end{vmatrix}, \begin{vmatrix} 0 & 2 \\ -1 & 1 \end{vmatrix} \; i.e., 7,3,2.$$

The cofactor of the elements of the second row of P are

$$-\begin{vmatrix} 1 & 0 \\ 1 & 2 \end{vmatrix}, \begin{vmatrix} 1 & 0 \\ -1 & 2 \end{vmatrix}, -\begin{vmatrix} 1 & 1 \\ -1 & 1 \end{vmatrix} \; i.e., -2, 2, -2$$

The cofactor of the elements of the third row of P are

$$\begin{vmatrix} 1 & 0 \\ 2 & -3 \end{vmatrix}, -\begin{vmatrix} 1 & 0 \\ 0 & -3 \end{vmatrix}, \begin{vmatrix} 1 & 1 \\ 0 & 2 \end{vmatrix} \; i.e., -3, 3, 2.$$

$$\therefore Adj\, P = \text{transpose of the matrix} \begin{bmatrix} 7 & 3 & 2 \\ -2 & 2 & -2 \\ -3 & 3 & 2 \end{bmatrix} = \begin{bmatrix} 7 & -2 & -3 \\ 3 & 2 & 3 \\ 2 & -2 & 2 \end{bmatrix}$$

$$\therefore P^{-1} = \frac{1}{\det P} AdjP = \frac{1}{10} \begin{bmatrix} 7 & -2 & -3 \\ 3 & 2 & 3 \\ 2 & -2 & 2 \end{bmatrix}$$

Now, $e_1 = 1e_1 + 0e_2 + 0e_3$.

$\therefore$ Coordinates matrix of e_1 relative to the basis $B = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$.

$\therefore$ Coordinates matrix of e_1 relative to the basis $B' = [e_1]_{B'} = P^{-1} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$

$$= \frac{1}{10} \begin{bmatrix} 7 & -2 & -3 \\ 3 & 2 & 3 \\ 2 & -2 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

$$= \frac{1}{10} \begin{bmatrix} 7 \\ 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 7/10 \\ 3/10 \\ 2/10 \end{bmatrix}$$

$$\therefore e_1 = \frac{7}{10}\alpha_1 + \frac{3}{10}\alpha_2 + \frac{2}{10}\alpha_3.$$

Also $[e_2]_B = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ and $[e_3]_B = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$.

$\therefore [e_2]_{B'} = P^{-1}\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ and $[e_3]_{B'} = P^{-1}\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$

Thus $[e_2]_{B'} = \frac{1}{20}\begin{bmatrix} -2 \\ 2 \\ -2 \end{bmatrix}, [e_3]_{B'} = \frac{1}{10}\begin{bmatrix} -3 \\ 3 \\ 2 \end{bmatrix}$.

$\therefore e_2 = -\frac{2}{10}\alpha_1 + \frac{2}{10}\alpha_2 - \frac{2}{10}\alpha_3.$

and $e_3 = -\frac{3}{10}\alpha_1 + \frac{3}{10}\alpha_2 + \frac{2}{10}\alpha_3$.

Check your progress

Problem 1: If T be the linear operator on R^2 defined by $T(a,b) = (a,0)$ then write the matrix of T in the standard ordered basis $B = \{(1,0),(0,1)\}$.

Also if $B' = \{(1,1),(2,1)\}$ is another ordered basis for R^3, find the transition matrix P from the basis B'. Hence find the matrix of T relative to the basis B'.

Problem 2: If the matrix of a linear transformation T on $V_2(C)$, with respect to the ordered basis $B = \{(1,0),(0,1)\}$ is $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$, what is the matrix of T with respect to the ordered basis

$B' = \{(1,1),(1,-1)\}$?

Problem 3: Is it true that only matrix similar to the identity matrix I is itself.

5.7 *SUMMARY*

In this unit, we have learned about the important concept of change of basis, similarity of matrices, determinant and trace of matrices. The overall summarization of this units are as follows:

- The coordinates of each vector in V relative to the basis B can be expressed to the coordinates relative to the basis B'.
- The relation of similarity is an equivalence relation in the set of all $n \times n$ matrices over the field F
- Similar matrices have the same determinant
- The relation of similarity is an equivalence relation in the set of all linear transformations on a vector space $V(F)$.
- Trace of the similar matrices are same.

5.8 *GLOSSARY*

- Change of basis
- Similarity of matrices
- Determinant of linear transformation
- Trace of matrices

5.9 *REFERENCES*

- Gel′fand I.M. Lectures on linear algebra (1989), Courier Corporation.
- Kenneth Hoffman & Ray Kunze, Linear Algebra (2nd edition)(2015), Prentice-Hall.
- David C. Lay, Linear Algebra and its Application (3rd Edition) (2007) Pearson Education Asia, India Reprint.
- Seymour Lipshutz and Marc Lipson, Schaum's outlines "Linear Algebra" (3rd Edition)(2012), Mc Graw Hill Education.
- J. N. Sharma and A. R. Vasistha, Linear Algebra (29th Edition) (1999), Krishna Prakashan.

5.10 *SUGGESTED READING*

- Minking Eie & Shou-Te Chang (2020), A First Course In Linear Algebra, World Scientific.
- Axler, Sheldon (2015), Linear algebra done right. Springer.
- https://nptel.ac.in/courses/111106051

- https://archive.nptel.ac.in/courses/111/104/111104137

5.11 *TERMINAL QUESTION*

Long Answer Type Question:

1. Find the matrix relative to the basis $\alpha_1 = \left(\frac{2}{3}, \frac{2}{3}, -\frac{1}{3}\right)$, $\alpha_2 = \left(\frac{1}{3}, -\frac{2}{3}, -\frac{2}{3}\right)$, $\alpha_3 = \left(\frac{2}{3}, -\frac{1}{3}, -\frac{2}{3}\right)$ of R^3, of the linear transformation $T : R^3 \to R^3$ whose matrix relative to the standard basis is $\begin{bmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 3 \end{bmatrix}$

2. Find the co-ordinates of the vector $(2,1,3,4)$ relative to the basis vectors $\alpha_1 = (1,1,0,0)$, $\alpha_2 = (1,0,1,1)$, $\alpha_3 = (2,0,0,2)$, $\alpha_4 = (0,0,2,2)$

3. If F be a field and V, the set of all polynomials in x over F of degree ≤ 5. If $D : V \to V'$ is defined by $D[f(x)] = f''(x)$, where $f'(x)$ is the derivative of $f(x)$, show that D is a linear transformation on V. Find the matrix of D in the basis $\{1, x, x^2, x^3, x^4\}$.

4. Let V be the vector space of those polynomial functions from the reals into itself which have ≤ 3. Let $B = \{f_1, f_2, f_3, f_4\}$ where $f_i(x) = x^{i-1}(1 \leq i \leq 4)$. Then show that B forms a basis for V. For any real number t let $g_i(x) = (x+t)^{i-1}$. Show that $B' = \{g_1, g_2, g_3, g_4\}$ is also a basis for V. If D is the differentiation operator on V, write the matrices of D in the ordered bases B and B'.

5. If A and B are $n \times n$ complex matrices, then show that $AB - BA = I$ is impossible.

6. Let T be a linear operator on R^3 defined by $T(x, y, z) = (3x + z, -2x + y, -x + 2y + 4z)$. Prove that T is invertible and find a formula for T^{-1}.

7. Let T be a linear transformation on an $n-$dimensional vector space $V(F)$ and let B and B' be two ordered basis for V. Then show that the matrix of T relative to B' is similar to the matrix of T relative to B.

Short answer type question:

1. If T and S are similar linear transformation on a finite dimensional vector space $V(F)$, $\det T = \det S$.

2. If A and B are linear transformation on the same vector space and if at least one of them is invertible, then AB and BA are similar.

3. If two linear transformations A and B on $V(F)$ are similar, then show that A^2 and B^2 are also similar and if A, B are invertible, then A^{-1}, B^{-1} are also similar.

4. Show that identity matrix (I) is the only matrix similar to itself.

5. Consider the vector space $V(R)$ of all 2×2 matrices over the field R of real numbers. Let T be the linear transformation on V that sends each matrix X onto AX, where $A=\begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix}$. Find the matrix of T with respect to the ordered basis $B=\{\alpha_1,\alpha_2,\alpha_3,\alpha_4\}$ for V where $\alpha_1=\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}, \alpha_2=\begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}, \alpha_3=\begin{bmatrix}0 & 0\\ 1 & 0\end{bmatrix}, \alpha_4=\begin{bmatrix}0 & 0\\ 0 & 1\end{bmatrix}$

6. Prove that similar matrices have the same determinant.

Fill in the blanks:

1. The relation of similarity is an ……………… relation in the set of all $n\times n$ matrices over the field F

2. Similar matrices have same ……………

3. The sum of the diagonal element of any square matrix ……………

4. The trace of similar matrices are …………………

5. $tr(AB)=$…………………..

6. $\det(A+B)=$……………………..

5.12 *ANSWERS*

Answers of check your progress:

1: $[T]_B=\begin{bmatrix}1 & 0\\ 1 & 0\end{bmatrix}; P=\begin{bmatrix}1 & 2\\ 1 & 1\end{bmatrix}; [T]_{B'}=\begin{bmatrix}-1 & -2\\ 1 & 2\end{bmatrix}$

2: $[T]_{B'}=\begin{bmatrix}2 & 0\\ 0 & 0\end{bmatrix}$

3: Matrix similar to I is I itself

Answers of long answer type question:

1: $$\begin{bmatrix} 3 & -\frac{2}{3} & -\frac{2}{3} \\ -\frac{2}{3} & \frac{10}{3} & 0 \\ -\frac{2}{3} & 0 & \frac{8}{3} \end{bmatrix}$$

2: $(2,1,3,4) = \alpha_1 + \frac{1}{2}\alpha_3 + \frac{3}{2}\alpha_4$

Answer of fill in the blanks questions:

1: Equivalence

2: Determinant

3: Trace of the matrix

4: Same

5: $tr(BA)$

6: $\det(A) + \det(B)$

BLOCK- II

QUOTIENT SPACE AND LINEAR FUNCTION

UNIT-6: QUOTIENT SPACE

CONTENTS

6.1 *INTRODUCTION*

Let $X = R^2$ be the standard Cartesian plane, and let Y be a line through the origin in X. The space occupied by all X lines that are parallel to Y is known as the quotient space X/Y. In other words, lines in X that are parallel to Y make up the elements of the set X/Y. Because their difference vectors belong to Y, the points along any given such line will satisfy the equivalence relation. This provides a geometric method of visualizing quotient spaces. The quotient space can be more often described as the space of all points along a line through the origin that is not parallel to Y by re-parameterizing these lines. The set of all co-parallel lines or, alternatively, the vector space made up of a plane that only crosses the line at the origin can be used to represent the quotient space for R^3 by a line through the origin.

6.2 *OBJECTIVES*

After reading this unit learners will be able to

- Understand the concept of quotient space.
- Implement the application of dimension of quotient space.
- Understand the concept of direct sum of subspaces.
- Visualized and understand the concept of disjoint subspaces, complementary subspaces, direct sum of several subspaces and co-ordinates in a vector space.

6.3 *QUOTIENT SPACE*

In this section we will discuss about the quotient space.

Definition: Let W be any quotient subspace of a vector space $V(F)$. Also let α be any element of V . Then the set

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

is called the a right coset of W in V generated by α . Similarly the set,

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

is called the left coset of W in V generated by α .

Here, it is obvious that $W+\alpha$ and $\alpha+W$ are both subsets of V . Since in addition V is commutative, therefore we have $W+\alpha=\alpha+W$. Thus we shall call $W+\alpha$ as simply a coset of W in V generated by α.

The following results about the cossets to be remembered.

(i) We have $0\in V$ and $W+0=W$. Therefore W itself is a coset of W in V .

(ii) $\alpha \in W \Rightarrow W + \alpha = W$

Proof : First we shall prove that $W + \alpha \subseteq W$. Let $\gamma + \alpha$ be any arbitrary element of $W + \alpha$. Then $\gamma \in W$. Now W is a subspace of V. Therefore,

$\gamma \in W, \alpha \in W \Rightarrow \gamma + \alpha \in W$

So, each element of $W + \alpha$ is also element of W. Hence $W + \alpha \subseteq W$.

Now we have only to prove that $W \subseteq W + \alpha$.

Let $\beta \in W$. Since W is a subspace, therefore

$\alpha \in W \Rightarrow -\alpha \in W$

Thus, $\beta \in W$, $-\alpha \in W \Rightarrow \beta - \alpha \in W$. Now we can write,

$\beta = (\beta - \alpha) + \alpha \in W + \alpha$ since $\beta - \alpha \in W$.

Thus $\beta \in W \Rightarrow \beta \in W + \alpha$. Therefore $W \subseteq W + \alpha$

Hence $W = W + \alpha$ in $W + \alpha$

(iii) If $W + \alpha$ and $W + \beta$ are two cosets of W in V, then

$W + \alpha = W + \beta \Leftrightarrow \alpha - \beta \in W$

Proof: Since $0 \in W$, therefore $0 + \alpha \in W + \alpha$. Thus

$\alpha \in W + \alpha$.

Now, $W + \alpha = W + \beta \Rightarrow \alpha \in W + \beta$

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

$\Rightarrow \alpha - \beta \in W + 0 \Rightarrow \alpha - \beta \in W$

Conversely, $\alpha - \beta \in W \Rightarrow W + (\alpha - \beta) = W$

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

$$\Rightarrow W+\alpha = W+\beta$$

Let V/W denotes the set of all cosets of W in V i.e., let

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

We have just seen that if $\alpha - \beta \in W$, then $W+\alpha = W+\beta$. Thus a coset of W in V can have more than one representation.

Now if $V(F)$ is a vector space, then we shall give a vector space structure to the set V/W over the same field F. For this we shall have to define addition in V/W i.e., addition of coset of W in V and multiplication of a coset by an elelement of F i.e., scalar multiplication.

Theorem 1: If W is any subset of a vector space $V(F)$, then the set V/W of all cosets $W+\alpha$ where α is any arbitrary element of V, is a vector space over F for the addition and scalar multiplication compositions defined as follows:

$(W+\alpha)+(W+\beta) = W+(\alpha+\beta)\ \forall\ \alpha, \beta \in V$

and $a(W+\alpha) = W + a\alpha; a \in F, \alpha \in V$.

Proof: We have $\alpha, \beta \in V \Rightarrow \alpha + \beta \in V$

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

Therefore $W+(\alpha+\beta) \in V/W$ and also $W + a\alpha \in V/W$. Therefore, V/W is closed with regard to the aforementioned definitions of scalar multiplication and coset addition. Initially, we will demonstrate that these two compositions are well-defined, meaning they are not dependent on the specific representative selected to signify a coset.

Let $W+\alpha = W+\alpha', \alpha, \alpha' \in V$

and $W+\beta = W+\beta', \beta, \beta' \in V$

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

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

Now W is a subspace, therefore

$\alpha-\alpha^{'}\in W,\beta-\beta^{'}\in W\Rightarrow(\alpha-\alpha^{'})+(\beta-\beta^{'})\in W$

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

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

$$\Rightarrow W+(\alpha+\beta)=W+(\alpha^{'}+\beta^{'})$$

$$\Rightarrow(W+\alpha)+(W+\beta)=(W+\alpha^{'})+(W++\beta^{'})$$

Therefore addition in V/W is well defined.

Again, $a^{'}\in F,\alpha-\alpha^{'}\in W\Rightarrow a(\alpha-\alpha^{'})\in W$

$$\Rightarrow a\alpha-a\alpha^{'}\in W$$

$$\Rightarrow W+a\alpha=W+a\alpha^{'}$$

$\therefore$ scalar multiplication in V/W is also defined.

Commutativity of addition: Let $W+\alpha,W+\beta$ be any two elements of V/W . Then

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

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

Associativity of addition: Let $W+\alpha,W+\beta,W+\gamma$ be any three elements of V/W .Then

$(W+\alpha)+[(W+\beta)+(W+\gamma)]=(W+\alpha)+[W+(\beta+\gamma)]$

$=W+[\alpha+(\beta+\gamma)]$

$=W+[(\alpha+\beta)+\gamma)]$

$=[W+(\alpha+\beta)]+(W+\gamma)$

$=[(W+\alpha)+(W+\beta)]+(W+\gamma)$

Existence of additive identity: If 0 is the zero vector of V , then $W+0=W\in V/W$. If $W+\alpha$ is any element of V/W , If $W+\alpha$ is any element of V/W , then

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

$\therefore W+0=W$ is the additive identity.

Existence of additive inverse: If $W+\alpha$ is any element of V/W, then
$W+(-\alpha)=W-\alpha\in V/W$

then $W+(-\alpha)=W-\alpha\in V/W$. Also we have,

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

$\therefore W-\alpha$ is the additive inverse of $W+\alpha$.

Thus V/W is an abelian group with respect to addition composition. Further we observed that if

$a,b\in F$ and $W+\alpha, W+\beta\in V/W$, then

1. $a[(W+\alpha)+(W+\beta)]=a[W+(\alpha+\beta)]$

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

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

2. $(a+b)(W+\alpha)=W+(a+b)\alpha$

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

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

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

3. $(ab)(W+\alpha)=W+(ab)\alpha=W+a(b\alpha)$

$$=a(W+b\alpha)=a[b(W+\alpha)]$$

4. $1(W+\alpha)=W+1\alpha=W+\alpha$

$\therefore$ Across the field F in these two compositions, V/W is a vector space. The quotient space of V with respect to W is known as the vector space V/W. The zero vector in this vector space is denoted by the coset W.

6.4 *DIMENSION OF QUOTIENT SPACE*

In this section we will discuss about the dimension of the quotient space.

Theorem 2: If W be a subspace of a finite dimensional vector space $V(F)$, then

$\dim(V/W) = \dim V - \dim W$

Proof: Let m be the dimension of the subspace W of the vector space $V(F)$. Let

$S = \{\alpha_1, \alpha_2, ..., \alpha_m\}$

be a basis of W. Since S is a linearly independent subset of V. Let

$S' = \{\alpha_1, \alpha_2, ..., \alpha_m, \beta_1, \beta_2, ..., \beta_l\}$ be a basis of V. Then $\dim V = m + l$

$\therefore \dim V - \dim W = (m+l) - m = l$

So, we have to prove that $\dim V/W = l$

For it, let we claim that the set l cosets

$S_1 = \{W + \beta_1, W + \beta_2, ..., W + \beta_l\}$

Is a basis of V/W.

First, we will show that the set S_1 is linearly independent. Also the zero vector of V/W is W.

Let we consider, $a_1(W+\beta_1) + a_2(W+\beta_2) + ... + a_l(W+\beta_l) = W$

$\Rightarrow (W + a_1\beta_1) + (W + a_2\beta_2) + ... + (W + a_l\beta_l) = W + 0$

$\Rightarrow W + (a_1\beta_1 + a_2\beta_2 + ... + a_l\beta_l) = W + 0$

$\Rightarrow a_1\beta_1 + a_2\beta_2 + ... + a_l\beta_l \in W$

$\Rightarrow a_1\beta_1 + a_2\beta_2 + ... + a_l\beta_l = b_1\alpha_1 + b_2\alpha_2 + ... + b_m\alpha_m$ [Since any vector can be written as a linear combination of its basis vector]

$\Rightarrow a_1\beta_1 + a_2\beta_2 + ... + a_l\beta_l - b_1\alpha_1 - b_2\alpha_2 - ... - b_m\alpha_m = 0$

$\Rightarrow a_1 = 0, a_2 = 0, ..., a_l = 0$ because the vectors $\beta_1, \beta_2, ..., \beta_l, \alpha_1, \alpha_2, ..., \alpha_m$ are linearly independent.

$\therefore$ The set S_1 is linearly independent.

Now, we have only to prove that $L(S_1) = V/W$. Let $W + \alpha$ be any element of V/W. The vector $\alpha \in V$ can be expressed as

$\alpha = c_1\alpha_1 + c_2\alpha_2 + ... + c_m\alpha_m + d_1\beta_1 + d_2\beta_2 + ... + d_l\beta_l$

$= \gamma + d_1\beta_1 + d_2\beta_2 + ... + d_l\beta_1$ where $\gamma = c_1\alpha_1 + c_2\alpha_2 + ... + c_m\alpha_m \in W$

So, $W + \alpha = W + (\gamma + d_1\beta_1 + d_2\beta_2 + ...d_l\beta_l)$

$= (W + \gamma) + d_1\beta_1 + d_2\beta_2 + ...d_l\beta_l$

$= W + (d_1\beta_1 + d_2\beta_2 + ...d_l\beta_l)$ [$\because \gamma \in W \Rightarrow W + \gamma = W$]

$= (W + d_1\beta_1) + (W + d_2\beta_2) + ... + (W + d_l\beta_l)$

$= d_1(W + \beta_1) + d_2(W + \beta_2) + ... + d_l(W + \beta_l)$

Thus any element $W+\alpha$ of V/W can be expressed as a linear combination of S_1.

$\therefore V/W = L(S_1)$

$\therefore S_1$ is a basis of V/W.

$\therefore \dim V/W = l$

Hence the theorem.

6.5 *DIRECT SUM OF SPACES*

In this section we will learn about the direct sum of spaces

Definition: Let $V(F)$ be the vector spaces and let $W_1, W_2, ..., W_m$ be subspaces of V. Then V is said to be the direct sum of $W_1, W_2, ..., W_m$ if every element $\alpha \in V$ can be written in one and only one way as $\alpha = \alpha_1, \alpha_2, ..., \alpha_m$ where

$$\alpha_1 \in W_1, \alpha_2 \in W_2, ..., \alpha_m \in W_m$$

If a vector space $V(F)$ is a direct sum of its two subspaces W_1 and W_2 the we should have not only $V = W_1 + W_2$ but also that each vector of V can be uniquely expressed as sum of an element of W_1 and an element of W_2. Symbolically the direct sum is represented by the notation $V = W_1 \oplus W_2$.

Example 1: Let $V_2(F)$ be the vector space of all ordered pairs of F. Then $W_1 = \{(a,0) : a \in F\}$ and $W_2 = \{(0,b) : b \in F\}$ are two subspaces of $V_2(F)$. Obviously any element $x, y \in V_2(F)$ can be uniquely expressed as sum two elements from which one of them belongs to W_1 and other element will be belong in W_2. The unique expression is defined by $(x, y) = (x,0) + (0, y)$. Thus $V_2(F)$ is the direct sum of W_1 and W_2. Only the zero element $(0,0)$ is the only common element in both W_1 and W_2.

6.6 *DISJOINT SUBSPACE*

Definition: Two subspaces W_1 and W_2 of the vector space $V(F)$ are said to be disjoint if their intersection is the zero subspace i.e., $W_1 \cap W_2 = \{0\}$.

Theorem 3: The necessary and sufficient conditions for a vector space $V(F)$ to be a direct sum of its two subspaces W_1 and W_2 are that

(i) $V = W_1 + W_2$ and

(ii) $W_1 \cap W_2 = \{0\}$ i.e., W_1 and W_2 are disjoint

Proof: (Necessary condition) Let V be the direct sum of its two subspaces W_1 and W_2. Then element of V is expressed uniquely as sum of the element of W_1 and element of W_2. Therefore we have $V = W_1 + W_2$.

Let if possible $0 \neq \alpha \in W_1 \cap W_2$. Then $\alpha \in W_1, \alpha \in W_2$. Also $\alpha \in V$ and we can write

$\alpha = 0 + \alpha$ where $0 \in W_1$, $\alpha \in W_2$

And $\alpha = 0 + \alpha$ where $\alpha \in W_1, 0 \in W_2$.

Thus $\alpha \in V$ can be expressed by two different ways as a sum of an element belongs to W_1 and an element of W_2. This is contradict the fact that V be the direct sum of W_1 and W_2. Hence 0 is the only common vector in the both subspaces W_1 and W_2 i.e., $W_1 \cap W_2 = \{0\}$.

Therefore, the condition is necessary.

Sufficient condition: Let $V = W_1 + W_2$ and $W_1 \cap W_2 = \{0\}$. Then we have to show that V is the direct sum of W_1 and W_2.

Since $V = W_1 + W_2$, then each element of V can be expressed as linear sum of the elements of W_1 and W_2. So, we have to prove that this expression is unique. For it we assume that,

$\alpha = \alpha_1 + \alpha_2, \alpha \in V, \alpha_1 \in W_1, \alpha_2 \in W_2$

And $\alpha = \beta_1 + \beta_2, \beta \in V, \beta_1 \in W_1, \beta_2 \in W_2$. Now we only to prove that $\alpha_1 = \beta_1$ and $\alpha_2 = \beta_2$

We have $\alpha_1 + \alpha_2 = \beta_1 + \beta_2$

$\Rightarrow \alpha_1 - \beta_1 = \beta_2 - \alpha_2$

Since W_1 is subspace, therefore

$\alpha_1 \in W_1, \beta_1 \in W_1 \Rightarrow \alpha_1 - \beta_1 \in W_1$

Similarly, $\beta_2 - \alpha_2 \in W_2$.

$\therefore \alpha_1 - \beta_1 = \beta_2 - \alpha_2 \in W_1 \cap W_2$

But we know zero vector is only common vector in W_1 and W_2 i.e., $\{0\} \in W_1 \cap W_2$. Therefore $\alpha_1 - \beta_1 = 0 \Rightarrow \alpha_1 = \beta_1$. Also $\beta_2 - \alpha_2 = 0 \Rightarrow \alpha_2 = \beta_2$.

Thus each vector $\alpha \in V$ is uniquely expressible as sum of an element of W_1 and an element of W_2.

Hence $V = W_1 \oplus W_2$.

6.7 *DIMENSION OF A DIRECT SUM*

Theorem 4: If $V(F)$ be a finite dimensional vector space and $V(F)$ is a direct sum of two subspaces W_1 and W_2, then $\dim V = \dim W_1 + \dim W_2$

Proof: Let $\dim W_1 = m$ and $\dim W_2 = l$. Also let the sets of vectors

$S_1 = \{\alpha_1, \alpha_2, ..., \alpha_m\}$ and $S_2 = \{\beta_1, \beta_2, ..., \beta_l\}$ be the bases of W_1 and W_2 respectively.

$\dim W_1 + \dim W_2 = m + l$.

In order to prove that $\dim V = m + l$. We claim that the set

$S = S_1 \cup S_2 = \{\alpha_1, \alpha_2, ..., \alpha_m, \beta_1, \beta_2, ..., \beta_l\}$ is a basis of V.

First we will prove that the set S is linearly independent. Let

$$a_1\alpha_1 + a_2\alpha_2 + ... + a_m\alpha_m + b_1\beta_1 + b_2\beta_2 + ... + b_l\beta_l = 0$$

$$\Rightarrow a_1\alpha_1 + a_2\alpha_2 + ... + a_m\alpha_m = -(b_1\beta_1 + b_2\beta_2 + ... + b_l\beta_l).$$

Now $a_1\alpha_1 + a_2\alpha_2 + ... + a_m\alpha_m \in W_1$

And $-(b_1\beta_1 + b_2\beta_2 + ... + b_l\beta_l) \in W_2$

Thus, $a_1\alpha_1 + a_2\alpha_2 + ... + a_m\alpha_m \in W_1 \cap W_2$

and $-(b_1\beta_1 + b_2\beta_2 + ... + b_l\beta_l) \in W_1 \cap W_2$. Since V is the direct sum of W_1 and W_2. Therefore 0 is the only vector belonging to $W_1 \cap W_2$. Then we have

$$a_1\alpha_1 + a_2\alpha_2 + ... + a_m\alpha_m = 0, b_1\beta_1 + b_2\beta_2 + ... + b_l\beta_l = 0$$

As we know both the set $\{\alpha_1, \alpha_2, ..., \alpha_m\}$ and $\{\beta_1, \beta_2, ..., \beta_l\}$ are linearly independent therefore we have,

$a_1 = 0, a_2 = 0, ..., a_m = 0, b_1 = 0, b_2 = 0, ..., b_l = 0$

Therefore S is linearly independent.

Now we have to show that $L(S) = V$. Let the vector α be any arbitrary element of V. Then,

$\alpha =$ an element of $W_1 +$ an element of W_2

$=$ a linear combination of $S_1 +$ a linear combination of S_2

$=$ a linear combination of element of S.

$\therefore S$ is basis of V. Therefore $\dim V = m + l$

Hence the theorem

Theorem 5: Let V be a finite dimensional vector space and let W_1 and W_2 be subspaces of V such that $V = W_1 + W_2$ and $\dim V = \dim W_1 + \dim W_1$. Then $V = W_1 \oplus W_2$.

Proof: Let $\dim W_1 = l$ and $\dim W_2 = m$. Then,

$\dim V = l + m$

Let $S_1 = \{\alpha_1, \alpha_2, ..., \alpha_l\}$ be a basis of W_1 and $S_2 = \{\beta_1, \beta_2, ..., \beta_m\}$ be the basis of W_2. First, we will prove that $S_1 \cup S_2$ is a basis of V.

Let $\alpha \in V$. Since $V = W_1 + W_2$, therefore we can rewrite

$\alpha = \gamma + \delta$ where $\gamma \in W_1, \delta \in W_2$.

Now $\gamma \in W_1$ can be write as a linear combination of elements of S_1 and $\delta \in W_2$ can be write as linear combination of the element of S_2. Thus each vector $\alpha \in V$ can be written as linear combination of the element of $S_1 \cup S_2$.

$\Rightarrow V = L(S_1 \cup S_2)$.

Since $\dim V = l + m$ and $L(S_1 \cup S_2) = V$, it means number of distinct element in $S_1 \cup S_2$ cannot be less than $l + m$. Thus $S_1 \cup S_2$ is a basis of V. Therefore, the set

$\{\alpha_1, \alpha_2, ..., \alpha_l, \beta_1, \beta_2, ..., \beta_m\}$ is linearly independent.

Now we have to prove that $W_1 \cap W_2 = \{0\}$.

Let $\alpha \in W_1 \cap W_2$. Then $\alpha_1 \in W_1 \cap W_2$. Then $\alpha \in W_1, \alpha \in W_2$.

Therefore, $\alpha = a_1\alpha_1 + a_2\alpha_2 + ... + a_l\alpha_l$

and $\alpha = b_1\beta_1 + b_2\beta_2 + ... + b_m\beta_m$

for some a's and b's $\in F$.

$\therefore a_1\alpha_1 + a_2\alpha_2 + ... + a_l\alpha_l = b_1\beta_1 + b_2\beta_2 + ... + b_m\beta_m$

$\Rightarrow a_1\alpha_1 + a_2\alpha_2 + ... + a_l\alpha_l - b_1\beta_1 - b_2\beta_2 - ... - b_m\beta_m = 0$

$\Rightarrow a_1 = 0, a_2 = 0, ..., a_l = 0, b_1 = 0, b_2 = 0, ..., b_m = 0$

$\Rightarrow \alpha = 0$

$\therefore W_1 \cap W_2 = \{0\}$.

6.8 *COMPLEMENTARY SUBSPACES*

Definition: Let W_1 and W_2 be the subspaces of the vector space $V(F)$. Then the subspace W_2 is called the complement of W_1 in V if V is the direct sum of W_1 and W_2.

Theorem 6: (Existence of complementary subspaces) Corresponding to each subspaces W_1 of a finite dimensional vector space $V(F)$, there exists a subspace W_2 such that V is the direct sum of W_1 and W_2.

Proof: Let $\dim W_1 = m$. Let the set $S_1 = \{\alpha_1, \alpha_2, ..., \alpha_m\}$ be the basis of W_1. Since S_1 is a linearly independent subset of V, therefore S_1 can be extended to form a basis of V. Let the set

$S = \{\alpha_1, \alpha_2, ..., \alpha_m, \beta_1, \beta_2, ..., \beta_l\}$ be a basis of V.

Let W_2 be the subspace of V generated by the set

$S_2 = \{\beta_1, \beta_2, ..., \beta_l\}$.

We shall prove that V is the direct sum of W_1 and W_2. So, we have to show that $V = W_1 + W_2$ and $W_1 \cap W_2 = \{0\}$.

Let α be any element of V. Then we can express

$\alpha =$ a linear combination of element of S.

= a linear combination of element of S_1 + a linear combination of S_2

= an element of W_1 + an element of W_2

$\therefore V = W_1 + W_2$.

Again let $\beta \in W_1 \cap W_2$. Then β can be expressed as a linear combination of S_1 and also as a linear combination of S_2. So we have

$$\beta = a_1\alpha_1 + a_2\alpha_2 + ... + a_m\alpha_m = b_1\beta_1 + b_2\beta_2 + ... + b_l\beta_l$$

$$\therefore a_1\alpha_1 + a_2\alpha_2 + ... + a_m\alpha_m - b_1\beta_1 - b_2\beta_2 - ... - b_l\beta_l = 0$$

$\Rightarrow a_1 = 0, a_2 = 0, ..., a_m = 0, b_1 = 0, b_2 = 0, ..., b_l = 0$. Since $\alpha_1, \alpha_2, ..., \alpha_m, \beta_1, \beta_2, ..., \beta_l$ are linearly independent.

$\therefore \beta = 0$ (Zero vector)

Thus, $W_1 \cup W_2 = \{0\}$.

Now we can say that V is the direct sum of W_1 and W_2.

Theorem 7: If W_1 and W_2 are complementary subspaces of a vector space V, then the mapping f which assign to each vector β in W_2 the coset $W_1 + \beta$ is an isomorphism between W_2 and V / W_1.

Proof: We have given that,

$V = W_1 \oplus W_2$

And $f : W_2 \to V/W_1$ such that

$f(\beta) = W_1 + \beta \ \forall \ \beta \in W_2$

We shall show that f is an isomorphism of W_2 onto V/W_1.

(i) f **is one-one:** If $\beta_1, \beta_2 \in W_2$, then

$f(\beta_1) = f(\beta_2) \Rightarrow W_1 + \beta_1 = W_1 + \beta_2$ [By definition of f]

$\Rightarrow \beta_1 - \beta_2 \in W_1$

$\Rightarrow \beta_1 - \beta_2 \in W_1 \cap W_2$ [$\because \beta_1 - \beta_2 \in W_2$ because W_2 is subspace]

$\Rightarrow \beta_1 - \beta_2 = 0$ [$\because W_1 \cap W_2 = \{0\}$]

$\Rightarrow \beta_1 = \beta_2$

$\therefore f$ is one-one.

(ii) f **is onto:** Let $W_1 + \alpha$ be any coset in V/W_1, where $\alpha \in V$. Since V is direct sum of W_1 and W_2, therefore we can write

$\alpha = \gamma + \beta$ where $\gamma \in W_1, \beta \in W_2$

This gives $\gamma = \alpha - \beta \in W_1$

Since $\alpha - \beta \in W_1$, therefore $W_1 + \alpha = W_1 + \beta$

Now $f(\beta) = W_1 + \beta = W_1 + \alpha$ [By def. of f]

Thus, $W_1 + \alpha \in V/W_1 \Rightarrow \exists \beta \in W_2$ such that

$f(\beta) = W_1 + \alpha$

$\Rightarrow f$ is onto.

(iii) f **is linear transformation:** Let $a, b \in F$ and $\beta_1, \beta_2 \in W_2$. Then

Let $f(a\beta_1 + b\beta_2) = W_1 + (a\beta_1 + b\beta_2)$

$$= (W_1 + a\beta_1) + (W_1 + b\beta_2)$$

$$= a(W_1 + \beta_1) + b(W_1 + \beta_2)$$

$$= af(\beta_1) + bf(\beta_2).$$

Therefore f is a linear transformation.

Hence f is an isomorphism between W_2 and V/W_1.

Theorem 8: (Dimension of quotient space) If W is subspace of dimensional m of a $n-$ dimensional vector space V, then the dimension of the quotient space V/W is $n-m$.

Proof: As we have given that W is a subspace of vector space V. It means there exist a subspace W_1 of V such that $V = W \oplus W_1$.

Also, $\dim V = \dim W + \dim W_1$

Or, $\dim W_1 = \dim V - \dim W = n - m$

Thus by the theorem 7, we have

$V/W \cong W_1$

$\therefore \dim V/W = \dim W_1 = n - m$

6.9 *DIRECT SUM OF SEVERAL SUBSPACES*

We will now talk about the direct sum of many subspaces. In order to accomplish this, we must first define the idea of subspace independence, which is comparable to the disjointness requirement of two subspaces.

Definition: Suppose $W_1, W_2, ..., W_k$ are subspaces of the vector space V. We shall say that $W_1, W_2, ..., W_k$ are independent if $\alpha_1, \alpha_2, ..., \alpha_k = 0, \alpha_i \in W_i$ implies that each $\alpha_i = 0$.

Some important properties:

1. If $V(F)$ be a vector space and $W_1, W_2, ..., W_k$ are subspaces of the vector space V such that $W = W_1 + W_2 + ... + W_k$. Then following are the equivalent

(i) $W_1, W_2, ..., W_k$ are independent.

(ii) Each vector $\alpha \in W$ can uniquely expressed in the form $\alpha = \alpha_1 + \alpha_2 + ... + \alpha_k$ with $\alpha_i \in W_i$ for $i = 1, 2, ..., k$

(iii) For each $i, 2 \le i \le k$, the subspaces W_i is disjoint from the sum $(W_1 + W_2 + ... + W_{i-1})$.

2. If $V(F)$ be a vector space and $W_1, W_2, ..., W_n$ are subspaces of the vector space V. Suppose that $V = W_1 + W_2 + ... + W_n$ and that $W_i \cap (W_1 + W_2 + ... + W_{i-1} + W_{i+1} + ... + W_n) = \{0\}$ for ever $i = 1, 2, ..., n$. Prove that V is the direct sum of $W_1, W_2, ..., W_n$.

3. If $V(F)$ be the finite dimensional vector space and let $W_1, W_2, ..., W_k$ be subspace of V then these statement are equivalent.

(i) V is the direct sum of $W_1, W_2, ..., W_k$.

(ii) If B_i is a basis of $W_i, i = 1, 2, ..., k$, then the union $B = \bigcup_{i=1}^{k} B_i$ is also a basis for V.

4. If a finite dimensional vector space $V(F)$ is the direct sum of its subspaces $W_1, W_2, ..., W_k$, then $\dim V = \dim W_1 + ... + \dim W_k$

6.10 *CO-ORDINATES*

Let $V(F)$ be a finite dimensional vector space and consider $B = \{\alpha_1, \alpha_2, ..., \alpha_n\}$ be an ordered basis for V. When we refer to an ordered basis, we indicate that the vectors of B have been listed in a precise manner; that is, the vectors that are fixed and occupy the first, second, ..., n^{th} positions in the set B.

Let $\alpha \in V$. Then there exists a unique n-tuple $(x_1, x_2, ..., x_n)$ of scalars such that

$$\alpha = x_1\alpha_1 + x_2\alpha_2 + ... + x_n\alpha_n = \sum_{i=1}^{n} x_i\alpha_i \ .$$

The $n-$tuple $(x_1, x_2, ..., x_n)$ is called the $n-$tuple of co-ordinates of α relative to the ordered basis B. The scalars x_i is called the i^{th} coordinates of α relative to the ordered basis B. Then $n \times 1$ matrix,

$$X = \begin{bmatrix} x_1 \\ x_2 \\ . \\ . \\ x_n \end{bmatrix}$$

Is called the coordinate matrix of α relative to the ordered basis B . Here, we use the symbol $[\alpha]_B$. For the coordinate matrix of the vector α to the ordered basis *B*.

It should be emphasized that the vector α 's coordinates are unique only for a specific ordering of *B* for the same basis set *B*. There are various ways to arrange the base set *B*. A modification in *B*'s ordering could result in a change in α 's coordinates.

Solved examples

Example 1: Prove that the set $S = \{(1,0,0),(1,1,0),(1,1,1)\}$ is a basis of $R^3(R)$ where R is field of real numbers. Hence find the coordinates of the vector (a,b,c) with respect to the above basis.

Solution: As we know that the dimension of the vector space $R^3(R)$ is 3. If the given set *S* is linearly independent, then *S* will form basis of $R^3(R)$. Let us consider any scalars (x, y, z) in *R* such that,

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

$$\Rightarrow (x + y + z, y + z, z) = (0,0,0)$$

$$\Rightarrow x + y + z = 0, y + z = 0, z = 0$$

$$\Rightarrow x = 0, y = 0, z = 0$$

$\Rightarrow S$ is linearly independent.

Thus, S is basis of $R^3(R)$.

Now to find the coordinates of (a,b,c) with respect to the ordered basis S . Let p,q,r be scalars in R such that

$$(a,b,c) = p(1,0,0) + q(1,1,0) + r(1,1,1)$$

$\Rightarrow (a,b,c) = p(1,0,0) + q(1,1,0) + r(1,1,1)$

$\Rightarrow (a,b,c) = (p+q+r, q+r, r=c)$

$\Rightarrow p+q+r = a, q+r = b, r = c$

$\Rightarrow r = c, q = b-c, p = a-b$

Hence the co-ordinates of the vector (a,b,c) are (p,q,r) i.e., $(a-b, b-c, c)$.

Check your progress

Problem 1: If $V(F)$ be a finite dimensional vector space and $V(F)$ is a direct sum of two subspaces W_1 and W_2, then $\dim V =$ ....

Problem 2: Find the coordinates of the vector $(2,1,-6)$ of R^3 relative to the basis $\alpha_1 = (1,1,2)$, $\alpha_2 = (3,-1,0), \alpha_3 = (2,0,-1)$.

Problem 3: Check the necessary and sufficient conditions for a vector space $V(F)$ to be a direct sum of its two subspaces W_1 and W_2.

6.11 *SUMMARY*

In this unit, we have learned about the important concept of quotient space, dimension of quotient space, direct sum of spaces, disjoint subspaces, complementary subspaces and co-ordinates of vector spaces. The overall summarization of this units are as follows:

- If $V(F)$ is a vector space, then we shall give a vector space structure to the set V/W over the same field F.
- $\dim(V/W) = \dim V - \dim W$
- If a vector space $V(F)$ is a direct sum of its two subspaces W_1 and W_2 the we should have not only $V = W_1 + W_2$ but also that each vector of V can be uniquely expressed as sum of an element of W_1 and an element of W_2
- Two subspaces W_1 and W_2 of the vector space $V(F)$ are said to be disjoint if their intersection is the zero subspace.

- The necessary and sufficient conditions for a vector space $V(F)$ to be a direct sum of its two subspaces W_1 and W_2 are that
  - $V = W_1 + W_2$ and
  - $W_1 \cap W_2 = \{0\}$ i.e., W_1 and W_2 are disjoint
- If W_1 and W_2 be the subspaces of the vector space $V(F)$. Then the subspace W_2 is called the complement of W_1 in V if V is the direct sum of W_1 and W_2.
- Corresponding to each subspaces W_1 of a finite dimensional vector space $V(F)$, there exists a subspace W_2 such that V is the direct sum of W_1 and W_2.

6.12 *GLOSSARY*

- Quotient space
- Dimension of Quotient space
- Direct sum of spaces
- Co-ordinates

6.13 *REFERENCES*

- Gel′fand I.M. Lectures on linear algebra (1989), Courier Corporation.
- Kenneth Hoffman & Ray Kunze, Linear Algebra (2nd edition)(2015), Prentice-Hall.
- David C. Lay, Linear Algebra and its Application (3rd Edition) (2007) Pearson Education Asia, India Reprint.
- Seymour Lipshutz and Marc Lipson, Schaum's outlines "Linear Algebra" (3rd Edition)(2012), Mc Graw Hill Education.
- J. N. Sharma and A. R. Vasistha, Linear Algebra (29th Edition) (1999), Krishna Prakashan.

6.14 *SUGGESTED READING*

- Minking Eie & Shou-Te Chang (2020), A First Course In Linear Algebra, World Scientific.
- Axler, Sheldon (2015), Linear algebra done right. Springer.
- https://nptel.ac.in/courses/111106051
- https://archive.nptel.ac.in/courses/111/104/111104137

6.15 *TERMINAL QUESTION*

Long Answer Type Question:

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

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

2. Show that the set $S = \{(1,0,0),(1,1,0),(1,1,1)\}$ is a basis of $C^3(C)$ where C is the field of complex numbers. Hence find the coordinates of the vector $(3+4i, 6i, 3+7i)$ in C^3 with respect to the mentioned basis.

3. Let $B = \{\alpha_1, \alpha_2, \alpha_3\}$ be an ordered basis for R^3, where $\alpha_1 = (1,0,-1), \alpha_2 = (1,1,1), \alpha_3 = (1,0,0)$. Obtain the coordinates of the vector (a,b,c) in the ordered basis B.

4. Let V be the vector space of all polynomial functions of degree less than or equal to two from the field of real number R into itself. For a fixed $t \in R$, let $g_1(x) = 1, g_2(x) = x + t, g_3(x) = (x+t)^2$. Prove that $\{g_1, g_2, g_3\}$ is a basis for V and obtain the coordinates of $c_0(x) + c_1 x + c_2 x^2$ is the ordered basis.

5. Let V be a finite-dimensional vector space and let $W_1, W_2, ..., W_k$ be subspaces of V such that $V = W_1 + W_2 + ... + W_k$ and $\dim V = \dim W_1 + \dim W_2 + ... + \dim W_k$. Then show that $V = W_1 \oplus W_2 \oplus ... \oplus W_k$

6. If W is any subset of a vector space $V(F)$, then show that the set V/W of all cosets $W + \alpha$ where α is any arbitrary element of V, is a vector space over F for the addition and scalar multiplication compositions defined as follows:

$$(W+\alpha) + (W+\beta) = W + (\alpha+\beta) \,\forall\, \alpha, \beta \in V$$

and $a(W+\alpha) = W + a\alpha; a \in F, \alpha \in V$

7. Show that if W be a subspace of a finite dimensional vector space $V(F)$, then

$$\dim(V/W) = \dim V - \dim W$$

8. If W_1 and W_2 are complementary subspaces of a vector space V, then show that the mapping f which assign to each vector β in W_2 the coset $W_1 + \beta$ is an isomorphism between W_2 and V/W_1.

Short Answer Type Question:

1. If $V(F)$ be a finite dimensional vector space and $V(F)$ is a direct sum of two subspaces W_1 and W_2, then prove that $\dim V = \dim W_1 + \dim W_2$

2. Prove that the set $S = \{(1,0,0),(1,1,0),(1,1,1)\}$ is a basis of $R^3(R)$ where R is field of real numbers. Hence find the coordinates of the vector (a,b,c) with respect to the above basis.

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

4. If V be a finite dimensional vector space and let W_1 and W_2 be subspaces of V such that $V = W_1 + W_2$ and $\dim V = \dim W_1 + \dim W_1$. Then show that $V = W_1 \oplus W_2$.

5. Prove that $\dim(V/W) = \dim(V) - \dim(W)$.

6. Corresponding to each subspaces W_1 of a finite dimensional vector space $V(F)$, there exists a subspace W_2 such that V is the direct sum of W_1 and W_2.

Fill in the blanks:

1. The dimension of quotient space V/W is …………………

2. The field of any vector space and its quotient space is …………………

3. Two subspaces W_1 and W_2 of the vector space $V(F)$ are said to be disjoint if ………….

4. If $V(F)$ be a finite dimensional vector space and $V(F)$ is a direct sum of two subspaces W_1 and W_2, then $\dim V =$…………………..

5. If a finite dimensional vector space $V(F)$ is the direct sum of its subspaces $W_1, W_2, ..., W_k$, then $\dim V =$…………………

6.16 *ANSWERS*

Answers of check your progress:

1: $\dim W_1 + \dim W_2$

2: $(-7/8, -15/8, 17/4)$

3: $V = W_1 + W_2$ and $W_1 \cap W_2 = \{0\}$ i.e., W_1 and W_2 are disjoint

Answers of long answer type question:

2. $(3-2i, -3-i, 3+7i)$

3. $(b-c, b, a-2b+c)$

4. $(c_0 - c_1 t + c_2 t^2, c_1 - 2c_2 t, c_2)$

Answer of short answer type question

1. $(a-b, b-c, c)$

2. Take vector space $V = R^2$ and $W_1 = \{(a,0) : a \in R\}$, $W_2 = \{(0,a) : a \in R\}$ and $W_3 = \{(a,a) : a \in R\}$

Answer of fill in the blanks questions:

1. $\dim V - \dim W$ **2.** Same **3.** $W_1 \cap W_2 = \{0\}$

4. $\dim W_1 + \dim W_2$ **5.** $\dim W_1 + \ldots\ldots.. + \dim W_k$

UNIT-7: LINEAR FUNCTION

CONTENTS

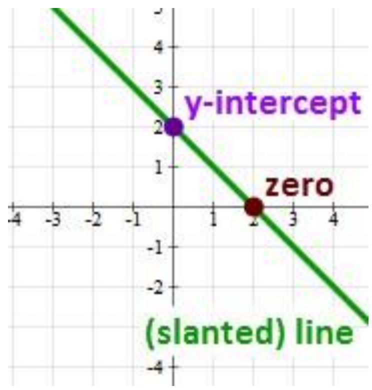
7.1 *INTRODUCTION*

See Linear function (calculus) for the definition of this term in that context.

The word "linear function" in mathematics refers to two different but related concepts:

- A linear function, or polynomial function of degree zero or one, is defined in calculus and related fields as a function whose graph is a straight line. The term "affine function" is frequently used to distinguish such a linear function from the other idea.
- A linear function is a linear map in linear algebra, mathematical analysis, and functional analysis.

A linear function from the real numbers to the real numbers in calculus and related mathematical domains is a function whose graph in Cartesian coordinates is a non-vertical line in the plane. The fact that the change in the output is proportionate to the change in the input is a defining characteristic of linear functions.



Graph of the linear function: $y(x) = -x+2$

7.2 OBJECTIVES

After reading this unit learners will be able to

- Understand the basic difference between linear functional and linear transformation.
- Visualized some special types of linear functional and properties of linear functional.
- Implementation the concept of dual space.
- Visualized and understand the important theorems of linear functional and dual space.
- Understand the concept of reflexivity and annihilators.

7.3 SOME OVERVIEW ON LINEAR TRANSFORMATION

In the previous unit we have already learned about the linear transformation. To learn the basic difference between linear transformation and linear function we initially recall or summarized about the linear transformation briefly.

Definition: Let *U* and *V* are two vector space then a mapping $T: U \to V$ is called a Linear Transformation if it satisfies the following condition:

1. $\forall\, x, y \in U$, $T(x + y) = T(x) + T(y)$

2. $\forall\, x \in U$, $\alpha \in R$, $T(\alpha x) = \alpha T(x)$

Definition: Let $T: U \to V$ and $S: U \to V$ be two Linear Transformation then the sum of *T* and *S* is denoted by $T + S$ and defined as $T + S$: U → V

$$(T+S)(\bar{x}) = T(\bar{x}) + S(\bar{x})\ \forall\, \bar{x} \in U$$

Example 1 $T : \mathbb{R}^2 \to \mathbb{R}^3,\ T(\overline{x}) = (x+y, x-y, 0),\ S : \mathbb{R}^2 \to \mathbb{R}^3,\ S(\overline{x}) = (x-y, x+y, 2x),$ *then find* $(T+S)$.

Solution:

$$\begin{aligned}(T+S)(\overline{x}) &= T(\overline{x}) + S(\overline{x})\\ &= (x+y, x-y, 0) + (x-y, x+y, 2x)\\ &= (x+y+x-y, x-y+x+y, 0+2x)\\ &= (2x, 2x, 2x)\end{aligned}$$

Definition: Let $T: U \to V$ be a Linear Transformation and let α be a scalar then the scalar multiplication of a linear transformation *T* by α denoted by αT and defined as $\alpha T : U \to V$

$$(\alpha T)(\bar{x}) = \alpha T(\bar{x}),\ \forall\ \bar{x} \in U$$

Definition: The set *f* all Linear Transformation from *U* to *V* is denoted by $L(U, V)$.

$$L(U,V) = \{T \mid T : U \to V \text{ is a linear transformation}\}$$

Definition: Let $T: U \to V$ be a linear transformation and let $S: V \to W$ be a linear transformation then, the composition of *S* and *T* is denoted by *SoT* and defined as $SoT: U \to W$.

$$SoT = S(T(\bar{x})), \forall\, \bar{x} \in U$$

Theorem 1: Prove that the sum of two linear transformations is also linear transformation.

OR

If $T, S \in L(U, V)$ then prove that $S + T \in L(U, V)$.

Proof: Here $T, S \in L(U, V)$ i.e. $T: U \rightarrow V$ and $S: U \rightarrow V$ are linear transformation and we have to prove

$S + T: U \rightarrow V$ is also linear transformation.

(i) Let $\overline{x}, \overline{y} \in U$ *to prove that* $(S+T)(\overline{x}+\overline{y}) = (S+T)(\overline{x}) + (S+T)(\overline{y})$

$$\begin{aligned}(S+T)(\overline{x}+\overline{y}) &= S(\overline{x}+\overline{y}) + T(\overline{x}+\overline{y})\\ &= (S(\overline{x}) + S(\overline{y})) + (T(\overline{x}) + T(\overline{y}))\\ &= S(\overline{x}) + T(\overline{x}) + S(\overline{y}) + T(\overline{y})\\ &= (S+T)(\overline{x}) + (S+T)(\overline{y})\end{aligned}$$

(ii) Let $\alpha \in \mathbb{R}$ *and let* $x \in U$ *to prove that* $(S+T)(\alpha\overline{x}) = \alpha(S+T)(\overline{x})$.

$$\begin{aligned}(S+T)(\alpha\overline{x}) &= S(\alpha\overline{x}) + T(\alpha\overline{x})\\ &= \alpha S(\overline{x}) + \alpha T(\overline{x})\\ &= \alpha(S(\overline{x}) + T(\overline{x}))\\ &= \alpha(S+T)(\overline{x})\end{aligned}$$

So from (i) and (ii) $S+T : U \rightarrow V$ *is also linear transformation*

Theorem 2: If $T \in L(U,V)$ and $\alpha \in R$ then prove that $\alpha T \in L(U,V)$

Proof: Here $T : U \rightarrow V$ is a linear transformation and α be a scalar to prove that $\alpha T : U \rightarrow V$ is a linear transformation.

(i) Let $x, y \in U$ *to prove that* $(\alpha T)(\overline{x}+\overline{y}) = (\alpha T)(\overline{x}) + (\alpha T)(\overline{y})$

$$\begin{aligned}(\alpha T)(\overline{x}+\overline{y}) &= \alpha(T(\overline{x}+\overline{y}))\\ &= \alpha(T(\overline{x}) + T(\overline{y}))\\ &= \alpha T(\overline{x}) + \alpha T(\overline{y})\\ &= (\alpha T)(\overline{x}) + (\alpha T)(\overline{y})\end{aligned}$$

(ii) Let $x \in U$ and let β be a scalar $\beta \in \mathbb{R}$ to prove that $(\alpha T)(\beta \overline{x}) = \beta((\alpha T)(\overline{x}))$

$$\begin{aligned}(\alpha T)(\beta \overline{x}) &= \alpha(T(\beta(\overline{x}))) \\ &= \alpha(\beta(T(\overline{x}))) \\ &= (\alpha\beta)T(\overline{x}) \\ &= (\beta\alpha)T(\overline{x}) \\ &= \beta((\alpha T)(\overline{x}))\end{aligned}$$

From (i) and (ii) $\alpha T : U \to V$ is a linear transformation.

Theorem 3 *The composition of two linear transformation is also a linear transformation.*

OR

If $T \in L(U, V)$ and $S \in L(V, W)$,then prove that $SoT \in L(V, W)$.

Proof: *Here $T \in L(U, V)$, so $T : U \to V$ is a linear transformation and $S \in L(V, W)$ so $S : V \to W$ is a linear transformation.*
And we have to prove that $SoT : U \to W$ is also linear transformation.

(i) Let $\overline{x}, \overline{y} \in U$ to prove that $(SoT)(\overline{x} + \overline{y}) = (SoT)(\overline{x}) + (SoT)(\overline{y})$.

$$\begin{aligned}(SoT)(\overline{x} + \overline{y}) &= S(T(\overline{x} + \overline{y})) \\ &= S(T(\overline{x}) + T(\overline{y})) \\ &= S(T(\overline{x})) + S(T(\overline{y})) \\ &= (SoT)(\overline{x}) + (SoT)(\overline{y})\end{aligned}$$

(ii) Let $\overline{x} \in U$ and let α be a scalar to prove that $(SoT)(\alpha \overline{x}) = \alpha((SoT)(\overline{x}))$.

$$\begin{aligned}(SoT)(\alpha \overline{x}) &= S(T(\alpha \overline{x})) \\ &= S(\alpha T(\overline{x}))) \\ &= \alpha(S(T(\overline{x}))) \\ &= \alpha((SoT)(\overline{x}))\end{aligned}$$

So from (i) and (ii) $SoT : U \to W$ is a linear transformation.

Example 2 *Let* $T : \mathbb{R}^2 \to \mathbb{R}^2$, $T(x,y) = (x-y, x+y)$, $S : \mathbb{R}^2 \to \mathbb{R}^2$, $S(x,y) = (x+y, x-y)$ *then find* SoT *and* ToS.

Solution: *Let* $(x,y) \in \mathbb{R}^2$

$$\begin{aligned}
(SoT)(x,y) &= S(T(x,y)) \\
&= S(x-y, x+y) \\
&= (x-y+x+y, x-y-x-y) \\
&= (2x, -2y) \\
(ToS)(x,y) &= T(S(x,y)) \\
&= T(x+y, x-y) \\
&= (x+y-x+y, x+y+x-y) \\
&= (2y, 2x)
\end{aligned}$$

7.4 *LINEAR FUNCTIONAL*

Definition: Consider a vector space *V*(*F*). It is known that a vector space over *F* can be thought of as the field *F*. This is *F*(*F*) or F^1, the vector space. We'll just refer to it as *F*. A linear functional on *V* is a linear translation from *V* into *F*. The independent definition of a linear functional will now be provided. In this unit, we often take *R* to be a filed in terms of *F*.

Definition: Let V(F) be a vector space. A mapping from V into F is said to be a linear functional on V if,

$$f(a\alpha + b\beta) = af(\alpha) + bf(\beta) \ \forall \ a,b \in F \text{ and } \forall \ \alpha, \beta \in V$$

If *f* is a linear functional on *V*(*F*), then $f(\alpha)$ is in *F* for each α belonging to V. Since $f(\alpha)$ is a scalar, therefore a linear functional on *V* is a scalar valued function.

Example 3: Let $V_n(F)$ be the vector space of ordered *n*-tuples of the elements of the field *F*.

Let $x_1, x_2, ..., x_n$ be *n*-field elements of F.

If $\alpha = (a_1, a_2, ..., a_n) \in V_n(F)$.

Let *f* be a function form $V_n(F)$ into F defined by

$$f(\alpha) = x_1a_1 + x_2a_2 + ... + x_na_n$$

Let $\beta = (b_1, b_2, ..., b_n) \in V_n(F)$. If $a, b \in F$, we have

$$f(a\alpha + b\beta) = f[a(a_1, a_2, ..., a_n) + b((b_1, b_2, ..., b_n)$$

$$= f(aa_1 + bb_1, aa_2 + bb_2, ...,, aa_n + bb_n)$$

$$= x_1(aa_1 + bb_1) + x_2(aa_2 + bb_2),...,+x_n(aa_n + bb_n)$$

$$= af(a_1, a_2, ..., a_n) + bf(b_1, b_2, ..., b_n)$$

$$= af(\alpha) + bf(\beta)$$

Hence f is a linear functional on $V_n(F)$

Example 4: We will now present a important illustration of a linear functional.

We shall prove that the trace function is a linear functional on the space of all $n \times n$ matrices over a field *F*.

Let n be a positive integer and a field. Let $V(F)$ be the vector space of all $n \times n$ matrices over *F*. If $A = [a_{ij}]_{n\times n} \in V,$ then the trace of A is the scalar.

$$tr\ A = a_{11} + a_{12} + ... + a_{nn} = \sum_{i=1}^{n} a_{ii}.$$

Therefore, the scalar that results from summing the components of *A* that are located along the principal diagonal is the trace of *A*.

The trace function is a linear functional on V because if

$a, b \in F$ and $A = [a_{ij}]_{n\times n}, B = [b_{ij}]_{n\times n} \in V,$ then

$$tr(aA + bB) = tr(a[a_{ij}]_{n\times n} + b[b_{ij}]_{n\times n}) = tr([aa_{ij} + bb_{ij}]_{n\times n})$$

$$== \sum_{i=1}^{n}(aa_{ii} + bb_{ii}) = a\sum_{i=1}^{n} a_{ii} + \sum_{i=1}^{n} b_{ii} = a(trA) + b(trB)$$

7.4.1 *SOME SPECIAL LINEAR FUNCTIONAL*

1. **Zero functional:** Let V be a vector space over the field F. The function f from V into F defined by

$$f(\alpha) = 0 \, \forall \alpha \in V$$

is a linear functional on V.

Proof: Let $\alpha, \beta \in V$ and $a, b \in F$. We have

$f(a\alpha + b\beta) = 0$

$= a0 + b0 = af(\alpha) + bf(\beta)$

$\therefore f$ is a linear function on V. It is called the zero functional and we shall in future denoted it by $\hat{0}$

2. **Negative of linear functional:** Let V be a vector space over the field F. Let f be a linear functional on V. The correspondence $-f$ defined by

$$(-f)(\alpha) = -[f(\alpha)] \forall \alpha \in V$$

is a linear functional on V.

Proof: Since $f(\alpha) \in F \Rightarrow -f(\alpha) \in F$, therefore $-f$ is a function from V into F.

Let $a, b \in F$ and $\alpha, \beta \in V$. Then

$(-f)(a\alpha + b\beta) = -[f(a\alpha + b\beta)]$ [By definition of $-f$]

$= -[af(\alpha) + bf(\beta)]$ [Since f is a linear functional]

$= a[-f(\alpha)] + b[-f(\beta)]$

$= a[(-f)\alpha] + b[(-f)\beta]$

$\therefore -f$ is a linear functional on V.

Properties of linear functional

7.4.2 *PROPERTIES OF LINEAR FUNCTIONAL*

Theorem 4: Let f be a linear functional on a vector space $V(F)$. Then

(i) $f(0)=0$ where 0 on the left hand side is zero vector of V, and 0 on the right hand side is zero element of F.

(ii) $f(-\alpha)=-f(\alpha)\,\forall\,\alpha\in V$.

Proof: Let $\alpha\in V$. Then $f(\alpha)\in F$.

We have $f(\alpha)+0=f(\alpha)\in F$ [$\because 0$ is zero element of F]

$=f(\alpha+0)$ [$\because 0$ is zero element of V]

$=f(\alpha)+f(0)$ [$\because f$ is a linear functional]

Now F is a field. Therefore,

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

$\Rightarrow f(0)=0$, by left cancellation law for addition in F.

(iii) We have $f[\alpha+(-\alpha)]=f(\alpha)+f(-\alpha)$ [$\because f$ is a linear functional by (i)]

But $f[\alpha+(-\alpha)]=f(0)=0$

Thus in F, we have

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

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

7.5 *DUAL SPACES AND DUAL BASES*

On a vector space $V(F)$, let V' be the set of all linear functionals. This set is sometimes denoted as V^*. Our current goal is to apply a vector space structure over the same field F to the set V'. We must appropriately define addition in V' and scalar multiplication in V' over F in order to do this.

Definition: The set of all linear functional from V to F is denoted by $L(V, F)$ or V^*.

Note:

The set of all linear functional from V to $\mathbb{R}$ is denoted by $L(V,\mathbb{R})$ or V^*.

$$L(V,\mathbb{R})=V^*=\{f/f:V\to\mathbb{R}\ is\ a\ linear\ functional\}$$

Theorem 5: (State and prove the existence theorem of dual basis)

Statement: *Let* $B = \{v_1, v_2, \ldots, v_n\}$ *be a basis for a vector space* V*.let* V^* *be a dual space of* V*, suppose* $f_1, f_2, \ldots f_n \in V^*$ *such that*

$$\begin{aligned} f_i(v_j) &= 1 \qquad i = j \\ &= 0 \qquad i \neq j \quad i, j = 1, 2, \ldots n \end{aligned}$$

Then prove that $B^* = \{f_1, f_2, \ldots f_n\}$ *is a basis for* V^*.

Proof: *Here* $B = \{v_1, v_2, \ldots, v_n\}$ *be a basis for a vector space* V *and* V^* *be a dual space of* V*, and* $f_1, f_2, \ldots f_n \in V^*$ *such that*

$$\begin{aligned} f_i(v_j) &= 1 \qquad i = j \\ &= 0 \qquad i \neq j \quad i, j = 1, 2, \ldots n \end{aligned} \tag{1}$$

we have tp prove $B^* = \{f_1, f_2, \ldots f_n\}$ *is a basis for* V^*.

(i) First we shall prove that B^* *is Linearly Independent*

Consider,

$$\alpha_1 f_1 + \alpha_2 f_2 + \ldots + \alpha_n f_n = \overline{0} \quad \text{where} \ \ \alpha_i \in \mathbb{R}, \ \ i = 1, 2, \ldots, n.$$

$$(\alpha_1 f_1 + \alpha_2 f_2 + \ldots + \alpha_n f_n)(v_1) = \overline{0}(v_1)$$

$$(\alpha_1 f_1)(v_1) + (\alpha_2 f_2)(v_1) + \ldots + (\alpha_n f_n)(v_1) = 0$$

$$\alpha_1(f_1)(v_1) + \alpha_2(f_2)(v_1) + \ldots + \alpha_n(f_n)(v_1) = 0$$

$$\alpha_1(1) + \alpha_2(0) + \ldots + \alpha_n(0) = 0$$

$$\alpha_1(1) = 0$$

$$\alpha_1 = 0$$

Similarly, we can prove $\alpha_2 = 0, \alpha_3 = 0, \ldots, \alpha_n = 0$*.*
so $B^* = \{f_1, f_2, \ldots f_n\}$ *is Linearly Independent.*

(ii) Now we have to prove that $[B^*] = V^*$.

we know that $[B^*] \subseteq V^*$.
so only to prove $V^* \subseteq [B^*]$
take $f \in V^*$*, so* $f : V \to \mathbb{R}$ *is a linear functional.*
Suppose,

$$f(v_1) = \alpha_1$$
$$f(v_2) = \alpha_2$$
$$\vdots$$
$$f(v_n) = \alpha_n, \quad where \;\; \alpha_i \in \mathbb{R}, \;\; i = 1, 2, \ldots, n.$$

Let us define a function $\phi : V \to \mathbb{R}$ *such that*

$$\phi = \alpha_1 f_1 + \alpha_2 f_2 + \ldots + \alpha_n f_n \qquad (2)$$

Now,

$$\begin{aligned}
\phi(v_1) &= (\alpha_1 f_1 + \alpha_2 f_2 + \ldots + \alpha_n f_n)(v_1) \\
&= (\alpha_1 f_1)(v_1) + (\alpha_2 f_2)(v_1) + \ldots + (\alpha_n f_n)(v_1) \\
&= \alpha_1 (f_1)(v_1) + \alpha_2 (f_2)(v_1) + \ldots + \alpha_n (f_n)(v_1) \\
&= \alpha_1 (1) + \alpha_2 (0) + \ldots + \alpha_n (0) \\
\phi(v_1) &= \alpha_1
\end{aligned}$$

Similarly, we can prove

$$\phi(v_2) = \alpha_2$$
$$\phi(v_3) = \alpha_3$$
$$\vdots$$
$$\phi(v_n) = \alpha_n$$
$$So, \;\; \phi(v_i) = \alpha_i, \quad where \; i = 1, 2, \ldots n$$

also here

$$f(v_i) = \alpha_i$$
$$\phi(v_i) = f(v_i)$$
$$\phi = f$$

so by equation (2)

$$f = \alpha_1 f_1 + \alpha_2 f_2 + \ldots + \alpha_n f_n$$

$$f \in [B^*]$$
$$V^* \subseteq [B^*]$$

so by equation (2)

$$f = \alpha_1 f_1 + \alpha_2 f_2 + \ldots + \alpha_n f_n$$
$$f \in [B^*]$$
$$V^* \subseteq [B^*]$$

so,

$$[B^*] = V^*$$

so from (i) and (ii) $B^* = \{f_1, f_2, \ldots f_n\}$ *is a basis for* V^*.

Definition: Let V be a vector space and V^* be dual space of a vector space V. Let $\dim V = n$ then $\dim V^* = n$ and basis $B^* = \{f_1, f_2, ..., f_n\}$ of V^* corresponding to a basis $B = \{v_1, v_2, ..., v_n\}$ of a vector space V is called a dual basis for a vector space V.

Example 5: Discuss about the dual basis corresponding to a basis $\{(2,1),(3,1)\}$ of R^2.

Solution: As we know that R^2 is a vector space.

$\therefore \dim \mathbb{R}^2 = 2$
Let $(\mathbb{R}^2)^*$ *be a dual space of* $\mathbb{R}^2$.
$\therefore \dim(\mathbb{R}^2)^* = 2$
Also here $B = \{(2,1),(3,1)\}$ *is a basis for* $\mathbb{R}^2$.
let $v_1 = (2,1)$ *and* $v_2 = (3,1)$
to find $B^* = \{f_1, f_2\}$ *a dual basis for* $\mathbb{R}^2$.
Define function $f_1 : \mathbb{R}^2 \to \mathbb{R}$ *such that*

$$f_1(x,y) = ax + by, \quad a, b \in \mathbb{R}$$

$$f_1(x,y) = ax + by$$
$$f_1(v_1) = ax + by$$

$$f_1(2,1) = 2a + b$$
$$1 = 2a + b$$
$$2a + b = 1 \qquad (3)$$

$$\begin{aligned} f_1(x,y) &= ax+by \\ f_1(v_2) &= ax+by \\ f_1(3,1) &= 3a+b \\ 0 &= 3a+b \\ 3a+b &= 0 \end{aligned} \tag{4}$$

Solve equation (3) and (4) we get $a=-1$.
Substitute $a=-1$ *in equation (3) we get* $b=3$.
So,we get

$$f_1(x,y) = -x+3y$$

Now, we define function $f_2:\mathbb{R}^2 \to \mathbb{R}$ *such that*

$$f_2(x,y) = cx+dy, \quad c,d \in \mathbb{R}$$

$$\begin{aligned} f_2(x,y) &= cx+dy \\ f_2(v_1) &= cx+dy \\ f_2(2,1) &= 2c+d \\ 0 &= 2c+d \\ 2c+d &= 0 \end{aligned} \tag{5}$$

$$\begin{aligned} f_2(x,y) &= cx+dy \\ f_2(v_2) &= cx+dy \\ f_2(3,1) &= 3c+d \\ 1 &= 3c+d \\ 3c+d &= 1 \end{aligned} \tag{6}$$

Solve equation (5) and (6) we get $c=1$.
Substitute $c=1$ *in equation (5) we get* $d=-2$.
So,we get

$$f_2(x,y) = x-2y$$

Thus $B^* = \{f_1, f_2\}$ *is a dual basis for* $\mathbb{R}^2$.
where,

$$f_1(x,y) = -x+3y$$

$$f_2(x,y) = x - 2y$$

Example 6: Discuss about the dual basis corresponding to a basis $\{(1,0,1),(1,1,0),(0,1,1)\}$ of R^3.

Solution: As we know that R^3 is a vector space.

$\therefore \dim \mathbb{R}^3 = 3$
Let $(\mathbb{R}^3)^*$ *be a dual space of* $\mathbb{R}^3$.
$\therefore \dim(\mathbb{R}^3)^* = 3$
Also here $B = \{(1,0,1),(1,1,0),(0,1,1)\}$ *is a basis for* $\mathbb{R}^3$.
let $v_1 = (1,0,1)$, $v_2 = (1,1,0)$ *and* $v_3 = (0,1,1)$
to find $B^* = \{f_1, f_2, f_3\}$ *a dual basis for* $\mathbb{R}^3$
Let $v_1 = (1,0,1), v_2 = (1,1,0)$ *and* $v_3 = (0,1,1)$
to find $B^* = \{f_1, f_2, f_3\}$ *a dual basis for* $\mathbb{R}^3$.
Define function $f_1 : \mathbb{R}^3 \to \mathbb{R}$ *such that*

$$f_1(x,y,z) = ax + by + cz, \quad a,b,c \in \mathbb{R}$$

$$\begin{aligned} f_1(x,y,z) &= ax + by + cz \\ f_1(v_1) &= ax + by + cz \\ f_1(1,0,1) &= a + c \\ 1 &= a + c \\ a + c &= 1 \end{aligned} \tag{7}$$

$$\begin{aligned} f_1(x,y,z) &= ax + by + cz \\ f_1(v_2) &= ax + by + cz \\ f_1(1,1,0) &= a + b \\ 0 &= a + b \\ a + b &= 0 \\ a &= -b \end{aligned} \tag{8}$$

$$\begin{aligned} f_1(x,y,z) &= ax + by + cz \\ f_1(v_3) &= ax + by + cz \\ f_1(0,1,1) &= b + c \\ 0 &= b + c \\ b + c &= 0 \end{aligned} \tag{9}$$

from equation (8) $a = -b$ in equation (7) we get

$$b - c = -1 \tag{10}$$

solve equation (9) and (10) we get $b = \frac{-1}{2}$.

Substitute $b = \frac{-1}{2}$ in equation (8) we get $a = \frac{1}{2}$.
from equation (9) we get $c = \frac{1}{2}$.
Thus we get,

$$f_1(x, y, z) = \frac{1}{2}x - \frac{1}{2}y + \frac{1}{2}z$$
$$f_1(x, y, z) = \frac{1}{2}(x - y + z)$$

Similarly we define function $f_2 : \mathbb{R}^3 \to \mathbb{R}$ such that

$$f_2(x, y, z) = ax + by + cz, \quad a, b, c \in \mathbb{R}$$

$$\begin{aligned} f_2(x, y, z) &= ax + by + cz \\ f_2(v_1) &= ax + by + cz \\ f_2(1, 0, 1) &= a + c \\ 0 &= a + c \\ a + c &= 0 \\ a &= -c \end{aligned} \tag{11}$$

$$\begin{aligned} f_2(x, y, z) &= ax + by + cz \\ f_2(v_2) &= ax + by + cz \\ f_2(1, 1, 0) &= a + b \\ 1 &= a + b \\ a + b &= 1 \end{aligned} \tag{12}$$

$$\begin{aligned} f_2(x, y, z) &= ax + by + cz \\ f_2(v_3) &= ax + by + cz \\ f_2(0, 1, 1) &= b + c \\ 0 &= b + c \\ b + c &= 0 \end{aligned} \tag{13}$$

from equation (11) $a = -c$ *in equation (12) we get*

$$b - c = 1 \tag{14}$$

solve equation (13) and (14) we get $b = \frac{1}{2}$.
Substitute $b = \frac{1}{2}$ *in equation (14) we get* $c = \frac{-1}{2}$.
from equation (11) we get $a = \frac{1}{2}$.
Thus we get,

$$f_2(x, y, z) = \frac{1}{2}x + \frac{1}{2}y - \frac{1}{2}z$$
$$f_2(x, y, z) = \frac{1}{2}(x + y - z)$$

Now we define function $f_3 : \mathbb{R}^3 \to \mathbb{R}$ *such that*

$$f_3(x, y, z) = ax + by + cz, \quad a, b, c \in \mathbb{R}$$

$$\begin{aligned} f_3(x, y, z) &= ax + by + cz \\ f_3(v_1) &= ax + by + cz \\ f_3(1, 0, 1) &= a + c \\ 0 &= a + c \\ a + c &= 0 \\ a &= -c \end{aligned} \tag{15}$$

$$\begin{aligned} f_3(x, y, z) &= ax + by + cz \\ f_3(v_2) &= ax + by + cz \\ f_3(1, 1, 0) &= a + b \\ 0 &= a + b \\ a + b &= 0 \end{aligned} \tag{16}$$

$$\begin{aligned} f_3(x, y, z) &= ax + by + cz \\ f_3(v_3) &= ax + by + cz \\ f_3(0, 1, 1) &= b + c \\ 1 &= b + c \\ b + c &= 1 \end{aligned} \tag{17}$$

from equation (15) $a = -c$ in equation (16) we get

$$b - c = 0 \tag{18}$$

solve equation (17) and (18) we get $b = \frac{1}{2}$.
Substitute $b = \frac{1}{2}$ *in equation (17) we get* $c = \frac{1}{2}$.
from equation (15) we get $a = \frac{-1}{2}$.
Thus we get,

$$f_3(x, y, z) = \frac{-1}{2}x + \frac{1}{2}y + \frac{1}{2}z$$
$$f_3(x, y, z) = \frac{1}{2}(-x + y + z)$$

Thus $B^* = \{f_1, f_2, f_3\}$ *is a dual basis for* $\mathbb{R}^3$.

where,

$$f_1(x, y, z) = \frac{1}{2}(x - y + z)$$
$$f_2(x, y, z) = \frac{1}{2}(x + y - z)$$
$$f_3(x, y, z) = \frac{1}{2}(-x + y + z)$$

Theorem 6: Let $\{v_1, v_2, ..., v_n\}$ be a basis for V and $\{f_1, f_2, ..., f_n\}$ be a basis for V^* then

prove that for any $v \in V$

$$v = f_1(v)v_1 + f_2(v)v_2 + \ldots + f_n(v)v_n$$

and for any $f \in V^*$

$$f = f(v_1)f_1 + f(v_2)f_2 + \ldots + f(v_n)f_n$$

Proof: Since $B = \{v_1, v_2, ..., v_n\}$ is a basis of a vector basis for V and $B^* = \{f_1, f_2, ..., f_n\}$ be a basis for V^*

$\therefore$ *B is linearly independent and* $[B] = V$ *and*
B^* is linearly independent and $[B^*] = V^*$

(i) Let

$$\begin{aligned}
&v \in V \\
&v \in V = [B] \\
&v \in [B] \\
&v = \alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_v, \ \ \forall\, \alpha_i \in \mathbb{R}, \ \ i = 1, 2, \ldots n
\end{aligned} \tag{19}$$

$$\begin{aligned}
f_1(v) &= f_1(\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_v) \\
&= f_1(\alpha_1 v_1) + f_1(\alpha_2 v_2) + \ldots + f_1(\alpha_n v_n) \\
&= \alpha_1(f_1(v_1)) + \alpha_2(f_1(v_2)) + \ldots + \alpha_n(f_1(v_n)) \\
&= \alpha_1(1) + \alpha_2(0) + \ldots + \alpha_n(0) \\
f_1(v) &= \alpha_1
\end{aligned}$$

Similarly we can prove that,

$$\begin{aligned}
f_2(v) &= \alpha_2 \\
f_3(v) &= \alpha_3 \\
&\vdots \\
f_{1n}(v) &= \alpha_n
\end{aligned}$$

Substitute this values in equation (19) we get,

$$v = f_1(v)v_1 + f_2(v)v_2 + \ldots + f_n(v)v_n$$

(ii) Let

$$\begin{aligned}
&f \in V^* \\
&f \in V^* = [B^*] \\
&f \in [B^*] \\
&f = \alpha_1 f_1 + \alpha_2 f_2 \ldots + \alpha_n f_n, \ \ \forall\, \alpha_i \in \mathbb{R}, \ \ i = 1, 2, \ldots n
\end{aligned} \tag{20}$$

$$\begin{aligned}
f(v_1) &= (\alpha_1 f_1 + \alpha_2 f_2 \ldots + \alpha_n f_n)(v_1) \\
&= (\alpha_1 f_1)(v_1) + (\alpha_2 f_2)(v_1) + \ldots + (\alpha_n f_n)(v_1) \\
&= \alpha_1(f_1(v_1)) + \alpha_2(f_2(v_1)) + \ldots + \alpha_n(f_n(v_1)) \\
&= \alpha_1(1) + \alpha_2(0) + \ldots + \alpha_n(0) \\
f(v_1) &= \alpha_1
\end{aligned}$$

Similarly we can prove that,

$$f(v_2) = \alpha_2$$
$$f(v_3) = \alpha_3$$
$$\vdots$$
$$f(v_n) = \alpha_n$$

Substitute this values in equation (20) we get,

$$f = f(v_1)f_1 + f(v_2)f_2 + \ldots + f(v_n)f_n$$

7.6 *REFLEXIVITY*

Every vector space V has a dual space V' that contains all linear functionals on V, as we know. And now V' is a vector space as well. It will therefore likewise have a dual space $(V')'$ made up of linear functionals on V'. For the sake of convenience, we will refer to this dual space as V'' and term it the **second dual space** of *V*.

Note: If *V* is finite-dimensional then $\dim V = \dim V' = \dim V''$. Which means these are isomorphic to each other.

Theorem 7: Let *F* be the field in a finite dimensional vector space *V*. If $\alpha \in V$ and L_α on V' defined by $L_\alpha(f) = f(\alpha) \forall f \in V'$ is a linear functional on V' i.e., $L_\alpha \in V''$. Also the mapping $\alpha \to L_\alpha$ is an isomorphism of *V* onto V''.

Proof: Let $\alpha \in V$ and $f \in V'$, then $f(\alpha)$ is a unique element of F. Then L_α defined by

$$L_\alpha(f) = f(\alpha) \forall f \in V' \quad \ldots\ldots (1)$$

is a mapping from V' into F.

Let $a, b \in F$ and $f, g \in V'$. Then

$$L_\alpha(af + bg) = (af + bg)(\alpha) = (af)(\alpha) + (bg)(\alpha)$$

$$= (af)(\alpha) + (bg)(\alpha) \quad \text{[By (1)]}$$

$$= af(\alpha) + bg(\alpha) \quad \text{[By the property of scalar multiplication in linear functional]}$$

$= a[L_\alpha(f)] + bL_\alpha(g)$ [By (1)]

Thus L_α is a linear functional on V' and thus $L_\alpha \in V''$.

Now we assume that ϕ be the function from V into V'' defined by,

$\phi(\alpha) = L_\alpha \; \forall \alpha \in V$.

First we will prove that ϕ is one-one: If $\alpha, \beta \in V$, then

$\phi(\alpha) = \phi(\beta)$

$\Rightarrow L_\alpha = L_\beta \Rightarrow L_\alpha(f) = L_\beta(f) \forall f \in V'$

$\Rightarrow f(\alpha) = f(\beta) \forall f \in V'$ [From (1)]

$\Rightarrow f(\alpha) - f(\beta) = 0 \forall f \in V' \Rightarrow f(\alpha - \beta) = 0 \forall f \in V'$

$\Rightarrow \alpha - \beta = 0$

[By the theorem that if $\alpha - \beta \neq 0$, then there exist linear functional f on V such that $f(\alpha - \beta) \neq 0$. Here we have $f(\alpha - \beta) = 0 \forall f \in V'$ so $\alpha - \beta$ must be 0]

$\Rightarrow \alpha = \beta$

$\Rightarrow \phi$ is one-one.

Now we will prove that ϕ is a linear transformation:

Let $\alpha, \beta \in V \; and \; a, b \in F$. Then,

$\phi(a\alpha + b\beta) = L_{a\alpha + b\beta}$ [By definition of ϕ]

For each $f \in V'$, we have

$L_{a\alpha + b\beta}(f) = f(a\alpha + b\beta)$

$= af(\alpha) + bf(\beta)$ [From (1)]

$= aL_\alpha(f) + bL_\beta(f)$ [From (1)]

$= (aL_\alpha)(f) + (bL_\beta)(f) = (aL_\alpha + bL_\beta)(f)$

$\therefore L_{a\alpha+b\beta} = aL_\alpha + bL_\beta = a\phi(\alpha) + b\phi(\beta)$

So, $\phi(a\alpha + b\beta) = a\phi(\alpha) + b\phi(\beta)$.

Hence, ϕ is linear transformation from V into $V^{''}$. Since ϕ is on-to also.

Hence ϕ is an isomorphism from V into $V^{''}$.

Note: The above theorem is called the natural correspondence between V and V''. It is significant to remember that the aforementioned theorem not only demonstrates that V and V'' are isomorphic—this much is evident from the fact that they have the same dimension—but also that an isomorphism is the natural correspondence between them. We refer to this characteristic of vector space as reflexivity. We have therefore demonstrated that every finite-dimensional vector space is reflexive in the aforementioned theorem.

7.7 *ANNIHILATORS*

Definition: Let V be a real vector space and S be a non-empty subset of a vector space V, then the set $\{f \in V^* \mid f(x) = 0, \forall\, x \in S\}$ is called an annihilators of a set S and it is denoted by $S^0 = \{f \in V^* \mid f(x) = 0 \;\; \forall x \in S\}$

Theorem 7 *Let S be a non-empty subset of a vector space V, then prove that S^0 is a subspace of V^*.*

Proof: *Here S is a non-empty subset of a vector space V.*
let V be a real vector space and V^* be a dual space of a vector space V.

$$S^0 = \{f \in V^* / f(x) = 0, \; \forall\, x \in S\}$$

$$\overline{0}(x) = 0, \;\; \forall\, x \in S$$

$$\overline{0} \in S^0$$

$$S^0 \neq \phi$$

(i) Let $f_1, f_2 \in S^0$, we have to prove that $f_1 + f_2 \in S^0$
Here $f_1, f_2 \in S^0$
So $f_1(x) = 0, \quad f_2(x) = 0, \quad \forall\, x \in S$

$$\begin{aligned}(f_1 + f_2)(x) &= f_1(x) + f_2(x)\\ &= 0 + 0\\ &= 0\\ f_1 + f_2 &\in S^0\end{aligned}$$

(ii) Let α be a scalar and let $f \in S^0$ then to prove that $\alpha f \in S^0$.
Here $f \in S^0$ so $f(x) = 0, \quad \forall\, x \in S$

$$\begin{aligned}((\alpha f)(x) &= \alpha f(x)\\ &= \alpha 0\\ &= 0\\ \alpha f &\in S^0\end{aligned}$$

So from (i) and (ii) S^0 is a subspace of dual space of V^*.

Note: If $S = \overline{0}$ then $S^0 = V^*$.

Check your progress

Problem 1: Check the dual basis of the basis set $B = \{(1,-1,3),(0,1,-1),(0,3,-2)\}$ for the vector space $V_3(R)$.

Problem 2: If the vectors $\alpha_1 = (1,1,1), \alpha_2 = (1,1,-1), \alpha_3 = (1,-1,-1)\}$ form the basis of $V_3(C)$. If $\{f_1, f_2, f_3\}$ is the dual basis and if $\alpha = (0,1,0)$, then find the value of $f_1(\alpha), f_2(\alpha), f_3(\alpha)$.

Problem 3: Check the dual basis of the basis set

$B = \{(1,0,0),(0,1,0),(0,0,1)\}$ for $V_3(R)$

7.8 *SUMMARY*

In this unit, we have learned about the important concept of linear transformation, linear functional, some special types and properties of linear functional, dual spaces, dual basis, reflexivity and annihilator. After completion of this unit learners will be able to:

- Find out the basic differences between the linear transformation and dual function.
- Conceptualized some special linear functions and their properties.
- Find out dual basis of any vector space corresponding to any given basis.
- Implement the concept existence theorem of dual basis.

7.9 *GLOSSARY*

- Linear Transformation
- Linear functional
- Dual basis
- Annihilator
- Reflexivity

7.10 *REFERENCES*

- Gel′fand I.M. Lectures on linear algebra (1989), Courier Corporation.
- Kenneth Hoffman & Ray Kunze, Linear Algebra (2nd edition)(2015), Prentice-Hall.
- David C. Lay, Linear Algebra and its Application (3rd Edition) (2007) Pearson Education Asia, India Reprint.
- Seymour Lipshutz and Marc Lipson, Schaum's outlines "Linear Algebra" (3rd Edition)(2012), Mc Graw Hill Education.
- J. N. Sharma and A. R. Vasistha, Linear Algebra (29th Edition) (1999), Krishna Prakashan.

7.11 *SUGGESTED READING*

- Minking Eie & Shou-Te Chang (2020), **A First Course In Linear Algebra, World Scientific.**
- Axler, Sheldon (2015), Linear algebra done right. Springer.
- https://nptel.ac.in/courses/111106051
- https://archive.nptel.ac.in/courses/111/104/111104137
- https://epgp.inflibnet.ac.in/

- https://old.sggu.ac.in/wp-content/uploads/2020/07/Linear-AlgebraTY_506.pdf

7.12 *TERMINAL QUESTION*

Long Answer Type Question:

1. If *f* is a linear functional on an *n*-dimensional vector space *V*(*F*), then show that the set of that subspace.

2. If *V* is a vector space over the field *F* and let *f* be a non-zero linear functional on *V* and let *N* be the null space of *f*. For a fix α_0 in *V* which is not in *N*. Prove that for each α in *V* there is a scalar *c* and a vector β in *N* such that $\alpha = c\alpha_0 + \beta$. Prove that *c* and β are unique.

3. If *V* is a vector space over the field *F*. Let f_1 and f_2 be linear functional on *V*. The function $f_1 + f_2$ defined by

$(f_1 + f_2)(\alpha) = f_1(\alpha) + f_2(\alpha)\ \forall \alpha \in V$ is a linear functional on *V*. If *c* is any element of *F*, the function *cf* defined by

$$(cf)(\alpha) = cf(\alpha) \forall \alpha \in V$$

is a linear functional on V. the set V'. The set V' of all linear functional on *V*, together with the addition and scalar multiplication defined as above is a vector space over the field *F*.

4. State and prove the existence theorem of dual basis.

5. Let *V* be the n-dimensional vector space over the field F and let $B = \{\alpha_1, \alpha_2, ..., \alpha_n\}$ be a basis for V. Then there exist a uniquely determined basis $B' = \{f_1, f_2, ..., f_n\}$ for V' such that $f_i(\alpha_j) = \delta_{ij}$. Consequently, the dual space of *n*-dimensional space is *n*-dimensional.

6. Find the dual basis of the basis set $B = \{(1,-2,3),(1,-1,1),(2,-4,7)\}$ of $V_3(R)$.

Short answer type question:

1. Prove that sum of two linear transformations is again a linear transformation.

2. Prove that composition of two linear transformations is again a linear transformation.

3. Let $T: R^2 \to R^2$ s.t., $T(x, y) = (x - y, x + y)$ and $S: R^2 \to R^2$ s.t., $S(x, y) = (x + y, x - y)$ then show that the compositions $S \circ T$ and $S \circ T$ are equal.

4. Define the linear functional and some special types of linear functional.

5. Show that trace function is a linear functional on the space of all $n \times n$ matrices over a field *F*.

6. If *f* is a non-zero linear functional on a vector space *V* and if *x* is an arbitrary scalar, does there necessarily exist a vector α in *V*, such that $f(\alpha)=x$?

7. Let *V* be an *n*-dimensional vector space over the field *F*. If α is a non-zero vector in *V*, there exist a linear functional *f* on *V* such that $f(\alpha)\neq 0$.

8. If the vectors $\alpha_1=(1,1,1),\alpha_2=(1,1,-1),\alpha_3=(1,-1,-1)\}$ form the basis of $V_3(C)$. If $\{f_1,f_2,f_3\}$ is the dual basis and if $\alpha=(0,1,0)$, then find the value of $f_1(\alpha),f_2(\alpha),f_3(\alpha)$.

9. Check the dual basis for the set $B=\{(1,0,0),(0,1,0),(0,0,1)\}$ for $V_3(R)$.

Fill in the blanks:

1. Sum of two linear transformations is also ……………

2. Composition of two linear transformations is also ……………………

3. Negative of the linear functional is also a ……………………

7.13 *ANSWERS*

Answers of check your progress:

1: $f_1(x,y,z)=x,\ f_2(x,y,z)=7x-2y-3z,\ f_3(x,y,z)=-2x+y+z$, i.e., the set $\{f_1,f_2,f_3\}$ is the dual basis of *B*.

2: $f_1(\alpha)=0,\ f_2(\alpha)=\dfrac{1}{2},\ f_3(\alpha)=-\dfrac{1}{2}$

3: $f_1(x,y,z)=-3x-5y-2z,\ f_2(x,y,z)=2x+y,\ f_3(x,y,z)=x+2y+z$ i.e., the set $\{f_1,f_2,f_3\}$ is dual basis.

Answer of long answer type questions:

1: The dimension of that subspace is *n*-1.

2: $B^{'}=\{f_1,f_2,f_3\}$

$f_1(x,y,z)=-3x-5y-2c,\ f_2(x,y,z)=2x+y,\ f_3(x,y,z)=x+2y+z$

Answer of short answer type questions:

8: $f_1(\alpha)=0,\ f_2(\alpha)=\dfrac{1}{2},\ f_3(\alpha)=-\dfrac{1}{2}$

9: The set $\{f_1,f_2,f_3\}$ is dual basis

$$f_1(x,y,z) = -3x - 5y - 2z,\ f_2(x,y,z) = 2x + y,\ f_3(x,y,z) = x + 2y + z\,.$$

Answer of fill in the blanks questions:

1: Linear transformation

2: Linear transformation

3: Linear functional

BLOCK- III

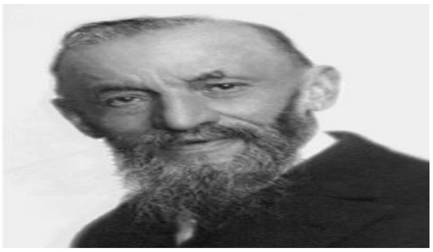
LINEAR OPERATOR, EIGEN VALUES AND EIGEN VECTORS

UNIT-8: LINEAR OPERATOR

CONTENTS

8.1 *INTRODUCTION*

The idea of a linear operator, which is essential to linear algebra together with the idea of a vector space, is used in many different areas of mathematics and science, most notably analysis and its applications. ***G.***



***Peano** provided the first definition of a linear operator as we know it today.*

Giuseppe Peano
27 August 1858-20 April 1932 (aged 73)
https://en.wikipedia.org/wiki/Giuseppe_Peano

We covered bases and how to alter a vector's basis representation in the previous units. The representation (the n-tuple of components) was multiplied by the appropriate matrix, which represented the relationship between the two sets of basis vectors, in order to achieve this change of basis. It's crucial to keep in mind that the vector itself does not change while thinking about the change of basis operation only the coordinate system in which it is written does.

Linear operators will be discussed in this unit. Although they will also be described in terms of a matrix multiplication, linear operators are functions on the vector space that differ fundamentally from changes of basis. A linear operator, often known as a linear transformation, is a method that converts a given vector into a completely other vector. As we will see, linear operators have the ability to change a vector in one space into a different vector in the same space, implicitly conduct a change of basis, or simply transform a vector in one space into another.

8.2 *OBJECTIVES*

After reading this unit learners will be able to

- Understand the basic difference between linear functional, linear transformation and linear operator.
- Visualized the matrix representation of linear operator.
- Implementation of linear operator in different basis.
- Visualized and understand the important of matrix operator and change of basis.

8.3 *LINEAR OPERATOR*

Linear operator is the special case of a linear transformation. Sometimes, *linear transformation* is also called *linear operator*. All that a linear operator does is assign a vector, which may or may not be in the same linear vector space, to another vector. Additionally, it needs to meet the linearity requirements. To be precise, we provide the definition that follows:

Definition 1: Let $V(F)$ be a vector space. A linear operator on V is a function T from V into V such that

$T(a\alpha + b\beta) = aT(\alpha) + bT(\beta)$, for all $\alpha, \beta \in V$ *and* $a, b \in F$

Thus T is a linear operator on V if T is linear transformation from V into V itself.

Definition 2: Linear Operator (Transformation): An operator A from linear vector space X to linear vector space Y, denoted $A{:}X \to Y$, is linear if,

$A(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 A x_1 + \alpha_2 A x_2$

for any x_1 x_2 , $\in X$, and scalars α_1 and α_2.

8.4 *RANGE AND NULL SPACES*

A linear operator, denoted by $A{:}X \to Y$ in the definition above, can also be represented as $Ax = y$, and it generates the vector $y \in Y$. This operation yields a vector y that is considered to be within operator A's range; it is also sometimes referred to as the image of x in Y, where x is the preimage of y. Although linear operators generally have multidimensional ranges, the idea of an operator's range is comparable to that of scalar functions. These ranges will be vector spaces that are linear in nature.

Range space: The range space of an operator $A: X \to Y$, denoted by $R(A)$, is the set of all vectors $y_i \in Y$ such that for every $y_i \in R(A)$ there exist an $x \in X$ such that $Ax = y$.

It is claimed that operator A is onto, or surjective, if its range contains all of space Y. A is said to be one-to-one, or injective, if it maps elements in X to unique values in Y, that is, if $x_1 \neq x_2$ implies that $A(x_1) \neq A(x_2)$. Operator A is also invertible if it is bijective, which occurs when it is both one-to-one and onto. When an operator A is invertible, it means that there is another operator, $A^{-1}: Y \to X$, such that $A^{-1}(A(x)) = x$ and $A^{-1}(A(y)) = y$. Then, $A^{-1}A$ is represented as I_X and AA^{-1} as I_Y, which are the identity operators in the corresponding spaces. $A: X \to X$ is an example of how A maps a space into itself. In this case, we just write $A^{-1}A = AA^{-1} = I$.

It is frequently necessary to determine which vector, out of all the ones in $x \in X$, will map to the zero vector in Y. Because of this, we define the identity operators in their respective spaces as the null space of an operator.

Null space: The null space of operator A, denoted by $N(A)$, is the set of all vectors $x_i \in X$ such that $A(x_i) = 0$:

$N(A) = \{x_i \in X \mid Ax_i = 0\}$

8.5 *MATRIX REPRESENTATION OF LINEAR OPERATORS*

To be more inclusive, we will talk about the matrix representation of linear operators that have the potential to modify the vector's basis as well. This is evidently required when discussing an operator that manipulates vectors in a one-dimensional space and yields vectors in a different dimension. It is obvious that both vectors could not be adequately described by a single basis.

Let $x \in X^n$ and $y \in X^m$ be two vectors, one from an *n*-dimensional space and the other from an *m*-dimensional space. We shall obtain a matrix representation *A* for an operator $A : X^n \to X^m$, that converts a vector in one space into a vector in another. Indicate the two spaces' arbitrary bases as

$\{v_j\} = \{v_1, v_2, ..., v_n\}$ for space X^n

And

$\{u_i\} = \{u_1, u_2, ..., u_n\}$ for space X^m

By extending *x* as a representation in its basis as $x = \sum_{j=1}^{n} \alpha_j v_j$ and applying the linearity property of the operator *A*,

$$y = Ax = A\left(\sum_{j=1}^{n} \alpha_j v_j\right) = \sum_{j=1}^{n} \alpha_j A(v_j) \qquad \text{.....................} \ (1)$$

This straightforward but significant result means that, given the basis vectors in which *x* is expressed, we can determine the influence of *A* on any vector *x*. Another way to express equation (1) in vector-matrix notation is as follows:

$$y = Ax = \left[Av_1 \vdots Av_2 \vdots \cdots \vdots Av_n\right] \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}$$

It is now evident that every vector Av_j exists in the range space X^m by definition. Then, it can be extended in the basis defined for that space, $\{u_i\}$, just like any other vector in that space. This extension provides

$$Av_j = \sum_{i=1}^{m} a_{ij} u_i \qquad \ldots\ldots\ldots\ldots\ldots.. (2)$$

for a certain set of a_{ij} coefficients. Using coefficients β_i, y can also be directly expanded in terms of its own basis vectors.

$$y == \sum_{i=1}^{m} \beta_i u_i \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (3)$$

Substitute equation (3) and (2) in (1), we can write

$$y = \sum_{j=1}^{n} \alpha_j \left(\sum_{i=1}^{m} a_{ij} \alpha_j \right) u_i = \sum_{j=1}^{m} \beta_i u_i \quad \ldots\ldots\ldots\ldots (4)$$

By changing the order of summation in this expression, we obtain

$$\sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} \alpha_j \right) u_i = \sum_{j=1}^{m} \beta_i u_i \quad \ldots\ldots\ldots\ldots\ldots. (5)$$

But because of the uniqueness of any vector's expansion in a basis, the expansion of y in $\{u_i\}$ must be unique. This implies that

$$\beta_i = \sum_{j=1}^{n} \alpha_{ij} \alpha_j \quad \forall\ i = 1\ to\ m \quad \ldots\ldots\ldots\ldots\ldots\ldots (6)$$

The expression (6) above can be seen by the learners to be a matrix-vector multiplication. Actually, this is the way we'll often use the outcome. If $\alpha = [\alpha_1 \ \ \alpha_2 \cdots \alpha_n]^T$ is the representation of $x \in X^n$ in the $\{v_j\}$ basis and $\beta = [\beta_1 \ \ \beta_2 \cdots \beta_m]^T$ is the representation of $y \in X^m$ in the $\{u_i\}$ basis, then, using our new matrix representation for operator A, we can determine this representation of y using the matrix multiplication $\beta = A\alpha$.

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \qquad \text{………… (7)}$$

That is, the coefficient from (2) is the $(i, j)^{th}$ element of the $(m \times n)$ matrix *A*, which explains how basis vectors from one space are transformed into the basis of the range space via the operator *A*. This implies, of course, that before defining the matrix representation for any linear operator, the bases of each space must be given. A comparison between (2) with (7) generally leads to the following conclusion:

The representation of the vector that results from A operating on the j^{th} basis vector of X^n, where this new vector is expanded into the basis of the range space, can be formed as the j^{th} column of the matrix representation of any linear operator $A : X^n \to X^m$. To put it another way, the j^{th} column of A is just Av_j, expressed as a representation in the basis $\{u_i\}$.

As we will see in the following examples, this is a very helpful characteristic. It gives us a practical means of ascertaining the matrix representation of any given linear operator.

The range and null spaces computation is also affected by the linear operator's matrix representation. As we will demonstrate in this unit (definition of null space), the linear operation $y = Ax$ can be viewed numerically as,

$$\begin{aligned} y &= Ax \\ &= \begin{bmatrix} a_1 & a_2 \ldots a_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \end{aligned} \qquad \text{……….. (8)}$$

Where a_i denotes the i^{th} column of the matrix *A* and x_i is the i^{th} component of the vector *x* or of its representation. Consequently, the span of all of *A*'s columns can be used to represent the range space if it represents the space of all possible values of *Ax*. This suggests that the rank of matrix *A* is equal to $r(A) = \dim(R(A))$, which is the dimension of the range space of operator *A*. Similarly, the space containing all solutions to the simultaneous linear equations can be used to represent the null space of operator *A*.

$0 = Ax$

$$= \begin{bmatrix} a_1 & a_2 \ldots a_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \qquad \ldots\ldots\ldots\ldots\ldots (9)$$

As a result, $q(A) = \dim(N(A))$, which is the dimension of operator A's null space, equals the nullity of matrix A.

Example 1: (Related to rotation matrices)

Consider a linear operator $A: R^2 \to R^2$ function, as seen in Figure 1, takes a vector *x* and rotates it by an angle θ counterclockwise.2. The linear operator that carries out these planar rotations on any vector in R^2 can be represented by a matrix. Test this matrix by rotating the vector $x = [1\ 2]^T$ by an angle $\theta = 30^\circ$.

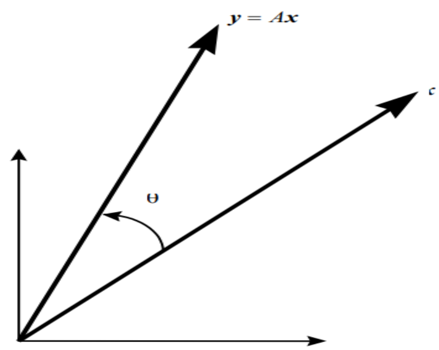


Figure 1: Vector *x* is transformed by Operator *A* converts the vector *x* into a vector *y* = *Ax*, which is just *x* rotated (by angle θ counterclockwise).

Solution: Each space's basis and the operator's impact on those basis vectors are necessary to determine the matrix representation of any operator. The standard basis vectors displayed in Figure 1 will be employed.

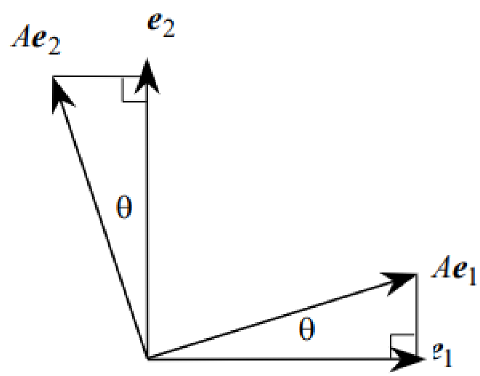


Figure 2: The effect of operator *A* on the basis vectors.

Basic trigonometry allows us to determine by decomposing the rotated vectors Ae_1 and Ae_2 along their original basis directions, e_1 and e_2.

$$Ae_1 = \cos\theta \cdot e_1 + \sin\theta \cdot e_2 = [e_1 \mid e_2]\begin{bmatrix}\cos\theta \\ \sin\theta\end{bmatrix}$$
$$Ae_2 = -\sin\theta \cdot e_1 + \cos\theta \cdot e_2 = [e_1 \mid e_2]\begin{bmatrix}-\sin\theta \\ \cos\theta\end{bmatrix} \quad \ldots\ldots\ldots\ldots\ldots (1)$$

Therefore, the matrix representation of A is

$$A = \begin{bmatrix}\cos\theta & -\sin\theta \\ \sin\theta & \cos\theta\end{bmatrix}$$

Keep in mind that every column in this A-matrix is only the matching representation from the previous (1).

Applying this to the vector $x = [1 \quad 2]^T$, we get

$$Ax = \begin{bmatrix}\cos 30° & -\sin 30° \\ \sin 30° & \cos 30°\end{bmatrix}\begin{bmatrix}1 \\ 2\end{bmatrix} = \begin{bmatrix}-0.134 \\ 2.23\end{bmatrix}$$

Figure 3 provides verification of this rotation.

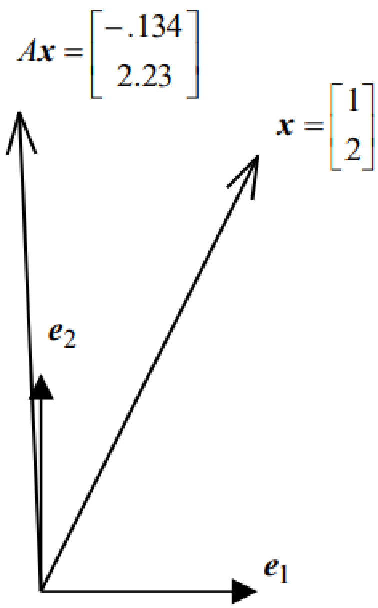


Figure 3: Pictorial representation of vector rotation.

Example 2: Composite Rotations: Roll, pitch and Yaw

We looked at plane rotations above. Three-dimensional rotations are another option. It's important to keep in mind that rotations around three different axes are feasible in three dimensions. Depending on which axis was chosen to represent the rotational axis, different matrices are employed to describe these rotations. Figure 4 shows the three possible rotations around e_1, e_2, or e_3.

These rotations about coordinate axes can be applied sequentially to create an arbitrary three-dimensional rotation, although such sequences wouldn't be special. There are typically an endless number of distinct axis rotation steps that can be taken to get from one three-dimensional orthonormal frame to another if the orientation of one is applied arbitrarily to the other.

To create the composite rotation matrix that rotates a coordinate system by θ_R around e_1, θ_p around e_2, and θ_Y around e_3, in that sequence, find the change of basis matrices for the three component rotations.

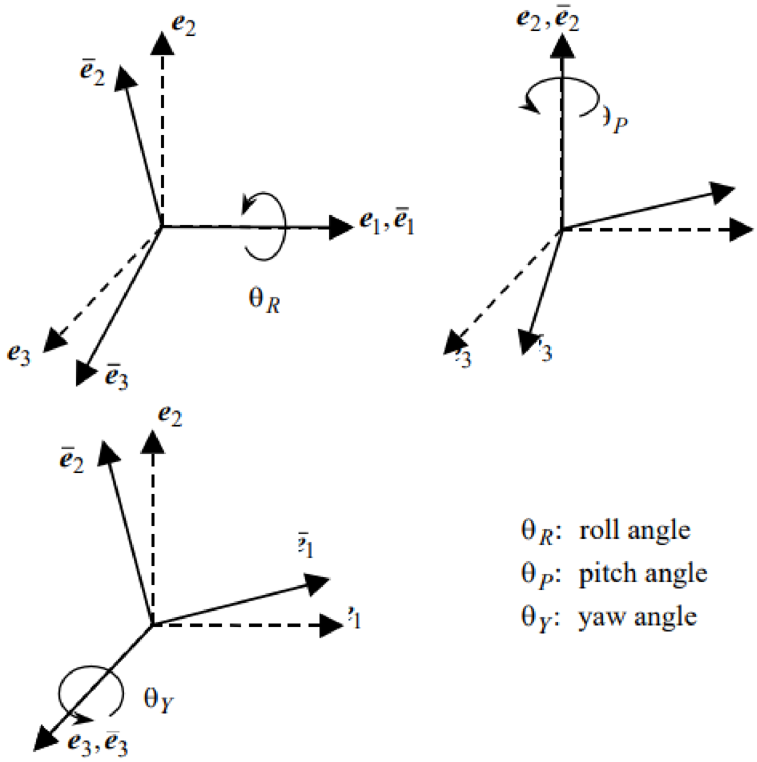


Figure 4: Axis rotations for pitch, roll, and yaw

Solution: As with the planar rotation example in Figure 2, the individual matrices can be determined by looking at the rotations in the plane in which they occur and decomposing the rotated axes in that plane. Equation may be used to get the change of basis matrix for each rotation using those same planar decompositions. This change of basis matrix can also be thought of as a rotation operator on a vector.

$$R_R = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_R & -\sin\theta_R \\ 0 & \sin\theta_R & \cos\theta_R \end{bmatrix} \qquad R_P = \begin{bmatrix} \cos\theta_P & 0 & \sin\theta_P \\ 0 & 1 & 0 \\ -\sin\theta_P & 0 & \cos\theta_P \end{bmatrix}$$

and

$$R_Y = \begin{bmatrix} \cos\theta_Y & -\sin\theta & 0 \\ \sin\theta_Y & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

The composite rotation can be expressed as the product of the roll, pitch, and yaw rotations that are applied to a vector *x* in that order.

$$R = R_Y R_P R_R$$
$$= \begin{bmatrix} \cos\theta_Y \cos\theta_P & \cos\theta_Y \sin\theta_P \sin\theta_R - \sin\theta_Y \cos\theta_R & \cdots \\ \sin\theta_Y \cos\theta_P & \sin\theta_Y \sin\theta_P \sin\theta_R + \cos\theta_Y \cos\theta_R & \cdots \\ -\sin\theta_P & \cos\theta_P \sin\theta_R & \end{bmatrix}$$

$$\begin{matrix} \cos\theta_Y \sin\theta_P \cos\theta_R + \sin\theta_Y \sin\theta_R \\ \sin\theta_Y \sin\theta_P \cos\theta_R - \cos\theta_Y \sin\theta_R \\ \cos\theta_P \cos\theta_R \end{matrix} \Bigg] \qquad \ldots\ldots\ldots. (1)$$

For free-floating objects whose orientation is best defined with regard to inertial coordinate frames, such as aircraft, missiles, spacecraft, and submarines, such rotations are helpful. There are some differences in rotations about body-centered coordinate axes.

8.6 *LINEAR OPERATORS IN DIFFERENT BASES*

We shall pay special attention to transformations from a space into itself, such $A: R^n \to R^n$ in the upcoming chapter. This will, of course, result in a square ($n \times n$) matrix and may also perform a basis change at the same time. Denote such an n-dimensional space with basis $\{v\}$ by writing X_v. The matrix *A* that modifies $X_v \to X_v$ is referred to as a transformation in the basis $\{v\}$.

For the time being, we will no longer be referring to vectors in a basis, but since a vector's representation will always change in tandem with its basis, it is reasonable to assume that an operator's representation on a vector will also alter in tandem with the basis. Hence, $\hat{A}$ represents the operator that converts vectors stated in a different basis, $\{\hat{v}\}$. By using this notation, we may state that $y = Ax$ and $\hat{y} = \hat{A}\hat{x}$, where $\hat{y}$ and $\hat{x}$ are obviously vectors represented in basis $\{\hat{v}\}$. The current topic of discussion is how to convert an operator's matrix representation from one basis to another. The change of basis matrices created in the preceding unit will be used for this. Using the change of basis matrix *B*, we can write $\hat{y} = By$ and $\hat{x} = Bx$. Using these changes of basis we have.

$$\hat{y} = \hat{A}\hat{x}$$
$$By = \hat{A}Bx$$

so

$$y = B^{-1}\hat{A}Bx$$

B^{-1} must exist because it is an $(n \times n)$ change of basis matrix and of full rank. Comparing this to the expression $y = Ax$, we see that

$$A = B^{-1}\hat{A}B$$

Equivalently, if we denote $M = B^{-1}$, then $A = M\hat{A}M^{-1}$, or

$$\hat{A} = M^{-1}AM$$

A similarity transformation occurs when the basis of a matrix representation of a linear operator is changed. There is a claim that A is similar to $\hat{A}$ and vice versa.

8.7 *MATRIX OPERATORS AND CHANGE OF BASIS*

Examine the linear vector space X comprising polynomials in s with real coefficients and a degree less than four across over the field of reals. The operator $A: X \rightarrow X$ that takes such vectors $v(s)$

and returns the new vectors $w(s) = v''(s) + 2v'(s) + 3v(s)$ (where prime denotes differentiation in s) is a linear operator (as is differentiation in general). The matrix form of this operator in two bases will be found in this example,

$$\{e_i\} = \{s^3, s^2, s, 1\} \equiv \{e_1(s), e_2(s), e_3(s), e_4(s)\}$$

and

$$\{\bar{e}_i\} = \{s^3 - s^2, s^2 - s, s - 1, 1\} \equiv \{\bar{e}_1(s), \bar{e}_2(s), \bar{e}_3(s), \bar{e}_4(s)\}$$

as well as the transformation between the two.

Solution:

First we will determine the effect of A on the basis vectors $\{e_i\}$:

$$\begin{aligned} A\boldsymbol{e}_1 &= \boldsymbol{e}_1''(s) + 2\boldsymbol{e}_1'(s) + 3\boldsymbol{e}_1(s) \\ &= 6s + 6s^2 + 3s^3 \\ &= \begin{bmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 & \boldsymbol{e}_3 & \boldsymbol{e}_4 \end{bmatrix} \begin{bmatrix} 3 \\ 6 \\ 6 \\ 0 \end{bmatrix} \end{aligned}$$

$$\begin{aligned} A\boldsymbol{e}_2 &= \boldsymbol{e}_2''(s) + 2\boldsymbol{e}_2'(s) + 3\boldsymbol{e}_2(s) \\ &= 2 + 4s + 3s^2 \\ &= \begin{bmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 & \boldsymbol{e}_3 & \boldsymbol{e}_4 \end{bmatrix} \begin{bmatrix} 0 \\ 3 \\ 4 \\ 2 \end{bmatrix} \end{aligned}$$

$$\begin{aligned} A\boldsymbol{e}_3 &= \boldsymbol{e}_3''(s) + 2\boldsymbol{e}_3'(s) + 3\boldsymbol{e}_3(s) \\ &= 2 + 3s \\ &= \begin{bmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 & \boldsymbol{e}_3 & \boldsymbol{e}_4 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 3 \\ 2 \end{bmatrix} \end{aligned}$$

$$\begin{aligned} A\boldsymbol{e}_4 &= \boldsymbol{e}_4''(s) + 2\boldsymbol{e}_4'(s) + 3\boldsymbol{e}_4(s) \\ &= 3 \\ &= \begin{bmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 & \boldsymbol{e}_3 & \boldsymbol{e}_4 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 3 \end{bmatrix} \end{aligned}$$

We obtain the matrix by taking the four columns of coefficients in these four forms.

$$A = \begin{bmatrix} 3 & 0 & 0 & 0 \\ 6 & 3 & 0 & 0 \\ 6 & 4 & 3 & 0 \\ 0 & 2 & 2 & 3 \end{bmatrix} \qquad \ldots\ldots. (1)$$

We can check this result by applying this transformation to the vector $v(s) = s^2 + 1$, whose representation in $\{e_i\}$ is $v = [0\ 1\ 0\ 1]^T$. Multiplying this representation of w by matrix A

results in

$$Av = \begin{bmatrix} 3 & 0 & 0 & 0 \\ 6 & 3 & 0 & 0 \\ 6 & 4 & 3 & 0 \\ 0 & 2 & 2 & 3 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \\ 4 \\ 5 \end{bmatrix} \qquad \ldots\ldots.. (2)$$

Applying the transformation by direct differentiation gives

$$w(s) = Av(s) = v''(s) + 2v'(s) + 3v(s) = 3s^2 + 4s + 5 \qquad \ldots\ldots\ldots (3)$$

which is clearly the same vector as computed in (2) above.

In what way does this operator manifest itself in the distinct base $\{\bar{e}_i\}$? There are two techniques to ascertain the response. Either we can apply a similarity transformation on the matrix A that was previously computed in (1), or we can derive the matrix transformation directly in this basis.

Initially, after determining the basis matrix change between the two bases, we need to calculate the expansion.

$$\boldsymbol{e}_j = \sum_{i=1}^{n} b_{ij} \bar{\boldsymbol{e}}_i$$

Upon closer examination, it becomes evident that calculating the inverse relationship is less complicated:

$$\bar{\boldsymbol{e}}_1 = \boldsymbol{e}_1 - \boldsymbol{e}_2 = [\boldsymbol{e}_1 \quad \boldsymbol{e}_2 \quad \boldsymbol{e}_3 \quad \boldsymbol{e}_4]\begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix}$$

$$\bar{\boldsymbol{e}}_2 = \boldsymbol{e}_2 - \boldsymbol{e}_3 = [\boldsymbol{e}_1 \quad \boldsymbol{e}_2 \quad \boldsymbol{e}_3 \quad \boldsymbol{e}_4]\begin{bmatrix} 0 \\ 1 \\ -1 \\ 0 \end{bmatrix}$$

$$\bar{\boldsymbol{e}}_3 = \boldsymbol{e}_3 - \boldsymbol{e}_4 = [\boldsymbol{e}_1 \quad \boldsymbol{e}_2 \quad \boldsymbol{e}_3 \quad \boldsymbol{e}_4]\begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \end{bmatrix}$$

$$\bar{\boldsymbol{e}}_4 = \boldsymbol{e}_4 = [\boldsymbol{e}_1 \quad \boldsymbol{e}_2 \quad \boldsymbol{e}_3 \quad \boldsymbol{e}_4]\begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

We get at the inverse matrix since this is the inverse relationship, $M = B^{-1}$. By gathering the coefficient columns, this matrix is

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

From which we can compute.

$$B = M^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

Now using the formula $A = B^{-1}\overline{A}B$, we find that

$$\overline{A} = BAB^{-1} = \begin{bmatrix} 3 & 0 & 0 & 0 \\ 6 & 3 & 0 & 0 \\ 8 & 4 & 3 & 0 \\ 6 & 4 & 2 & 3 \end{bmatrix}$$

We may now find the effect of the original operator on vectors that have already been expressed in the $\{\overline{e_i}\}$ basis in order to verify this result by determining $\overline{A}$ in a different method. Finding the expression for $\bar{v}(s)$, or the same vector v but in the "bar" basis, should come first:

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

Which is simply the matrix-vector notation for

$$1.\bar{e}_2(s) + 1.\bar{e}_3(s) + 2.\bar{e}_4(s) = (s^2 - s) + (s - 1) + 2 = s^2 + 1 = v(s)$$

Now applying the matrix operation in the $\{\overline{e_i}\}$ basis,

$$\overline{A}\bar{\boldsymbol{v}} = \begin{bmatrix} 3 & 0 & 0 & 0 \\ 6 & 3 & 0 & 0 \\ 8 & 4 & 3 & 0 \\ 6 & 4 & 2 & 3 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \\ 7 \\ 12 \end{bmatrix}$$

Writing this representation out explicitly in terms of the "barred" basis vectors,

$$\bar{A}\bar{v}(s)(s) = 3\bar{e}_2 + 7\bar{e}_3 + 12.\bar{e}_4 = 3(s^2 - s) + 7(s-1) + 12(1) = 3s^2 + 4s + 5$$

It is, naturally, just what we would anticipate: the identical vector that we calculated in Equation (2). Stated otherwise, we have proven that the Figure 5 is commutative. "Commucative diagrams" illustrate how various paths, each signifying a sequential action or transformation, lead to the same outcome. One executes the action stated next to the arrow, in the indicated direction, to move from one vector to another. In case an arrow needs to be traversed "against" or in reverse, the inverse transformation is employed. This graphic can be used to show that $\bar{A} = BAB^{-1}$.

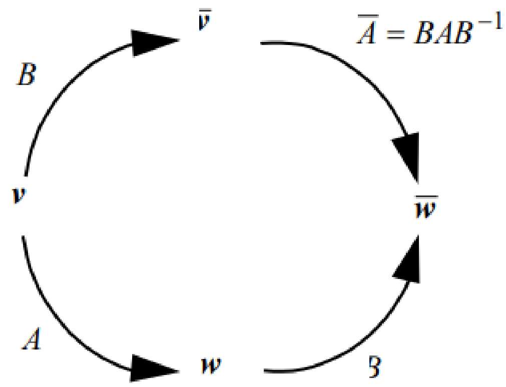


Figure 5: Commutative diagram showing different paths to the transformed vector $\bar{w}$.

Check your progress

Problem 1: Show that the transformation T: V_2(**R**) → V_2(**R**) defined by T(a, b) = (a + b, a) ∀ a, b ∈ **R** is a linear operator.

Solution: To show that T is a linear transformation, we need to prove that,

For any x, y ∈ V_2(R)

T(x + y) = T(x) + T(y) and T(ax) = aT(x) where a is a scalar in field.

Let (x_1, y_1) and (x_2, y_2) are arbitrary elements of V_2(**R**)

$T[(x_1, y_1) + (x_2, y_2)] = T[(x_1 + x_2, y_1 + y_2)] = (x_1 + x_2 + y_1 + y_2, x_1 + x_2)$ …..(i)

$T(x_1, y_1) + T(x_2, y_2) = (x_1 + y_1, x_1) + (x_2 + y_2, x_2) = (x_1 + x_2 + y_1 + y_2, x_1 + x_2)$ …..(ii)

From (i) and (ii), we get $T[(x_1, y_1) + (x_2, y_2)] = T(x_1, y_1) + T(x_2, y_2)$

Now, $T[a(x_1, y_1)] = T(ax_1, ay_2) = (ax_1 + ay_1, ax_1) = a(x_1 + y_1, x_1) = aT(x_1, y_1)$.

∴ T is a linear transformation.

Problem 2: Given a linear operator T on $V_3(\mathbf{R})$ defined by T (a, b, c) = (2b + c, a – 4b, 3a) corresponding to the basis B = {(1, 0, 0), (0, 1, 0), (0, 0, 1)}. Find the matrix representation of T.

Solution: Now, T(1, 0, 0) = (2 × 0 + 0, 1 – 4 × 0, 3 × 1) = (0, 1, 3)

= 0(1, 0, 0) + 1(0, 1, 0) + 3(0, 0, 1)

T(0, 1, 0) = (2 × 1 + 0, 0 – 4 × 1, 3 × 0) = (2, –4, 0)

= 2(1, 0, 0) –4(0, 1, 0) + 0(0, 0, 1)

And T(0, 0, 1) = (2 × 0 + 1, 0 – 4 × 0, 3 × 0) = (1, 0, 0)

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

Then, the matrix representation of T with respect to the basis B is

$$[T; B] = \begin{bmatrix} 0 & 2 & 1 \\ 1 & -4 & 0 \\ 3 & 0 & 0 \end{bmatrix}$$

8.8 *SUMMARY*

This unit has covered some of the foundational mathematics required to comprehend state spaces. In this case, the idea of a linear operator proved crucial. We introduce the linear operator technique as an alternative to thinking in terms of matrix-vector multiplication as it is used in matrix theory.

This provides a far deeper understanding of geometry for some of the fundamental math operations we have been carrying out all along. Certain ideas in control systems and linear system theory cannot be properly grasped without this approach. Other important concepts introduced in this unit were:

- It is possible to write any linear operator as a matrix, and any matrix can be conceptualized as an operator. Therefore, it should be evident that even though we emphasize the geometric comprehension of linear operators, our outdated computing techniques and habits are still valuable. Conceptualized some special linear functions and their properties.

- The linear operators are expressed in certain bases, which are either expressly stated or inferred from their context. Operations involving matrix multiplication can be used to alter these bases.

8.9 *GLOSSARY*

- Linear Operator
- Rotation matrices
- Composite rotation

8.10 *REFERENCES*

- Gel′fand I.M. Lectures on linear algebra (1989), Courier Corporation.
- Kenneth Hoffman & Ray Kunze, Linear Algebra (2nd edition)(2015), Prentice-Hall.
- David C. Lay, Linear Algebra and its Application (3rd Edition) (2007) Pearson Education Asia, India Reprint.

- Seymour Lipshutz and Marc Lipson, Schaum's outlines "Linear Algebra" (3rd Edition)(2012), Mc Graw Hill Education.

- J. N. Sharma and A. R. Vasistha, Linear Algebra (29th Edition) (1999), Krishna Prakashan.

8.11 *SUGGESTED READING*

- Minking Eie & Shou-Te Chang (2020), **A First Course In Linear Algebra, World Scientific.**

- Axler, Sheldon (2015), Linear algebra done right. Springer.

- https://nptel.ac.in/courses/111106051

- https://archive.nptel.ac.in/courses/111/104/111104137

- https://epgp.inflibnet.ac.in/

- https://old.sggu.ac.in/wp-content/uploads/2020/07/Linear-AlgebraTY_506.pdf
- https://orb.binghamton.edu

8.12 *TERMINAL QUESTION*

Long Answer Type Question:

1. Show that mapping T: $V_3(\mathbf{R}) \to V_3(\mathbf{R})$ defined by T(a, b, c) = (a, b, c) $\forall$ a, b, c $\in \mathbf{R}$ is a linear operator.

2. Show that the given subset of vectors of $\mathbf{R}^3$ forms a basis for $V_3(\mathbf{R})$.
{(1, 0, –1), (1, 2, 1), (0, –3, 2)}.

3. Given a linear operator T on $V_3(\mathbf{R})$ defined by T(a, b, c) = (2b + c, a – 4b, 3a) corresponding to the basis B = {(1, 1, 1), (1, 1, 0), (1, 0, 0)}. Find the matrix representation of T.

Short Answer Type Question:

1. Show that the mapping $T: F^3 \to F^3$ defined by,
$T(x, y, z) = (x - y + 2z, 2x + y, -x - 2y + 2z)$ is a linear operator

2. Which of the following functions $T: R^2 \to R^2$ are linear operator

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

b. $T(a,b) = (b,a)$

c. $T(a,b) = (a+b, a)$

3. Show that the mapping $T: V_2(R) \to V_2(R)$ defined by

$T(a,b) = (b,0)\, \forall\, a,b \in R$ is a linear operator.

4. Show that the mapping $T: V_3(R) \to V_3(R)$ defined by,
$T(x, y, z) = (3x, x - y, 2x + +y + z) \forall (x, y, z) \in V_3(R)$ is a linear operator.

5. Show that the mapping $T: R^3 \to R^3$ defined by $T(a,b,c) = (0, a, b)\, \forall\, a,b \in R$ is a linear operator.

8.13 *ANSWERS*

Answer of short answer type questions:

2 **(a)** T is a linear operator.
(b) T is a linear operator
(c) T is a linear operator.

UNIT-9: CHARACTERISTIC AND MINIMAL POLYNOMIAL OF AN OPERATOR

CONTENTS

9.1 *INTRODUCTION*

Generally, our focus on a polynomial reflects our interest either in its degree or coefficients from **F**. If we focus on the coefficient of highest degree of a polynomial, and try to make it a unit of the field **F**, then resultant polynomial is called monic polynomial. Such polynomials are relatively easy to factorize. One of such example is characteristic polynomial of linear operator or in particular, a matrix (square). All the roots of this equation are called **eigen value** or **characteristic value** or **latent root**. Now we try to find a non-zero vector v ∈ V such that T(V) = CV, where C is eigen value of T. We study **Cayley-Hamilton theorem,** which states that a linear operator satisfies its characteristic equation. Then we try to find a minimal polynomial which is also satisfied by T. This **inter-relationship** produces many standard results. Since a diagonal matrix is always easy to study so, we try to **diagonalize** every square matrix. But there are certain rules, which elaborate us the possibility and limitations of this concept.

Further we try to decompose a vector space into **independent subspaces**. Also, there are a variety of results on the **projections**.

9.2 *OBJECTIVE*

After studying this chapter, you will understand –

- Eigen value, eigen vector.
- Characteristic and minimal polynomial.
- Diagonalizable operators.
- Invariant and independent subspaces.
- Projection on a vector space.

9.3 *CHARACTERISTIC AND MINIMAL POLYNOMIAL*

Let T be a linear operator on a finite dimensional vector space V(**F**).

(1) A polynomial p(x) ∈ **F**[x] is called a **monic** polynomial, if the coefficient of the highest power of x in p(x) is unity. Thus $p(x) = x^n + a_1x^{n-1} + \ldots + a_{n-1}x^1 + a_n$ ∈ **F**[x] is a monic polynomial.

(2) The linear operator T satisfies the monic polynomial p(x) if,

$p(T) = T^n + a_1T^{n-1} + \ldots + a_{n-1}T^1 + a_n I = 0.$

If p(T) = 0, then we say that the polynomial p(x) **annihilates** T.

(3) Let T be a linear operator on a finite-dimensional vector space V(**F**). A monic polynomial p(x) ∈ F[x] of **lowest degree** such that p(T) = 0 is called a minimal polynomial for T over **F**.

(4) In a similar manner, we can define minimal polynomial for a square matrix A.

Theorem 1: Let T be a linear operator on a finite-dimensional vector space V(**F**). Then the characteristic and the minimal polynomials for T have the same roots, except for multiplicities.

Proof: Let p(x) be the minimal polynomial for T, so that p(T) = **0.** Suppose c be a root of p(x), so that p(c) = **0** .

Claim: We shall prove that c is a root of the characteristic polynomial of T (i.e. c is an eigen value of T). As c is a root of p(x).

$\Rightarrow$ (x – c) divides p(x) in **F**[x]. so by division algorithm, there exist q(x) $\in$ **F**[x] such that

$$p(x) = (x - c)\, q(x)\ ; \quad \deg q(x) < \deg p(x) \qquad \ldots.(1)$$

As p(x) is the minimal polynomial for T and deg q(x) < deg p(x). So q(T) $\neq$ **0**.

$\Rightarrow$ there exists some $\mathbf{0} \neq v \in V$, such that

$$q(T)(v) \neq \mathbf{0}.$$

Suppose $x = q(T)(v) \neq \mathbf{0}$. Then from equation (1), we have

$$p(T) = (T - cI)\, q(T)$$

$\Rightarrow$ $(T - cI)\, q(T) = \mathbf{0}$, as $p(T) = \mathbf{0}$.

$\Rightarrow$ $(T - cI)\, q(T)(v) = \mathbf{0}(v) = \mathbf{0}$.

$\Rightarrow$ $(T - cI)(x) = \mathbf{0}$.

$\Rightarrow$ $T(x) - cI(x) = \mathbf{0}$.

$\Rightarrow$ $T(x) = cx$.

$\Rightarrow$ c is an eigen value of T.

So c is a root of characteristic polynomial of T.

Step II: Let c be a root of the characteristic polynomial of T, i.e. c is an eigenvalue of T. Then there exists $\mathbf{0} \neq v \in V$, such that

$$T(v) = cv.$$

Since p(x) is a polynomial, p(T)(v) = p(c)(v)

$\Rightarrow$ $p(c)(v) = \mathbf{0}$ as $p(T) = \mathbf{0}$.

$\Rightarrow$ $p(c) = \mathbf{0}$ as $v \neq \mathbf{0}$.

Hence c is a root of the minimal polynomial for T.

Theorem 2: Let T be a diagonalizable linear operator on V and let $c_1, \ldots, c_k$ be the distinct eigenvalues of T. Then the minimal polynomial for T is the polynomial

$$p(x) = (x - c_1) \ldots (x - c_k).$$

Proof: Since we know that each eigen value of T is a root of the minimal polynomial for T. Hence each of $c_1, \ldots, c_k$ is a root of the minimal polynomial for T.

$\Rightarrow$ each of the polynomials $(x - c_1), \ldots, (x - c_k)$ is a factor of the minimal polynomial for T. Hence the polynomial $p(x) = (x - c_1) \ldots (x - c_k)$ will be the minimal polynomial for T, if $p(T) = \mathbf{0}$. Let v be an eigen vector of T. Then

$$(T - c_i I)(v) = \mathbf{0}, \text{ for some } i, 1 \leq i \leq k.$$

It follows that,

$$(T - c_1 I) \ldots (T - c_k I)(v) = \mathbf{0}, \text{ for each eigen vector } v. \qquad \ldots(1)$$

As T is diagonalizable, there exists a basis β of V, consisting of eigenvectors of T. Using this in equation (1), we get

$$(T - c_1 I) \ldots (T - c_k I)(x) = \mathbf{0}, \qquad \forall\ x \in V.$$

$$\therefore \qquad p(T) = (T - c_1 I) \ldots (T - c_k I) = \mathbf{0}.$$

Hence $p(x) = (x - c_1) \ldots (x - c_k)$ is the minimum polynomial for T.

Note: Above theorem **tells** us that, if T is a diagonalizable linear operator, then the minimal polynomial for T is a product of distinct linear factors.

Theorem 3: The minimal polynomial of a linear operator T divides its characteristic polynomial.

Proof: Let p(x) be the minimal polynomial of T.

$$\Rightarrow \qquad p(T) = \mathbf{0}.$$

Let f(x) be the characteristic polynomial of T. Then by Cayley-Hamilton theorem, $f(T) = \mathbf{0}$. Let c be a root of f(x). Then by **Division Algorithm,**

$$f(x) = (x - c)\, p(x) + q(x) \qquad \ldots (1)$$

where q(x) ∈ **F**[x] and either q(x) = **0** or deg q(x) < deg p(x). Suppose q(x) ≠ **0**. Then deg q(x) < deg p(x). Then from equation (1), we have

f(T) = (T – cI) p(T) + q(T)

⇒ **0** = (T – cI) **0** + q(T)

⇒ q(T) = **0**.

Contradiction! So our assumption was wrong.

⇒ q(x) = **0**.

Then from equation (1), we have f(x) = (x – c) p(x)

⇒ p(x) divides f(x).

Example 1: Let T be a linear operator on $\mathbf{R}^3$ which is represented in the standard ordered basis by the matrix A = $\begin{bmatrix} 1 & 2 & 0 \\ 2 & 1 & -6 \\ 2 & -2 & 3 \end{bmatrix}$. Prove that the characteristic polynomial of T is same as the minimal polynomial of T.

Solution: We can derive eigenvalues 5, 3, -3 of A as in the previous chapter. So characteristic polynomial of **T** is f(x) = (x – 5) (x – 3) (x + 3).

Since all characteristic values of T are distinct, so the minimal polynomial for T is

p(x) = (x – 5) (x – 3) (x + 3).

Hence f(x) = p(x).

Example 2: Let T be the linear operator on $\mathbf{R}^3$ which is represented in the standard ordered basis by A = $\begin{bmatrix} -9 & 4 & 4 \\ -8 & 3 & 4 \\ -16 & 8 & 7 \end{bmatrix}$. Find the minimal polynomial for T.

Solution: We can easily find |A – xI| = $\begin{vmatrix} -9-x & 4 & 4 \\ -8 & 3-x & 4 \\ -16 & 8 & 7-x \end{vmatrix}$

So, f(x) = det (A – xI) = (x – 3) $(x + 1)^2$.

So eigenvalues of T are 3, -1, -1.

Hence minimal polynomial is **either** p(x) = (x – 3) (x + 1) **or** p(x) = (x – 3) $(x + 1)^2$.

Case (i): Let p(x) = (x – 3) (x + 1) = $x^2 - 2x - 3$ then p(A) = $A^2 - A - 3I$.

Now, $A^2 = \begin{bmatrix} -9 & 4 & 4 \\ -8 & 3 & 4 \\ -16 & 8 & 7 \end{bmatrix}\begin{bmatrix} -9 & 4 & 4 \\ -8 & 3 & 4 \\ -16 & 8 & 7 \end{bmatrix}$

$$A^2 = \begin{bmatrix} -15 & 8 & 8 \\ -16 & 9 & 8 \\ -32 & 16 & 17 \end{bmatrix}$$

So $p(A) = \begin{bmatrix} -15 & 8 & 8 \\ -16 & 9 & 8 \\ -32 & 16 & 17 \end{bmatrix} - \begin{bmatrix} -18 & 8 & 8 \\ -16 & 6 & 8 \\ -32 & 16 & 14 \end{bmatrix} - \begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix}$

$$p(A) = \begin{bmatrix} -15 & 8 & 8 \\ -16 & 9 & 8 \\ -32 & 16 & 17 \end{bmatrix} + \begin{bmatrix} 15 & -8 & -8 \\ 16 & -9 & -8 \\ 32 & -16 & -17 \end{bmatrix}$$

$$p(A) = \mathbf{0}.$$

So p(x) = (x – 3) (x + 1) is the minimal polynomial. Now there is no need of second case.

Example 3: Let T be the linear operator on $\mathbf{C}^2$ which is represented in the standard ordered basis by the matrix $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$. Find the minimal polynomial for T.

Solution: The characteristic polynomial for T is det (A – xI) = $\begin{vmatrix} -x & -1 \\ 1 & -x \end{vmatrix} = x^2 + 1$. So eigen values of T are x = + i, – i. Since both eigen values are distinct . So minimal polynomial is

$$p(x) = (x - i)(x + i) = x^2 + 1.$$

Verification: $p(A) = A^2 + I = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

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

Example 4: Let V be a finite-dimensional vector space. Find the minimal polynomials for the operators I and **O**, where I is the identity operator and **O** is the zero operator on V.

Solution: (i) I(x) = x

Let p(x) = x – 1 then p(I) = I – I = **0**.

So obviously p(x) = x – 1 is the minimal polynomial for I.

(ii) Let p(x) = x then p(**0**) = **0**.

So p(x) = x is the minimal polynomial for **O**.

Example 5: Let T be the linear operator on $\mathbf{R}^3$ which is represented in the standard ordered basis by the matrix. A = $\begin{bmatrix} 2 & 1 & 0 \\ 0 & 1 & -1 \\ 0 & 2 & 4 \end{bmatrix}$. Find the minimal polynomial for T.

Solution: $|A - xI| = \begin{vmatrix} 2-x & 1 & 0 \\ 0 & 1-x & -1 \\ 0 & 2 & 4-x \end{vmatrix}$ = $(x-2)^2(3-x)$ = f(x). So eigen values are 2, 2, 3. So possible minimal polynomials are either p(x) = (3 – x) (x – 2) **or** p(x) = $(3-x)(x-2)^2$

Case I: Let p(x) = (3 – x) (x – 2)

$$p(A) = (3I - A)(A - 2I) = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 2 & 1 \\ 0 & -2 & -1 \end{bmatrix}\begin{bmatrix} 0 & 1 & 0 \\ 0 & -1 & -1 \\ 0 & 2 & 2 \end{bmatrix}$$

$$p(A) = \begin{bmatrix} 0 & 2 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \neq \mathbf{0}\ .$$

So p(x) = (3 –x) (x – 2) is not the minimal polynomial.

Case II: Now p(x) = $(3-x)(x-2)^2$. It can be easily verified that p(A) = $(3I - A)(A - 2I)^2$ = **0**.

So this is the minimal polynomial of T.

Example 6: Prove that the minimal polynomial of a linear operator T is a divisor of every polynomial that annihilates T.

Solution: Let p(x) be the minimal polynomial for T and h(x) be any polynomial that annihilates T.

Claim: We shall prove that p(x) divides h(x).

Since p(x), h(x) ∈ **F**[x]. So by **Division Algorithm** in **F**[x] , there exist two polynomials q(x), r(x) ∈ **F**[x] such that

$$h(x) = p(x)\, q(x) + r(x) \qquad \ldots (1)$$

where either $r(x) = \mathbf{0}$ or $\deg r(x) < \deg p(x)$.

So, $h(T) = p(T)\, q(T) + r(T)$

$\Rightarrow \quad \mathbf{0} = \mathbf{0} + r(T) \qquad \Rightarrow \quad r(T) = \mathbf{0}$.

If $r(x) \neq \mathbf{0}$, then $r(T) = \mathbf{0} \Rightarrow r(x)$ is the minimal polynomial.

Contradiction! So $r(x) = \mathbf{0}$, is the only choice. Then,

$h(x) = p(x)\, q(x)$

$\Rightarrow p(x)$ divides $h(x)$.

9.4 *INVARIANT SUBSPACES*

Let T be a linear operator on a vector space V(**F**). A subspace of V(**F**) is said to be invariant under T (or W is T-invariant) if $T(W) \subseteq W$. We can also say, W is invariant under T if $T(x) \in W$, for all $x \in W$.

Example 7: If T is any linear operator on a vector space V, then prove that ker (T) and Range (T) are invariant subspaces of V.

Solution: (**i**) We know that ker T = $\{ v \in V : T(V) = \mathbf{0} \}$. Since $T(\mathbf{0}) = \mathbf{0} \in$ ker T. For any $x \in$ ker T, $T(x) = \mathbf{0} \in$ ker T.

$\Rightarrow T(x) \in$ ker T, $\forall\, x \in$ ker T.

Hence ker T is an invariant subspace of V.

(**ii**) Let R(T) be the range of T. So for any $x \in R(T)$, $\exists$ some $v \in V$ such that $x = T(v)$

$\therefore T(x) = T(T(v)) = T(v_1)$, $v_1 = T(v) \in V$

$\Rightarrow \quad T(x) \in R(T) \quad \forall\; x \in R(T)$.

Hence Range T is an invariant subspace of V.

Example 8: Let T be a linear operator on a finite-dimensional vector space V(**F**) and let c be an eigenvalue of T. Prove that the eigenspace W_c is invariant under T.

Solution: Here W_c={ v ∈ V : T(v) = cv}. For any v ∈ W_c , we have T(v) = cv

⇒ T(T(v)) = T(cv).

⇒ T(T(v)) = c T(v)

If we take T(v) = v_1, then

T(v_1) = cv_1 ⇒ v_1 ∈ W_c *OR* T(v) ∈ W_c . So W_c is invariant under T.

Example 9: Prove that the space generated by (1, 1, 1) and (1, 2, 1) is an invariant subspace of $\mathbf{R}^3$ under T, where T(x, y, z) = (x + y – z, x + y, x + y – z) .

Solution: Suppose W be the subspace of $\mathbf{R}^3$ generated by (1, 1, 1) and (1, 2, 1). So

$$W = \{ \alpha (1, 1, 1) + \beta (1, 2, 1) : \alpha , \beta \in \mathbf{R}\}.$$

Given linear transformation is

$$T (x, y, z) = (x + y – z, x + y, x + y – z) \quad ….(1)$$

So T(1, 1, 1) = (1, 2, 1) ∈ W , and

T(1, 2, 1) = (2, 3, 2) = (1, 1, 1) + (1, 2, 1) ∈ W.

If w ∈ W. Then w = α (1, 1, 1) + β (1, 2, 1).

So, T(w) = α T(1, 1, 1) + β T(1, 2, 1) ∈ W.

⇒ T(w) ∈ W ∀ w ∈ W.

So W is invariant subspace of $\mathbf{R}^3$ under T.

Example 10: Let T be a linear operator on $\mathbf{R}^2$ whose matrix representation in the standard basis is $\begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}$. Let W be the subspace generated by e_1 ={ (1, 0)}. Prove that W is invariant under T.

Solution: As we know, standard basis of $\mathbf{R}^2(\mathbf{R})$ is {(1, 0), (0, 1)}. Let e_1 = (1, 0) , e_2 = = (0, 1). So by given matrix, T(e_1) = 2e_1 + 0e_2 and T (e_2) = 1e_1 + 2e_2 .

Given that

$W = L\{e_1\} = \{ \alpha e_1 : \alpha \in R\} = \{ \alpha (1, 0) : \alpha \in R\}$.

For any $x \in W$, we have $x = \alpha e_1$

$T(x) = T(\alpha e_1) = \alpha T(e_1) = \alpha (2e_1 + 0e_2)$

$T(x) = 2 \alpha e_1 \in W$ as $2\alpha \in \mathbf{R}$.

$\Rightarrow T(x) \in W \ \forall x \in W$.

$\Rightarrow$ W is invariant under T.

Example 11: Show that the subspace spanned by two subspaces, each of which is invariant under a linear operator T on V, is itself invariant under T.

Solution: Let W_1 and W_2 be two invariant subspaces of V under T. So

$$T(W_1) \subseteq W_1 \text{ and } T(W_2) \subseteq W_2 \quad \text{….(1)}$$

Let W be the subspace of V spanned by $W_1 \cup W_2$. So from the **vector spaces** chapter, we know that $W = W_1 + W_2$.

Claim: W is invariant under T. Let $w \in W$, then $w = w_1 + w_2$ where $w_1 \in W_1$, $w_2 \in W_2$. Then $T(w_1) \in W_1$ and $T(w_2) \in W_2$ ….(2)

So $T(w) = T(w_1 + w_2) = T(w_1) + T(w_2) \in W$.

Thus $T(w) \in W \ \forall \ w \in W$.

Hence W is invariant under T.

Example 12: Let T be a linear operator on a vector space V. If every subspace of V is invariant under T, then prove that T is a scalar multiple of the identity operator I on V.

Solution: Step I: Let $v \neq \mathbf{0}$ and $v \in V$ be arbitrary. Suppose $W = L\{v\} = \{ \alpha v : \alpha \in \mathbf{F} \}$. Then obviously W is a subspace of $V(\mathbf{F})$. By given hypothesis $T(W) \subseteq W$. We can write $v = 1v$ where $1 \in \mathbf{F}$. So $v \in W \Rightarrow T(v) \in W$

Hence $T(v) = \alpha v$, for some $\alpha \in \mathbf{F}$. …. (1)

Again, suppose $w \in W$ be arbitrary . Then $w = \beta v$ for some $\beta \in \mathbf{F}$.

$\Rightarrow T(w) = T(\beta v) = \beta\, T(v) = \beta\,(\alpha v)$; using (1)

$T(w) = (\beta\alpha)\, v = (\alpha\, \beta)\, v = \alpha\,(\beta v)$

$T(w) = \alpha w \quad \forall \quad w \in W$. ….(2)

So for every $w \in W$, T is a scalar multiple of the identity operator I on V.

Step II: If $v' \notin W$. Then obviously v, v' are linearly independent.

Let W' = L{v'}. Since, W' is a subspace of V, so by the given hypothesis, W' is invariant under T i.e. $T(W') \subseteq W'$. As discussed in equation (1) , we have

$T(v') = \alpha' v'$, for some $\alpha' \in \mathbf{F}$. …. (3)

Let W'' = L {v – v'}. **As argued above**, we have $T(v - v') = \gamma\,(v - v')$ for some $\gamma \in \mathbf{F}$.

$\Rightarrow \quad T(v) - T(v') = \gamma v - \gamma v'$

$\Rightarrow \quad \alpha v - \alpha' v' = \gamma v - \gamma v'$, using equations (1) and (3)

$\Rightarrow \quad (\alpha - \gamma)v + (\gamma - \alpha')\, v' = \mathbf{0}$.

But v and v' are linearly independent.

$\Rightarrow \quad \alpha - \gamma = \mathbf{0}$, and $\gamma - \alpha' = \mathbf{0}$.

$\Rightarrow \quad \alpha = \gamma = \alpha'$.

Putting α' in equation (3), we get

$T(v') = \alpha\, v' \quad \forall \quad v' \notin W$. …. (4)

From (2) and (4), we have $T(x) = \alpha\, x \quad \forall\, x \in V$.

$\Rightarrow \quad T(x) = \alpha\, I(x) = (\alpha I)\,(x) \quad \forall\, x \in V$.

Hence T = αI , $\alpha \in$ **F**.

Example 13: Let T be a linear operator on $\mathbf{R^2}$. Let its matrix in the standard ordered basis is

A = $\begin{bmatrix} 1 & -1 \\ 2 & 2 \end{bmatrix}$. Then

(i) Prove that the only subspaces of $\mathbf{R^2}$ invariant under T are $\mathbf{R^2}$ and the zero subspace.

(ii) If U is the linear operator on $\mathbf{C^2}$, the matrix of which in the standard ordered basis is A, show that U has 1-dimensional invariant subspaces.

Solution: (**i**) Suppose W be any proper subspace of $\mathbf{R^2}$ such that W is invariant under T. Then obviously dim W = 1.

$\Rightarrow$ W is spanned by some $\mathbf{0} \neq w \in$ W.

Since W is invariant under T. So T(w) $\in$ W $\forall$ w $\in$ W. Let T(w) = cw, for some c $\in$ **R**.

$\Rightarrow$ c is an eigenvalue of **T**.

All the eigenvalues of T are given by

$$|A - xI| = \begin{vmatrix} 1-x & -1 \\ 2 & 2-x \end{vmatrix} = x^2 - 3x + 4 = 0.$$

$\Rightarrow$ $x = \frac{1}{2}(3 \pm i\sqrt{7}) \notin$ **R.**

$\Rightarrow$ c $\notin$ **R** , a **contradiction !**

Hence there does **NOT** exist any proper subspace of $\mathbf{R^2}$ which is invariant under T.

$\Rightarrow$ Only subspaces of $\mathbf{R^2}$ invariant under T are $\mathbf{R^2}$ and {**0**}.

(**ii**) As discussed above, eigenvalues of U on $\mathbf{C^2}$ are

$c_1 = \frac{1}{2}(3 + i\sqrt{7}) \in$ **C** ; $c_2 = \frac{1}{2}(3 - i\sqrt{7}) \in$ **C.**

Now W_{c_1} = { v $\in$ V : T(v) = c_1v }.

Claim: W_{c_1} is invariant under T. For any v $\in W_{c_1}$, we have T(v) = c_1v.

$\Rightarrow \quad T(T(v)) = T(c_1 v)$

$\Rightarrow \quad T(T(v)) = c_1 T(v)$

Let $T(v) = v_1$.

$\Rightarrow \quad T(v_1) = c_1 v_1$

$\Rightarrow \quad v_1 \in W_{c_1}$ or $T(v) \in W_{c_1}\ \forall\ v \in W_{c_1}$.

So W_{c_1} is invariant under T. Similarly, we can show that W_{c_2} is also invariant. Also dim W_{c_1} = dim W_{c_2} = 1.

Theorem 4: Let T be a linear operator on a finite-dimensional vector space V(**F**). Let W be an invariant subspace of T. Then T has a matrix representation $\begin{bmatrix} A & B \\ O & C \end{bmatrix}$, where A is matrix of restriction of T on W.

Proof: Let $\beta_1 = \{ w_1, \ldots, w_r \}$ be a basis of W. Since $w_1, \ldots, w_r$ are linearly independent vectors of V, they can be extended to form a basis of V. Suppose $\beta = \{ w_1, \ldots, w_r, v_1, \ldots, v_s \}$ be a basis of V. Given that W is invariant under T, so $T(x) \in W$, for each $x \in W$ .Let us define a mapping $T_W : W \rightarrow W$ by

$$T_W(x) = T(x)\ \forall\ x \in W. \quad \ldots(1)$$

Definitely T_W will be a linear operator. Here T_W is called restriction of T on W. From equation (1), we have

$T(w_1) = T_W(w_1) = a_{11}w_1 + \ldots + a_{r1}w_r$

$T(w_2) = T_W(w_2) = a_{12}w_1 + \ldots + a_{r2}w_r$

….. …… ….. …… …… ….. ….. ….. ….

$T(w_r) = T_W(w_r) = a_{1r}w_1 + \ldots + a_{rr}w_r \quad ; \quad a_{ij} \in \mathbf{F}.$

Also we have

$T(v_1) = b_{11}w_1 + \ldots + b_{r1}w_r + c_{11}v_1 + \ldots + c_{s1}v_s$

$T(v_2) = b_{12}w_1 + \ldots + b_{r2}w_r + c_{12}v_1 + \ldots + c_{s2}v_s$

….. ….. ….. ….. ….. ….. ….. ….. ….. …. …..

$T(v_s) = b_{1s}w_1 + ….. + b_{rs}w_r + c_{1s}v_1 + ….. + c_{ss}v_s$

From these equations, we obtain

$$[T_W]_{\beta_1} = \begin{bmatrix} a_{11} & \cdots & a_{1r} \\ \vdots & \ddots & \vdots \\ a_{r1} & \cdots & a_{rr} \end{bmatrix}_{r\times r} = A, \text{ say}$$

$$[T]_\beta = \begin{bmatrix} a_{11} & \cdots & a_{1r} & b_{11} & \cdots & b_{1s} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{r1} & \cdots & a_{rr} & b_{r1} & \cdots & b_{rs} \\ 0 & \cdots & 0 & c_{11} & \cdots & c_{1s} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & c_{s1} & \cdots & c_{ss} \end{bmatrix} = \begin{bmatrix} A & B \\ O & C \end{bmatrix}$$

where $B = \begin{bmatrix} b_{11} & \cdots & b_{1s} \\ \vdots & \ddots & \vdots \\ b_{r1} & \cdots & b_{rs} \end{bmatrix}$; $C = \begin{bmatrix} c_{11} & \cdots & c_{1s} \\ \vdots & \ddots & \vdots \\ c_{s1} & \cdots & c_{ss} \end{bmatrix}$

Here A is the matrix of restriction of T on W.

Theorem 5: Let T be a linear operator on a finite-dimensional vector space V(**F**). Let W and U be invariant subspaces of **T** such that V = W ⊕ U. Then T has a matrix representation $\begin{bmatrix} A & O \\ O & C \end{bmatrix}$, where A and C are the matrices of restriction of T on W and U respectively.

Proof: Let $\beta_1 = \{ w_1 ,…., w_r \}$ be a basis of W and $\beta_2 = \{ v_1 ,…., v_s \}$ be a basis of U. Given that V = W ⊕ U . So W ∩ U = {**0**}. So dim V = dim W + dim U.

Hence $\beta = \{ w_1 ,…., w_r ,v_1 ,…., v_s\}$ is basis of V. As W is invariant under T. So T(x) ∈ W, for each x ∈ W. Let us define, $T_W : W \to W$ by $T_W(x) = T(x)\ \forall\ x \in W$. Then obviously, T_W is a linear operator on W. Here T_w is restriction of T on W.

Now, $T(w_1) = T_W(w_1) = a_{11}w_1 + …. + a_{r1}w_r$

…. …. …. …. …. …. …. …. …. …. … …

$T(w_r) = T_W(w_r) = a_{1r}w_1 + …. + a_{rr}w_r$; $a_{ij} \in$ **F.**

Hence $[T_W]_{\beta_1} = \begin{bmatrix} a_{11} & \cdots & a_{1r} \\ \vdots & \ddots & \vdots \\ a_{r1} & \cdots & a_{rr} \end{bmatrix} = A$, say

Similarly $[T_U]_{\beta_2} = \begin{bmatrix} c_{11} & \cdots & c_{1s} \\ \vdots & \ddots & \vdots \\ c_{s1} & \cdots & c_{ss} \end{bmatrix}$ = C, say

Then it can be easily observed that

$$[T]_\beta = \begin{bmatrix} a_{11} & \cdots & a_{1r} & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{r1} & \cdots & a_{rr} & 0 & \cdots & 0 \\ 0 & \cdots & 0 & c_{11} & \cdots & c_{1s} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & c_{s1} & \cdots & c_{ss} \end{bmatrix} = \begin{bmatrix} A & O \\ O & C \end{bmatrix}.$$

Example 14: Let V be the vector space of all polynomials in x over **F**, of degree ≤ 5. Let T : V→ V be defined by, $T(1) = x^2 + x^4$, $T(x) = x + 1$, $T(x^2) = 1$, $T(x^3) = x^3 + x^2 + 1$, $T(x^4) = x^4$, $T(x^5) = 0$.

(i) If W is the linear span of $\{1, x^2, x^4\}$, show that W is invariant under T.
(ii) Also find the matrix of T in a suitable basis of V.

Solution: (i) Let $w \in W$, $\Rightarrow$ $w = \alpha + \beta x^2 + \gamma x^4$; $\alpha, \beta, \gamma \in$ **F.** We have

$$T(w) = T(\alpha 1 + \beta x^2 + \gamma x^4)$$

$$= \alpha\, T(1) + \beta\, T(x^2) + \gamma\, T(x^4)$$

$$= \alpha\, (x^2 + x^4) + \beta 1 + \gamma\, x^4$$

$$= \beta 1 + \alpha x^2 + (\alpha + \gamma)\, x^4 \in W$$

So $T(w) \in W$, for each $w \in W$

$\Rightarrow$ W is invariant under T.

(ii) We shall find the matrix of the restriction of T on W. A basis of W is $\beta_1 = \{1, x^2, x^4\}$.

So
$$\left.\begin{aligned} T_W(1) &= T(1) = x^2 + x^4 = 0.1 + 1.x^2 + 1.x^4 \\ T_W(x^2) &= T(x^2) = 1 = 1.1 + 0.x^2 + 0.x^4 \\ T_W(x^4) &= T(x^4) = x^4 = 0.1 + 0.x^2 + 1.x^4 \end{aligned}\right\} \quad \ldots..(1)$$

Hence $[T_W]_\beta = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix}$ = A , say.

Now we find the matrix of T relative to the basis $\beta = \{1, x^2, x^4, x, x^3, x^5\}$. So we have

$$\left.\begin{aligned} T(x) &= x+1 = 1.1 + 0.x^2 + 0.x^4 + 1.x + 0.x^3 + 0.x^5 \\ T(x^3) &= x^3+x^2+1 = 1.1 + 1.x^2 + 0.x^4 + 0.x + 1.x^3 + 0.x^5 \\ T(x^5) &= 0 = 0.1 + 0.x^2 + 0.x^4 + 0.x + 0.x^3 + 0.x^5 \end{aligned}\right\} \quad \ldots\ldots (2)$$

From equations (1) and (2), the matrix of T with respect to the basis β is

$$[T]_\beta = \begin{bmatrix} 0 & 1 & 0 & \vdots & 1 & 1 & 0 \\ 1 & 0 & 0 & \vdots & 0 & 1 & 0 \\ 1 & 0 & 1 & \vdots & 0 & 0 & 0 \\ \cdots & \cdots & \cdots & \vdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \vdots & 1 & 0 & 0 \\ 0 & 0 & 0 & \vdots & 0 & 1 & 1 \\ 0 & 0 & 0 & \vdots & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} A & B \\ O & C \end{bmatrix}.$$

where $B = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, $C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$.

Example 15: Let V be the vector space of all polynomials of degree less than or equal to six. Let W be the subspace of V spanned by $\{1, x^2, x^4, x^6\}$. Let D be the differential operator on V i.e. $D\, f(x) = \frac{d}{dx}(f(x))$.

(i) Show that W is not invariant under D.

(ii) Let $T = D^2$, where $D^2 f(x) = \frac{d^2}{dx^2}(f(x))$. Show that W is invariant under T.

(iii) Find the matrix of T_W in a suitable basis of W, where T_W is the restriction of T on W.

(iv) Find the matrix of T in a suitable basis of V.

Solution: (**i**) Since W is spanned by $\{1, x^2, x^4, x^6\}$. So $\{1, x^2, x^4, x^6\} \subset W$. Now $x^2 \in W$, but $D(x^2) = \frac{d}{dx}(x^2) = 2x \notin W$. So W is not invariant under D.

(ii) Let $f(x) \in W$. Then $f(x) = \alpha_1.1 + \alpha_2.x^2 + \alpha_3.x^4 + \alpha_4.x^6$; $\alpha_i \in$ **F**

$\Rightarrow$ $D\,f(x) = 0 + 2\,\alpha_2.x + 4\,\alpha_3.x^3 + 6\,\alpha_4.x^5$

$\Rightarrow$ $D^2 f(x) = 2\,\alpha_2 + 12\,\alpha_3.x^2 + 30\,\alpha_4.x^4$

Given $D^2 = T$, so $T\,f(x) = (2\alpha_2).1 + (12\,\alpha_3)\,x^2 + (30\,\alpha_4)\,x^4 + 0.x^6 \in W$

Hence T(f(x)) $\in$ W , for each f(x) $\in$ W .

Hence W is invariant under T = D^2 .

(iii) A basis of W is $\beta_1 = \{ 1, x^2, x^4 , x^6 \}$. We have

$$T_W(1) = D^2(1) = 0 = 0.1 + 0.x^2 + 0.x^4 + 0.x^6$$

$$T_W(x^2) = D^2(x^2) = 2 = 2.1 + 0.x^2 + 0.x^4 + 0.x^6$$

$$T_W(x^4) = D^2(x^4) = 12x^2 = 0.1 + 12x^2 + 0.x^4 + 0.x^6$$

$$T_W(x^6) = D^2(x^6) = 30x^4 = 0.1 + 0.x^2 + 30x^4 + 0.x^6$$

Hence, the matrix of T_W in the basis β_1 of W is A = $\begin{bmatrix} 0 & 2 & 0 & 0 \\ 0 & 0 & 12 & 0 \\ 0 & 0 & 0 & 30 \\ 0 & 0 & 0 & 0 \end{bmatrix}$.

(iv) Now, we find the matrix of T, relative to the basis $\beta = \{ 1, x^2, x^4 , x^6, x , x^3 , x^5\}$ of V . Here T = D^2

$$T(1) = 0 = 0.1 + 0.x^2 + 0.x^4 + 0.x^6 + 0.x + 0.x^3 + 0.x^5$$

$$T(x^2) = 2 = 2.1 + 0.x^2 + 0.x^4 + 0.x^6 + 0.x + 0.x^3 + 0.x^5$$

$$T(x^4) = 12x^2 = 0.1 + 12.x^2 + 0.x^4 + 0.x^6 + 0.x + 0.x^3 + 0.x^5$$

$$T(x^6) = 30x^4 = 0.1 + 0.x^2 + 30.x^4 + 0.x^6 + 0.x + 0.x^3 + 0.x^5$$

$$T(x) = 0 = 0.1 + 0.x^2 + 0.x^4 + 0.x^6 + 0.x + 0.x^3 + 0.x^5$$

$$T(x^3) = 6x = 0.1 + 0.x^2 + 0.x^4 + 0.x^6 + 6.x + 0.x^3 + 0.x^5$$

$$T(x^5) = 20x^3 = 0.1 + 0.x^2 + 0.x^4 + 0.x^6 + 0.x + 20.x^3 + 0.x^5$$

So $$[T]_\beta = \begin{bmatrix} 0 & 2 & 0 & 0 & \vdots & 0 & 0 & 0 \\ 0 & 0 & 12 & 0 & \vdots & 0 & 0 & 0 \\ 0 & 0 & 0 & 30 & \vdots & 0 & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \vdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & \vdots & 0 & 6 & 0 \\ 0 & 0 & 0 & 0 & \vdots & 0 & 0 & 20 \\ 0 & 0 & 0 & 0 & \vdots & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} A & O \\ O & C \end{bmatrix}.$$

where $C = \begin{bmatrix} 0 & 6 & 0 \\ 0 & 0 & 20 \\ 0 & 0 & 0 \end{bmatrix}$.

9.5 *DIRECT SUM DECOMPOSITION*

Independent subspaces: The subspace $W_1, \ldots, W_k$ of V(**F**) are called independent if for $x_i \in W_i$; i = 1, 2, …., k, we have $x_1 + x_2 + \ldots + x_k = \mathbf{0}$,

$\Rightarrow \quad x_i = \mathbf{0}, \ \forall \ i = 1, 2, \ldots, k.$

Theorem 6: Let V be a finite-dimensional vector space over **F**. Let $W_1, \ldots, W_k$ be subspaces of V and $W = W_1 + \ldots + W_k$. Then the following conditions are equivalent:

(i) $W_1, W_2 \ldots, W_k$ are independent.
(ii) $W_j \cap (W_1 + \ldots + W_{j-1}) = \{\mathbf{0}\}$ for all j ; $2 \leq j \leq k$.
(iii) If β_i is an ordered basis of W_i, $1 \leq i \leq k$, then $\beta = \{\beta_1, \ldots, \beta_k\}$ is an ordered basis of W.

Proof: Step I: We shall prove **(i)** $\Rightarrow$ **(ii).** Suppose $W_1, \ldots, W_k$ be independent subspaces of V(**F**).

Let $x \in W_j \cap (W_1 + \ldots + W_{j-1})$; $2 \leq j \leq k$. Then $x \in W_j$ and $x \in W_1 + \ldots + W_{j-1}$.

So $x \in W_j$ and $x = x_1 + \ldots + x_{j-1}$, where $x_i \in W_i$ for I = 1, 2, …., j – 1.

$\Rightarrow \quad x_1 + \ldots + x_{j-1} + (-x) = \mathbf{0}.$

$\Rightarrow \quad x_1 = \mathbf{0}, \ldots, x_{j-1} = \mathbf{0}, x = \mathbf{0}$; since $W_1, \ldots, W_k$ are independent.

$\Rightarrow \quad x = \mathbf{0}.$

So, $W_j \cap (W_1 + \ldots + W_{j-1}) = \{\mathbf{0}\}$; $\quad 2 \leq j \leq k$.

Step II: Now we prove **(ii)** $\Rightarrow$ **(iii).**

Let $\beta_i = \{ b_1^i, b_2^i, \ldots, b_{d_i}^i \}$ be a basis of W_i ; I = 1, 2, …., k.

Claim: $\beta = \{\beta_1, \ldots, \beta_k\}$ is a basis of W.

(a) First we shall show linear independence of elements of β.

Let $\sum_{i=1}^{k}(\alpha_1^i b_1^i + \alpha_2^i b_2^i + \ldots + \alpha_{d_i}^i b_{d_i}^i) = \mathbf{0}$; $\alpha_j^i \in \mathbf{F}$.

or $\sum_{i=1}^{k} w_i = \mathbf{0}$ where $w_i = \sum_{j=1}^{d_i} \alpha_j^i b_j^i \in W_i$.

or $w_1 + w_2 + \ldots + w_k = \mathbf{0}$.

$\Rightarrow$ $w_i = \mathbf{0}$ for i = 1, 2, …., k , **because if**, j is the largest +ve integer

such that $w_j \neq \mathbf{0}$, then $w_1 + w_2 + \ldots + w_j = \mathbf{0}$ $\Rightarrow$ $-w_j = w_1 + \ldots + w_{j-1}$

$\Rightarrow$ $w_j \in W_j \cap (W_1 + \ldots + W_{j-1})$; by part **(ii)**

$\Rightarrow$ $w_j = \mathbf{0}$.

Now for each i, $w_i = \mathbf{0}$ $\Rightarrow$ $\sum_{j=1}^{d_i} \alpha_j^i b_j^i = \mathbf{0}$. Since β_i is a linearly independent set for each i, therefore $\alpha_j^i = 0$; $\forall$ j = 1, 2, …., d_i and $\forall$ i = 1, 2, …., k .Thus β is a linearly independent subset of W.

(b) Now we shall prove that β spans W. Let $x \in W = W_1 + \ldots + W_k$ be arbitrary.

Then $x = \sum_{i=1}^{k} w_i$, where $w_i \in W_i$. As β_i is a basis of W_i, so $w_i = \sum_{j=1}^{d_i} \lambda_j^i b_j^i$; $\lambda_j^i \in \mathbf{F}$.

So, $x = \sum_{i=1}^{k}\sum_{j=1}^{d_i} \lambda_j^i b_j^i$, $\forall x \in W$.

$\Rightarrow$ β spans W.

Hence β is a basis of W.

Step III: Now we shall show **(iii) ⇒ (i).**

Suppose $w_1 + \ldots + w_k = \mathbf{0}$, or $\sum_{i=1}^{k} w_i = \mathbf{0}$.

As β_i is a basis of W_i, so $w_i = \sum_{j=1}^{d_i} \lambda_j^i b_j^i$; $\lambda_j^i \in \mathbf{F}$.

$\Rightarrow$ $\sum_{i=1}^{k}\sum_{j=1}^{d_i} \lambda_j^i b_j^i = \mathbf{0}$.

⇒ $\lambda_j^i = 0$, $\forall$ j = 1, 2, …., d_i ; $\forall$ i = 1, 2, …., k as β is a basis of W, so linearly independent.

⇒ $w_i = \mathbf{0}$, for each i = 1, 2, …., k .

⇒ $W_1,.., W_k$ are independent.

Example 16: Prove that if W_1, W_2, W_3 are independent subspaces of V, then $W_i \cap W_j = \mathbf{0}$, for $i \neq j$; i, j = 1, 2, 3. Give an example to show that its converse **may not** be true.

Solution: For $i \neq j$, we have $W_j = \sum_{\substack{j=1 \\ j\neq 1}}^{3} W_j$. Since W_1, W_2, W_3 are independent . So by previous theorem, $W_i \cap W_j \subseteq W_i \cap \sum_{\substack{j=1 \\ j\neq 1}}^{3} W_j = \{\mathbf{0}\}$. Hence $W_i \cap W_j = \{\mathbf{0}\}$ for $i \neq j$ and i, j = 1, 2, 3.

Conversely, let us take $V = R^3(\mathbf{R})$, $W_1 = L\,\{(1, 0, 0) = e_1\}$

$W_2 = L\,\{(0, 1, 0) = e_2\}$ and $W_3 = L\,\{e_1 + e_2\}$. Obviously $W_1 \cap W_2 = \{\mathbf{0}\}$.

Let $x \in W_1 \cap W_3$. Then $x \in W_1$ and $x \in W_3$.

⇒ $x = \alpha e_1$ and $x = \beta\,(e_1 + e_2)$ where $\alpha, \beta \in \mathbf{R}$.

⇒ $\alpha e_1 = \beta e_1 + \beta e_2$

⇒ $(\alpha - \beta)e_1 - \beta e_2 = \mathbf{0}$.

But e_1 and e_2 are linearly independent vectors.

⇒ $\alpha - \beta = 0$ and $\beta = 0$.

⇒ $\alpha = \beta = 0$.

⇒ $x = \mathbf{0}$.

So $W_1 \cap W_3 = \{\mathbf{0}\}$.

Similarly we can prove that $W_2 \cap W_3 = \{\mathbf{0}\}$. Now suppose $x \in W_3$. Then $x = \gamma\,(e_1 + e_2)$; $\gamma \in$ **R.**

⇒ $x = \gamma e_1 + \gamma e_2 \in W_1 + W_2$.

So $W_3 \subseteq W_1 + W_2$.

$\Rightarrow \quad W_3 \cap (W_1 + W_2) = W_3 \neq \{\mathbf{0}\}$.

$\Rightarrow \quad W_1, W_2, W_3$ are **NOT** independent subspaces.

9.6 *PROJECTION ON A VECTOR SPACE*

A linear operator E on a vector space V is called a projection if $E^2 = E$. It means E is independent.

Theorem 7: If E is a projection on a vector space V(F), then $V = R \oplus N$, where R is the range space of E and N is the null space of E.

Proof: Here E is a linear operator. So

$$R = \{ E(x) : x \in V\} \text{ and } N \{ x \in V: E(x) = \mathbf{0}\}.$$

Step I: Here we shall prove that $x \in R \Leftrightarrow E(x) = x$. So by definition, $x \in R \Rightarrow x = E(y)$ for some $y \in R$.

So $E(x) = E(E(y)) = E^2(y) = E(y)$ as $E^2 = E$.

$\Rightarrow \quad E(x) = E(y) = x$.

Conversely , $x = E(x) \Rightarrow x \in R$.

Step II: Here we shall prove that $V = R + N$ and $R \cap N = \{\mathbf{0}\}$. Let $x \in V$ be arbitrary. Then we can write

$$x = E(x) + (x - E(x)) \qquad \text{....(1)}$$

Here $E(x) \in R$. Also $E - (x - E(x)) = E(x) - E^2(x) = E(x) - E(x) = \mathbf{0}$.

$\Rightarrow \quad x - E(x) \in N$.

So $x \in R + N \ \forall x \in V$.

Hence $V = R + N$. Again, let $x \in R \cap N$ be arbitrary.

⇒ $x \in R$ and $x \in N$.

Now $x \in R \Rightarrow x = E(x)$ and $x \in N \Rightarrow E(x) = \mathbf{0}$.

Thus $x = \mathbf{0}$.

⇒ $R \cap N = \{\mathbf{0}\}$.

Finally $V = R \oplus N$.

Note: (1) If $V = R \oplus N$, then we say that E is the projection on R along N .

(2) We discussed that $x \in R \Leftrightarrow E(x) = x$. Also $V = R \oplus N$. so each $v \in V$ is uniquely expressible as $v = r + n$, where $r \in R$ and $n \in N$. So $E(v) = E(r + n) = E(r) + E(n) = r + \mathbf{0} = r$.

Hence if E is the projection on R along N, then for each $v \in V$, such that $v = r + n$, $E(v) = r$.

(3) Any projection E on V is diagonalizable.

We have already discussed that $V = R \oplus N$. Let $\{ v_1, ..., v_r\}$ be a basis of R and $\{v_{r+1}, ..., v_n\}$ a basis of N. Then $\beta = \{v_1, ..., v_r, v_{r+1}, ..., v_n\}$ is a basis of $V = R \oplus N$. Hence the matrix of E with respect to the basis β is the diagonal matrix,

$$[E]_\beta = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}, \text{ I is } r \times r \text{ unit matrix.}$$

Example 17: A linear operator E on V is a projection if and only if I – E is a projection.

Solution: Let E be a projection. Then $E^2 = E$. So

$$(I - E)^2 = (I - E)(I - E) = I^2 - E - E + E^2$$

$$= I - 2E + E \text{ as } I^2 = I \text{ and } E^2 = E.$$

$$= I - E$$

⇒ I – E is also a projection.

Conversely, let I – E be a projection. Then $(I - E)^2 = I - E$

⇒ $I^2 - 2E + E^2 = I - E$

$\Rightarrow \quad E^2 = 2E - E$

$\Rightarrow \quad E^2 = E.$

$\Rightarrow$ E is also a projection.

Example 18: If E_1 and E_2 are projections on V such that $E_1E_2 = E_2E_1$. Prove that E_1E_2 and $E_1 + E_2 - E_1E_2$ are projections.

Solution: Given that $E_1E_2 = E_2E$ …. (1)

Now, $(E_1E_2)^2 = (E_1E_2)(E_1E_2) = E_1(E_2E_1)E_2 = E_1(E_1E_2)E_2$

$= E_1^2 E_2^2$

$(E_1E_2)^2 = E_1E_2$ as $E_1^2 = E_1$ and $E_2^2 = E_2$

$\Rightarrow$ E_1E_2 is also a projection.

Now, $(E_1 + E_2 - E_1E_2)^2 = (E_1 + E_2 - E_1E_2)(E_1 + E_2 - E_1E_2)$

$= E_1^2 + E_1 E_2 - E_1^2 E_2 + E_2 E_1 + E_2^2 - E_2E_1E_2 - E_1 E_2 E_1 - E_1 E_2^2 + E_1 E_2 E_1 E_2$

$= E_1 + E_1 E_2 - E_1 E_2 + E_1 E_2 + E_2 - E_1 E_2^2 - E_1^2 E_2 - E_1 E_2^2 + E_1^2 E_2^2$

$= E_1 + E_1 E_2 - E_1 E_2 + E_1 E_2 + E_2 - E_1 E_2 - E_1 E_2 - E_1 E_2 + E_1E_2$

$= E_1 + E_2 - E_1E_2$

So, $E_1 + E_2 - E_1E_2$ is also a projection.

Example 19: Let V be a real vector space and E be a projection on V. Prove that I + E is invertible and also find $(I + E)^{-1}$.

Solution: Since E is a linear operator. So αE is also $\forall\ \alpha \in R$. Let $\alpha = \frac{1}{2}$. So $\frac{1}{2}E$ is also a linear operator.

Now $(I + E)(I - \frac{E}{2}) = I(I - \frac{E}{2}) + E(I - \frac{E}{2})$

$= I - \frac{E}{2} + E - \frac{1}{2}E^2$

$= I + \frac{E}{2} - \frac{1}{2}E$ as $E^2 = E$

$(I + E)(I - \frac{1}{2}E) = I.$

Similarly, we can prove , $(I - \frac{1}{2}E)(I + E) = I$.

So $I + E$ is invertible and $(I + E)^{-1} = I - \frac{1}{2}E$.

Example 20: If E is the projection on R along N, then prove that I – E is the projection on N along R.

Solution: Since E is the projection on R along N. So V = R ⊕ N,

Claim: We shall prove that: Range space of (I – E) = N and null space of (I – E) = R.

(**i**) As we know, for any x ∈ R, x = E(x). So (I – E) (x) = I(x) – E(x) = x – x =**0.**

⇒ x ∈ ker (I – E).

⇒ R ⊆ ker (I – E). ….. (1)

Now for any x ∈ ker (I – E), we have, (I – E) (x) =**0.**

⇒ I(x) – E(x) = **0.**

⇒ x – E(x) = **0,** as I is identity operator.

⇒ x = E(x). So x ∈ R

⇒ ker (I – E) ⊆ R. …. (2)

Hence from equations (1) and (2), we have, R = ker (I –E) = null space of (I – E).

(**ii**) Since V = R ⊕ N , so each v ∈ V is **uniquely** expressible as v = r + n, where r ∈ R and n ∈ N. i.e. r = E(r) and E(n) =**0**.

So, E(v) = E(r + n) = E(r) + E(n) = r + 0 = r.

⇒ (I – E) (v) = I(v) – E(v) = v – r = n ∈ N.

So range space of (I – E) $\subseteq$ N. ….. (3)

Conversely, let n $\in$ N , then E(n) =**0.**

So (I – E) (n) = I(n) – E(n) = n – 0 = n.

$\Rightarrow$ n $\in$ range space of (I – E).

$\Rightarrow$ N $\subseteq$ range space of (I – E). …. (4)

From equations (3) and (4), we obtain, N = range space of (I – E).

Step II: Now,

$(I – E)^2 = (I – E) (I – E) = I(I – E) – E(I – E)$

$= I^2 – E – E + E^2.$

$(I – E)^2 = I – E$, as $E^2 = E$.

Also V = R ⊕ N = N ⊕ R. Hence I – E is the projection on N along R.

Example 21: If E is a projection and f is a polynomial, then f(E) = aI + bE. Discuss it.

Solution: Suppose $f(x) = a_0 + a_1x + a_2x^2 + …. + a_nx^n \in \mathbf{F}[x]$.

So, $f(E) = a_0I + a_1E + a_2E^2 + a_3E^3 + …. + a_nE^n.$

$= a_0I + a_1E + a_2E + a_3E + …. + a_nE$, as $E^2 = E$.

$= a_0I + (a_1 + a_2 + a_3 + ….+ a_n)E.$

= aI + bE, where $a = a_0$, $b = a_1 + a_2 + ….+ a_n$.

Example 22: Find a projection E which projects $\mathbf{R}^2$ onto the subspace spanned by (1, -1) along the subspace spanned by (1, 2).

Solution: Let W_1 and W_2 be subspaces of $\mathbf{R}^2$ spanned by (1, -1) and (1, 2) respectively. So W_1 = { x(1, -1) : x $\in \mathbf{R}$ } and W_2 = { y(1, 2) : y $\in \mathbf{R}$ }

Step I: We shall prove that S = {(1, -1), (1,2)} is a linearly independent set. Suppose

α (1, -1) + β (1, 2) = 0 , α , $\beta \in$ **R.**

$\Rightarrow$ $\alpha + \beta = 0.$

$-\alpha + 2\beta = 0.$

Coefficient matrix = A = $\begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix}$ $R_2 \to R_2 + R_1$

$= \begin{bmatrix} 1 & 1 \\ 0 & 3 \end{bmatrix}.$

So rank of A = 2. Hence above equation system has only one solution i.e. $\alpha = \beta = 0$. So S = {(1, -1), (1, 2) } is a linearly independent set.

Step II: The standard ordered basis of $\mathbf{R}^2$ is {e_1 = (1, 0), e_2 = (0, 1)}. Let (α , β) = a (1, -1) + b (1, 2).

$\Rightarrow$ $a + b = \alpha$

$-a + 2b = \beta$

On solving , we get $b = \frac{\alpha + \beta}{3}$, $a = \frac{2\alpha - \beta}{3}$

So $(\alpha , \beta) = \left(\frac{2\alpha - \beta}{3}\right)$ (1, -1) + $\left(\frac{\alpha + \beta}{3}\right)$ (1, 2) ….(1)

If $\alpha = 1$, $\beta = 0$, we get

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

If $\alpha = 0$, $\beta = 1$, we get

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

Also from equation (1), we conclude

$(\alpha , \beta) = \alpha e_1 + \beta e_2 = \left(\frac{2\alpha - \beta}{3}\right)(1, -1) + \left(\frac{\alpha + \beta}{3}\right)(1, 2) \in W_1 + W_2$ ….(4)

So $R^2 = W_1 + W_2.$

It can be easily proved that $W_1 \cap W_2 = \{(0, 0)\}$. Hence $R^2 = W_1 \oplus W_2$.

Step III: As we know if $V = R \oplus N$, then the projection E on R along N is given by E(v) = r, where $v \in V$ has a unique representation $v = r + n$; $r \in R$, $n \in N$. Hence by equation (4), the projection E on W_1 along W_2 is given by $E : \mathbf{R^2} \to \mathbf{R^2}$, where $E(\alpha, \beta) = \left(\frac{2\alpha - \beta}{3}\right)$ (1, -1)

$\Rightarrow$ $E(\alpha, \beta) = \left(\frac{2\alpha - \beta}{3}, \frac{\beta - 2\alpha}{3}\right)$; $\forall (\alpha, \beta) \in \mathbf{R^2}$.

Note: Here we observe that, E(1, -1) = (1, -1) and E(1, 2) = (0, 0).

Hence W_1 is range of E and W_2 is the null space of E.

Example 23: Assume that E be a projection on V and T be a linear operator on V. Prove that –

(i) The range of E is invariant under T if and only if E T E = T E.
(ii) Both the range and null space of E are invariant under T if and only if E T = T E

Solution: Given that E is a projection on V, so $V = R \oplus N$.

Here R = E(v) = range of E

N = ker E.

(i) Let the range of E be invariant under T. Then

$T(R) \subset R \Rightarrow T(E(v)) \subset E(v)$.

$\Rightarrow$ $T(E(x)) \in E(v)$, $\forall x \in V$.

So $T(E(x)) = E(v_0)$, for some $v_0 \in V$.

$\Rightarrow$ $(TE)(x) = E(v_0)$ $\forall x \in V$. …. (1)

Now $(E\,T\,E)(x) = E(T\,E(x)) = E(E(v_0)) = E^2(v_0) = E(v_0)$.

$\Rightarrow$ $(E\,T\,E)(x) = (T\,E)(x)$ $\forall x \in V$; using (1).

$\Rightarrow$ E T E = T E.

Conversely, let E T E = T E.

⇒ $(TE)(x) = (E\,T\,E)(x)$, $\forall\ x \in V$.

⇒ $T(E(x)) = E(T\,E(x))$

$= E(x_1)$, where $x_1 = (T\,E)(x) \in V$.

⇒ $T(E(x)) \in E(v)\ \forall\ x \in V$.

⇒ $T(E(v)) \subseteq E(v)$.

Hence the range of E is invariant under T.

(ii) Since E is a projection on V, so $V = R \oplus N$, where

$R = \{ v \in V : E(v) = v \}$ and $N = \ker E = \{ v \in V : E(v) = \mathbf{0} \}$.

Necessary condition: Suppose R and N both be invariant under T. Then

$T(R) \subset R$ and $T(N) \subset N$. …. (2)

Claim: We shall prove that $E\,T = T\,E$. Let $v \in V$ be arbitrary. Since $V = R \oplus N$, so $v = r + n$, where $r \in R$ and $n \in N$.

Also E is a projection, so $E(v) = r$. …. (3)

From equation (2), we have,

$T(r) \in R$ and $T(n) \in N$.

⇒ $E(T(r)) = T(r)$ and $E(T(n)) = \mathbf{0}$. …. (4)

Now $(ET)(v) = E(T(v)) = E(T(r + n)) = E(T(r) + T(n))$

$= E(T(r)) + E(T(n)) = T(r) + \mathbf{0}$, ; using (4)

$= T(r) = T(E(v))$. ; using (3)

⇒ $(E\,T)(v) = (T\,E)(v)\ \ \forall\ v \in V$.

Hence $E\,T = T\,E$.

Sufficient condition: Let E T = T E. Then E T E = E(E T) = E^2 T = E T. So by part (i) we conclude that range of E i.e., R is invariant under T.

Claim: Null space of E i.e. N is invariant under T i.e. T(N) $\subseteq$ N.

For any n $\in$ N, we have E(n) = **0**.

So E T = T E $\Rightarrow$ (E T) (n) = (T E) (n).

$\Rightarrow$ E(T(n)) = T(E(n)) = T(0) =**0**.

$\Rightarrow$ T(n) $\in$ N $\forall$ n $\in$ N.

$\Rightarrow$ N is invariant under T.

Example 24: Let V = $W_1 \oplus W_2$, where W_1 and W_2 are subspaces of V. If E_1 is the projection on W_1 along W_2 and E_2 is the projection on W_2 along W_1, then prove that,

(i) $E_1 + E_2 = I$. **(ii)** $E_1E_2 = E_2E_1 = \mathbf{0}$.

Solution: Given that V = $W_1 \oplus W_2$. So each v $\in$ V is uniquely expressible as $v = w_1 + w_2$; $w_1 \in W_1$, $w_2 \in W_2$.

(**i**) Since E_1 is the projection on W_1 along W_2 . So

$E_1(v) = w_1$.

Similarly, $E_2(v) = w_2$.

Now $(E_1 + E_2)(v) = E_1(v) + E_2(v) = w_1 + w_2 = v$.

$\Rightarrow$ $(E_1 + E_2)(v) = I(v)$ $\forall$ v $\in$ V.

Hence $E_1 + E_2 = I$.

(ii) Now $E_1E_2 = E_1 (I – E_1)$, by part (i)

$= E_1 – E_1^2$

$= E_1 – E_1$, as E_1 is projection $E_1E_2 = \mathbf{0}$.

Similarly, $E_2E_1 = \mathbf{0}$.

Example 25: If a diagonalizable operator has only eigenvalues 0 and 1, then prove that it is a projection.

Solution: Given that, linear operator, say T, is diagonalizable. Then there exists a basis $\beta = \{ v_1, \ldots, v_n\}$ of V such that, $[T]_\beta =$ diag $(c_1, \ldots, c_n)$, where $c_1, \ldots, c_n$ are eigen values of T. But 0 and 1 are only eigenvalues of T.

Suppose $c_1 = \ldots = c_m = 1$ and $c_{m+1} = \ldots = c_n = 0$. …. (1)

Let us take $v \in V$ as arbitrary. So

$$v = \alpha_1 v_1 + \ldots + \alpha_m v_m + \alpha_{m+1} v_{m+1} + \ldots + \alpha_n v_n \quad ; \qquad \alpha_i \in \mathbf{F}$$

$$T(v) = T(\alpha_1 v_1 + \ldots + \alpha_m v_m + \alpha_{m+1} v_{m+1} + \ldots + \alpha_n v_n)$$

$$= \alpha_1 T(v_1) + \ldots + \alpha_m T(v_m) + \alpha_{m+1} T(v_{m+1}) + \ldots + \alpha_n T(v_n)$$

$= \alpha_1 (c_1 v_1) + \ldots + \alpha_m (c_m v_m) + \alpha_{m+1} (c_{m+1} v_{m+1}) + \ldots + \alpha_n (c_n v_n)$, where $T(v_i) = c_i v_i$. Now using equation (1), we get

$T(v) = \alpha_1 v_1 + \ldots + \alpha_m v_m$

$\Rightarrow \quad T^2(v) = T(T(v)) = T(\alpha_1 v_1 + \ldots + \alpha_m v_m)$

$= T(\alpha_1 v_1 + \ldots + \alpha_m v_m) + \alpha_{m+1} (c_{m+1} v_{m+1}) + \ldots + \alpha_n (c_n v_n)$, where $c_{m+1} = \ldots = c_n = 0$.

$T^2(v) = T(\alpha_1 v_1 + \ldots + \alpha_m v_m) + \alpha_{m+1} T(v_{m+1}) + \ldots + \alpha_n T(v_n)$

$= T(\alpha_1 v_1 + \ldots + \alpha_m v_m + \alpha_{m+1} v_{m+1} + \ldots + \alpha_n v_n)$

$T^2(v) = T(v) \ \forall \ v \in V$.

$\Rightarrow \quad T^2 = T$.

$\Rightarrow$ T is a projection.

Example 26: Let T be a linear operator on a finite-dimensional vector space V(**F**). Let R be the range of T and let N be the null space of T. Prove that R and N are independent if and only if $V = R \oplus N$.

Solution: If part : Suppose V = R ⊕ N. It means V = R + N and R ∩ N = {**0**}.

Claim: We shall prove that R and N are independent.

Let r + n = **0** , where r ∈ R and n ∈ N . …. (1)

Then, T(r + n) = T(**0**).

⇒ T(r) + T(n) =**0**.

⇒ T(r) +**0**= **0**.

⇒ r ∈ N.

So, r ∈ R ∩ N = {**0**} ⇒ r =**0**.

So from equation (1), n = **0**.

⇒ r + n =**0** implies r = **0**, n = **0**.

⇒ R and N are independent subspaces of V.

Only if part: Let R and N be independent subspaces of V.

Claim: V = R ⊕ N. By rank-nullity theorem (**Sylvester's law**), we have

rank (T) + nullity (T) = dim V i.e. dim R + dim N = dim V and V = R + N. Every v ∈ V is expressible as v = x + y ; x ∈ R, y ∈ N. We shall prove **uniqueness** of this representation.

Suppose v = $x_1 + y_1$ also, where x_1 ∈ R, y_1 ∈ N

⇒ x + y = $x_1 + y_1$ or $(x - x_1) + (y - y_1) =$ **0**.

But R and N are independent subspaces.

⇒ $x - x_1 =$ **0** and $y - y_1 =$ **0.**

⇒ $x = x_1$ and $y = y_1$.

So v = x + y = $x_1 + y_1$.

⇒ representation is unique.

⇒ $V = R \oplus N$.

Projection on a subspace: Let $V = W_1 \oplus W_2$, where W_1 and W_2 are subspaces of a vector space V(**F**). Then each $v \in V$ is uniquely expressible as $v = w_1 + w_2$, where $w_1 \in W_1$ and $w_2 \in W_2$. The projection on W_1 (along W_2) is defined as a linear operator E on V such that E(v) = w_1 . Also two subspaces W_1 and W_2 of V are called independent if $W_1 \cap W_2 = \{\mathbf{0}\}$.

Example 27: If E_1 and E_2 are projections onto independent subspaces, then prove that $E_1 + E_2$ is also a projection .

Solution: Step I: Suppose E_1 and E_2 be projections onto independent subspaces W_1 and W_2, respectively. By definition , $V = W_1 \oplus W_1'$ and $V = W_2 \oplus W_2'$ $\forall v \in V$.

So, $E_1(v) = w_1$ where $v = w_1 + w_1'$; $w_1 \in W_1$, $w_1' \in W_1'$ and $E_2(v) = w_2$ where $v = w_2 + w_2'$; $w_2 \in W_2$, $w_2' \in W_2'$. Now suppose $v \in V$ be arbitrary.

Then $(E_1E_2)(v) = E_1(E_2(v)) = E_1(w_2)$. …. (1)

Now two cases may arise:

Case (i): If $w_2 = \mathbf{0}$, then $E_1(w_2) = E_1(\mathbf{0}) = \mathbf{0}$.

Case (ii): If $w_2 \neq 0$, then $w_2 \notin W_1$, because W_1 and W_2 are independent; consequently $W_1 \cap W_2 = \{\mathbf{0}\}$. Since $V = W_1 \oplus W_1'$ and $w_1 \notin W_1$, $\Rightarrow w_2 \in W_1'$.

So $E_1(w_2) = E_1(\mathbf{0} + w_2) = \mathbf{0}$. Hence from (1), we have $(E, E_2)(v) = \mathbf{0}$, $\forall v \in V$.

⇒ $E_1E_2 = \mathbf{0}$.

Similarly, we can prove that $E_2E_1 = \mathbf{0}$.

Step II: Finally, $(E_1 + E_2)^2 = (E_1 + E_2)(E_1 + E_2) = E_1(E_1 + E_2) + E_2(E_1 + E_2)$

$$= E_1E_1 + E_1E_2 + E_2E_1 + E_2E_2$$

$$= E_1^2 + 0 + 0 + E_2^2 = E_1 + E_2.$$

⇒ $E_1 + E_2$ is a projection .

Example 28: Let V be a vector space over **F** and $W_1, W_2, \ldots, W_k$ be subspaces of V. Then

$$V = W_1 \oplus W_2 \oplus \ldots \oplus W_k \text{ , if and only if there exist k linear operators}$$

$E_1, E_2, \ldots, E_k$ on V such that –

(i) Each E_i is a projection i.e. $E_i^2 = E_i$.
(ii) $E_i E_j = 0$ for $i \neq j$.
(iii) $I = E_1 + E_2 + \ldots + E_k$.
(iv) The range of E_i is W_i for $i = 1, 2, \ldots, k$.

Proof: Necessary Condition: Let $V = W_1 \oplus W_2 \oplus \ldots \oplus W_k$. Then each $x \in V$ is uniquely expressible as $x = x_1 + x_2 + \ldots + x_i + \ldots + x_k$; $x_i \in W_i$, $1 \leq i \leq k$.

Let us **define** a mapping $E_i : V \to V$ such that

$E_i(x) = x_i$. …. (1)

Claim: We shall prove that E_i is a linear operator on V $\forall$ i. Let x, y $\in$ V and α , $\beta \in$ **F**. So

$y = y_1 + y_2 + \ldots + y_i + \ldots + y_k$; $y_i \in W_i$, $1 \leq i \leq k$.

On the basis of equation (1), we have

$E_i(\alpha x + \beta y) = \alpha x_i + \beta y_i = \alpha E_i(x) + \beta E_i(y)$.

$\Rightarrow$ E_i is a linear operator on V ; $\forall$ i = 1, 2, …., k.

(i) Now $E_i^2(x) = E_i(E_i(x)) = E_i(x_i)$, using (1)

$= x_i$ using (1)

$= E_i(x)$. using (1)

$\Rightarrow$ $E_i^2(x) = E_i(x)$, $\forall$ $x \in V$.

$\Rightarrow$ E_i is a projection, $\forall$ i = 1, 2, …., k.

(ii) For $i \neq j$, we get

$(E_i\ E_j)\ (x) = E_i(E_j(x)) = E_i\ (x_j) = \mathbf{0} = \mathbf{0}\ (x).$

$\Rightarrow \quad E_i\ E_j = 0, \text{ for } i \neq j.$

(iii) For any $x \in V$, we see that

$(E_1 + E_2 + \ldots + E_k)\ (x) = E_1(x) + \ldots + E_k(x) = x_1 + \ldots + x_k = I(x).$

$\Rightarrow \quad E_1 + E_2 + \ldots + E_k = I.$

(iv) Range of $E_i = \{ E_i(x) : x \in V \} = \{ x_i : x_i \in W_i \} = W_i$.

Sufficient Condition: Here, we have given **(i) – (ii)** conditions.

Claim: We shall prove that $V = W_1 \oplus W_2 \oplus \ldots \oplus W_k$.

Let $x \in V$ be arbitrary.

Then from condition (**iii**), we have

$I(x) = (E_1 + E_2 + \ldots + E_k)\ (x).$

$\Rightarrow \quad x = E_1(x) + \ldots + E_k(x)$.

$\Rightarrow \quad x = x_1 + \ldots + x_k$.

So $V = W_1 + \ldots + W_k$.

Uniqueness: Let $x = z_1 + z_2 + \ldots + z_k$; $z_i \in W_i$ = Range E_i. Since $z_i \in$ Range E_i , so

$z_i = E_i(\ t_i) \quad ; \quad t_i \in V$. …. (3)

So $E_i(x) = E_i(z_1 + \ldots + z_k) = E_i(z_1) + \ldots + E_i(z_i) + \ldots + E_i(z_k).$

$= E_i(E_1(\ t_1)) + \ldots + E_i(E_i(\ t_i)) + \ldots + E_i(E_k(\ t_k))$, by (3)

$= (E_iE_1)\ (\ t_1) + \ldots + (E_iE_i)\ (\ t_i) + \ldots + (E_iE_k)\ (\ t_k).$

$= \mathbf{0} + \ldots + E_i^2\ (t_i)\ + \mathbf{0} \ldots + \mathbf{0}.$

$x_i = E_i(\ z_i) = z_i \quad \forall\ i = 1, 2, \ldots, k.$

This proves uniqueness of (2). Hence

$$V = W_1 \oplus W_2 \oplus \ldots.. \oplus W_k.$$

Theorem 8: Let T be a linear operator on a vector space V and $V = W_1 \oplus W_2 \oplus \ldots.. \oplus W_k$.

Define $E_i(v) = E_i(x_1 + \ldots.. + x_k) = x_i \in W_i$. Then –

(i) Each E_i is a projection on V.

(ii) $E_iE_j = 0$, for $i \neq j$.

(iii) $I = E_1 + \ldots.. + E_k$.

A **necessary and sufficient** condition that each W_i is invariant under T is that $T E_i = E_iT$ for all i = 1, 2, ….,k.

Proof: Conditions **(i), (ii)** and **(iii)** can be verified from previous theorem.

Given $V = W_1 \oplus W_2 \oplus \ldots.. \oplus W_k$. …. (1)

Necessary condition: Suppose each W_i be invariant under T, i = 1, 2, ….,k.

So $T(W_i) \subseteq W_i$, for i = 1, 2, ….,k. …. (2)

Since $I = E_1 + \ldots.. + E_k$, so

$$I(v) = (E_1 + \ldots.. + E_k)(v) = E_1(v) + \ldots. + E_k(v)$$

$$v = x_1 + \ldots.. + x_k. \qquad \text{by (1)}$$

Now, $T(v) = T(x_1 + \ldots.. + x_k) = T(x_1) + \ldots.. + T(x_k)$

$$T(v) = y_1 + \ldots. + y_i + \ldots. + y_k \text{ ; where } y_i = T(x_i) \in W_i \text{ ; by (2)}$$

So $E_i(T(v)) = E_i(y_1 + \ldots. + y_i + \ldots. + y_k)$

$$= E_i(y_1) + \ldots. + E_i(y_i) + \ldots. + E_i(y_k)$$

$$= \mathbf{0} + \ldots. + y_i + \mathbf{0} + \ldots. + \mathbf{0} = y_i$$

$$= T(x_i)$$

$$= T E_i(v). \qquad \text{by (1)}$$

⇒ $(E_i\, T)\,(v) = (T\, E_i)\,(v) \quad \forall\, v \in V$.

⇒ $E_iT = T\, E_i$ for all i = 1, 2, …., k.

Sufficient Condition: Let $T\, E_i = E_iT$, for all i = 1, 2, …., k.

Claim: We shall prove that each W_i is invariant under T. Suppose $w_i \in W_i$ be arbitrary. By given definition in statement, $E_i(w_i) = w_i$.

So $T(w_i) = T(E_i(w_i)) = (T\, E_i)\,(w_i) = (E_iT)\,(w_i)$

$= E_i\,(T(w_i)) \in$ Range $E_i = W_i$.

Thus $T(w_i) \in W_i \;\; \forall\;\; w_i \in W_i$.

Hence W_i is invariant under T, for I =1, 2, …., k.

Theorem 9: If T is a diagonalizable operator on a finite-dimensional vector space V and if c_1, …., c_k are distinct eigenvalues of T, then there exist linear operators E_1, …., E_k on V such that:

(i) $T = c_1E_1 + \ldots.. + c_kE_k$.
(ii) $I = E_1 + \ldots. + E_k$.
(iii) $E_i\, E_j = 0$ for $i \neq j$.
(iv) $E_i^2 = E_i$, for each i.
(v) The range of E_i is the eigenspace of T associated with the eigenvalue c_i of T.

Proof: Let W_i be the eigenspace of T corresponding to the eigenvalue c_i ; i = 1, 2, …., k. We have $W_i = \{v \in T(v) = c_iv\}$ ….(1)

Given that T is a diagonalizable operator. So

$$V = W_1 \oplus W_2 \oplus \ldots.. \oplus W_k.$$

Let us **define** $E_i : V \rightarrow V$ such that

$$E_i\,(v) = E_i\,(x_1 + \ldots.. + x_k) = x_i.$$

Claim: We have already proved conditions **(ii) – (v)**. Now we prove condition (**i**). Let $v \in V$ be arbitrary.

$\therefore \quad I = E_1 + \ldots + E_k \Rightarrow I(v) = (E_1 + \ldots + E_k)(v).$

$\Rightarrow \quad v = E_1(v) + \ldots + E_k(v).$

$\Rightarrow \quad v = y_1 + \ldots + y_k$ where $y_i = E_i(v) \in$ Range $E_i = W_i$.

But $y_i \in W_i$, so by equation (1), we have

$T(y_i) = c_i y_i$, for each i. …. (2)

Now $T(v) = T(y_1 + \ldots + y_k)$

$= T(y_1) + \ldots + T(y_k)$

$= c_1 y_1 + \ldots + c_k y_k$, using (2)

$= c_1 E_1(v) + \ldots + c_k E_k(v)$, as $y_i = E_i(v)$

$T(v) = (c_1 E_1 + \ldots + c_k E_k)(v) \quad \forall \ v \in V.$

$\Rightarrow \quad T = c_1 E_1 + \ldots + c_k E_k.$

Note: Converse of above theorem is **also true**. We shall prove it in next theorem.

Theorem 10: Let T be a linear operator on a finite-dimensional vector space V(**F**). Let $c_1, \ldots, c_k$ be distinct scalars and $E_1, \ldots, E_k$ be non-zero linear operators on V such that:

(i) $T = c_1 E_1 + \ldots + c_k E_k$.
(ii) $I = E_1 + \ldots + E_k$.
(iii) $E_i E_j = 0$ for $i \neq j$.

Then T is diagonalizable with $c_1, \ldots, c_k$ as its eigenvalues. Further $E_i^2 = E_i$, for each i and Range E_i = eigenspace of T associated with c_i.

Proof: From **(ii)** and **(iii)** conditions, we have

$E_i = E_i I = E_i (E_1 + \ldots + E_k) = E_i E_1 + \ldots + E_i^2 + \ldots + E_i E_k$

$= \mathbf{0} + \ldots + E_i^2 + \ldots + \mathbf{0}.$

$\Rightarrow \quad E_i = E_i^2.$ ------(iv)

This is condition (**iv**).

Now, $T E_i = (c_1E_1 + \ldots + c_iE_i + \ldots + c_kE_k) E_i$

$= c_i E_i^2 = c_iE_i$

$\Rightarrow$ $(T - c_iI) E_i = 0$, for each i.

As $E_i \neq 0$, there exists some $v_i \in V$ such that $E_i (v_i) \neq 0$.

So $(T - c_iI) E_i(v) = 0$ for each i.

$\Rightarrow$ $T(E_i(v)) = c_i(E_i(v))$; $E_i (v_i) \neq 0$.

$\Rightarrow$ c_i is an eigenvalue of T for each i.

If c is any scalar, then

$(T - cI) = (c_1E_1 + \ldots + c_kE_k) - c (E_1 + \ldots + E_k)$

$(T - cI) = (c_1 - c) E_1 + \ldots + (c_k - c) E_k$. …. (1)

If c is an eigen value of T, then there exists some $\mathbf{0} \neq v \in V$, such that

$Tv = cv$ or $(T - cI) (v) = \mathbf{0}$.

$\Rightarrow$ $(c_1 - c) E_1(v) + \ldots + (c_k - c) E_k(v) = \mathbf{0}$, using (1)

So $E_j [(c_1 - c) E_1(v) + \ldots + (c_k - c) E_k(v)] = E_j (\mathbf{0}) = \mathbf{0}$.

$\Rightarrow$ $(c_1 - c) E_j E_1(v) + \ldots + (c_k - c) E_j E_k(v) = \mathbf{0}$.

$\Rightarrow$ $(c_j - c) E_j E_j (v) = \mathbf{0}$, as $E_i E_j = \mathbf{0}$, for $i \neq j$.

$\Rightarrow$ $(c_j - c) E_j^2 (v) = \mathbf{0}$ or $(c_j - c) E_j (v) = \mathbf{0}$; $j = 1, 2, \ldots, k$. ….(2)

If we take $E_j (v) = 0$ for all $j = 1, 2, \ldots, k$, then

$I = E_1 + \ldots + E_k$ $\Rightarrow$ $I(v) = E_1(v) + \ldots + E_k(v)$

$\Rightarrow$ $v = \mathbf{0} + \ldots + \mathbf{0} = \mathbf{0}$. **Contradiction !**

So $E_j(v) \neq 0$ for some j. Using equation (2), we get $(c_j - c) = 0$, for some j.

$\Rightarrow \quad c_j = c$, for some j.

Hence $c_1, \ldots, c_k$ are the only eigenvalues of T.

Let W_i = Range of $E_i = E_i(v)$, for i = 1, 2, …., k. ------(v)

Then $I = E_1 + \ldots + E_k$.

$\Rightarrow \quad I(v) = E_1(v) + \ldots + E_k(v) \quad ; \forall\, v \in V$.

$\Rightarrow \quad v \in W_1 + \ldots + W_k \quad \forall\, v \in V$.

$\Rightarrow \quad V = W_1 + \ldots + W_k$.

Also conditions **(ii) – (v)** are satisfied. So we have,

$$V = W_1 \oplus W_2 \oplus \ldots \oplus W_k$$

$\Rightarrow \quad \dim V = \dim W_1 + \ldots + \dim W_k$. ….(3)

Step II: Now we shall show that –

W_i = eigenspace of T corresponding to c_i , $\forall$ i = 1, 2, …., k.

Then from previous knowledge, T is diagonalizable.

Let W_{c_i} denotes the eigenspace of T corresponding to the eigenvalue c_i .

Claim: We shall prove that $W_{c_i} = W_i$.

Let $x \in W_{c_i}$ be arbitrary. Then, $T(x) = c_i x \;;\; 1 \leq i \leq k$.

$\Rightarrow \quad (c_1E_1 + \ldots + c_kE_k)(x) = c_i\, I(x)$; using condition (i)

$\Rightarrow \quad c_1E_1(x) + \ldots + c_kE_k(x) = c_i\,[E_1(x) + \ldots + E_k(x)]$

$\Rightarrow \quad (c_1 - c_i)E_1(x) + \ldots + (c_k - c_i)E_k(x) = \mathbf{0}$.

$\Rightarrow \quad (c_j - c_i)E_j(x) = \mathbf{0}$, for all j =1, 2, …., k.

But $c_j - c_i \neq 0$ for $j \neq i$.

So $E_j(x) = \mathbf{0}$, $j \neq i$ . ....(4)

Since $I = E_1 + \ldots + E_k$, so

$I(x) = E_1(x) + \ldots + E_k(x)$.

$\Rightarrow$ $x = E_i(x)$; using (4)

$\Rightarrow$ $x \in$ Range of $E_i = W_i$.

So $W_{c_i} \subset W_i$ . ....(5)

Again, let $\mathbf{0} \neq x \in W_i = R(E_i)$.

Then $x = E_i(y_i)$.

As **proved above**, we have $T E_i(y_i) = c_i E_i(y)$.

$\Rightarrow$ $T(x) = c_i x$, where $x = E_i(y_i) \neq \mathbf{0}$.

$\Rightarrow$ $x \in W_{c_i}$

So $W_i \subset W_{c_i}$. ....(6)

$\Rightarrow$ $W_i = W_{c_i}$ = eigenspace of T corresponding to c_i ; $i = 1, 2, \ldots, k$.

$\Rightarrow$ T is diagonalizable.

Theorem 11: If T is a linear operator on a finite-dimensional vector space V(**F**) and minimal polynomial p(x) of T is a product of distinct linear factors i.e. $p(x) = (x - c_1)(x - c_2) \ldots (x - c_k)$, where $c_1, \ldots, c_k$ are distinct, then T is diagonalizable.

Proof: Let us define k polynomials as

$$p_1(x) = \frac{(x - c_2)(x - c_3) \ldots (x - c_k)}{(c_1 - c_2)(c_1 - c_3) \ldots (c_1 - c_k)} ,$$

$$p_2(x) = \frac{(x - c_1)(x - c_3) \ldots (x - c_k)}{(c_2 - c_1)(c_2 - c_3) \ldots (c_2 - c_k)} ,$$

…. …. …. …. …. …. …. …. ….

$$p_k(x) = \frac{(x-c_1)(x-c_2).....(x-c_{k-1})}{(c_k-c_1)(c_k-c_2).....(c_k-c_{k-1})} ,$$

We observe that , $p_1(c_1) = p_2(c_2) = …. = p_k(c_k) = 1$ and for other values, these are zero i.e.

$$p_i(c_j) = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases} \qquad ….(1)$$

Given that $p(x) = (x - c_1)(x - c_2) …..(x - c_k)$,

So deg $p_i(x) <$ deg $p(x)$; $i = 1, 2, …., k$.

Step I: Let W be the vector space of all polynomials over **F** of degree $\leq$ k. We shall prove that $p_1(x), p_2(x), …., p_k(x) \in W$ are linearly independent.

Let $\alpha_1 p_1(x) + \alpha_2 p_2(x) + …. + \alpha_k p_k(x) = 0$; $\alpha_i \in$ **F**.

$\Rightarrow \alpha_1 p_1(c_i) + \alpha_2 p_2(c_i) + …. + \alpha_k p_k(c_i) = 0$.

$\Rightarrow \alpha_i = 0 \ \forall i = 1, 2, …., k.$ using (1)

Since dim W = k. So $\{ p_1(x), …., p_k(x)\}$ is a basis of W.

As $1 \in W$. So $\exists \ \lambda_i \in$ **F** such that,

$$1 = \lambda_1 p_1(x) + …. + \lambda_k p_k(x).$$

Putting $x = c_i$ on both sides, we get

$1 = \lambda_1 p_1(c_i) + …. + \lambda_k p_k(c_i) = \lambda_i$; using (1)

$\Rightarrow \lambda_i = 1$ for $i = 1, 2, …., k$.

$\Rightarrow 1 = p_1(x) + …. + p_k(x).$ ….(2)

Since $x \in W$, so $x = \beta_1 p_1(x) + …. + \beta_k p_k(x)$; $\beta_i \in$ **F**.

Putting $x = c_i$, we get

$c_i = \beta_1 p_1(c_i) + …. + \beta_k p_k(c_i) = \beta_i$, using (1)

$\Rightarrow$ $\beta_i = c_i \quad \forall\ i = 1, 2, \ldots, k.$

So $x = c_1 p_1(x) + \ldots + c_k\, p_k(x).$ ….(3)

Step II: Let $p_i(T) = E_i$ for $i = 1, 2, \ldots, k$. If possible, let $E_j = 0$ for some j, then $p_j(T) = 0$ and $\deg p_j(x) < \deg p(x)$, but this is a **contradiction** to the minimality of p(x). So $E_j \neq 0$ for all j.

Now putting x = T in equations (2) and (3), we get

$$I = p_1(T) + \ldots + p_k(T) = E_1 + \ldots + E_k \qquad \ldots(4)$$

and $T = c_1 p_1(T) + \ldots + c_k\, p_k(T) = c_1 E_1 + \ldots + c_k\, E_k$ ….(5)

Since p(x) is the minimal polynomial of T, so p(T) = 0. ….(6)

Here we **remember** that p(x) divides $p_1(x)\, p_2(x)$, etc or in general p(x) divides

$$p_i(x)\, p_j(x), \text{ for all } i \neq j$$

So by **Division Algorithm**, $\exists\ q(x) \in \mathbf{F}[x]$, such that

$$p_i(x)\, p_j(x) = p(x)\, q(x).$$

$\Rightarrow$ $p_i(T)\, p_j(T) = p(T)\, q(T) = 0$ for all $i \neq j$; using (6)

$\Rightarrow$ $E_i\, E_j = 0$ for all $i \neq j$. ….(7)

If we use equations **(4), (5), (7)**, then from the knowledge of previous theorems, we **conclude** that –

T is diagonalizable with $c_1, \ldots, c_k$ as its eigenvalues.

Theorem 12: Let T be a linear operator on a finite-dimensional vector space V(**F**). Then T is

diagonalizable **if and only if** the minimal polynomial for T has he following form

$$p(x) = (x - c_1)(x - c_2) \ldots(x - c_k),$$

where $c_1, \ldots, c_k$ are distinct elements of **F**.

Proof: Necessary Condition: Let T be diagonalizable . Let $c_1, \ldots, c_k$ be distinct eigenvalues of T. Since we know that any eigenvalue of T is a root of the minimal polynomial for T. So each of the polynomials $x - c_1, x - c_2, \ldots, x - c_k$, is a factor of the minimal polynomial for T. Hence the polynomial $p(x) = (x - c_1) \ldots (x - c_k)$ will be the minimal polynomial for T, if p(T) = 0. Let v be any eigenvector of T. Then $(T - c_1I) \ldots (T - c_k I)(v) = \mathbf{0}$, for all eigenvectors v of T. Since T is diagonalizable, there exists a basis $\beta = \{ v_1, \ldots, v_k\}$ consisting of eigenvectors of T. As shown above,

$$(T - c_1I) \ldots (T - c_k I)(v_i) = \mathbf{0}, \text{ for } i = 1, 2, \ldots, k. \qquad \ldots(1)$$

Let $x \in V$ be arbitrary. Then

$$x = \alpha_1 v_1 + \ldots + \alpha_k v_k \; ; \quad \alpha_i \in \mathbf{F}.$$

So , $(T - c_1I)(T - c_2I) \ldots (T - c_k I)x = \mathbf{0}$, for each $x \in V$.

$\Rightarrow \quad (T - c_1I) \ldots (T - c_k I) = \mathbf{0}$, on V.

$\Rightarrow \quad p(T) = \mathbf{0}$.

Hence $p(x) = (x - c_1) \ldots (x - c_k)$ is the minimal polynomial for T.

Sufficient Condition: It has been proved in the previous theorem.

Check your progress

Problem 1: Find the minimal polynomial for the matrix $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$.

Problem 2: Find the minimal polynomial for the matrix $A = \begin{bmatrix} 3 & 1 & -1 \\ 2 & 2 & -1 \\ 2 & 2 & 0 \end{bmatrix}$.

Problem 3: Give an example of a matrix whose characteristic and minimal polynomials are same.

Problem 4: Give an example of a matrix whose characteristic and minimal polynomials are **NOT** same.

9.7 *SUMMARY*

After the study of this chapter, we have learnt about the difference between the characteristic polynomial and minimal polynomial. Then we studied about invariant subspaces of a vector space. After that we discussed direct sum decomposition of a vector space. At last we learned about projection on a vector space and a lot of exercises to have a good command over various concepts.

9.8 *GLOSSARY*

- **Invariant Subspaces:** Let T be a linear operator on a vector space V(**F**). A subspace of V(**F**) is said to be invariant under T (or W is T-invariant) if $T(W) \subseteq W$.
- **Independent subspaces:** The subspace $W_1, \ldots, W_k$ of V(**F**) are called independent if → for $x_i \in W_i$; i = 1, 2, …., k, we have $x_1 + x_2 + \ldots + x_k = \mathbf{0}$,

  $\Rightarrow \quad x_i = \mathbf{0}, \ \forall \ i = 1, 2, \ldots, k.$

- **Projection on a Vector Space:** A linear operator E on a vector space V is called a projection if $E^2 = E$. It means E is independent.

9.9 *REFERENCES*

- S Kumaresan; Linear Algebra-A Geometric Approach; PHI-2016.
- K Hoffman & Ray Kunze; Linear Algebra; PHI-1971.
- P Halmos; Finite-Dimensional Vector Spaces;
- G Strang; Introduction To Linear Algebra;

9.10 *SUGGESTED READING*

- NPTEL videos.
- Schaum series.
- A R Vashishtha, Krishna Prakashan; Meerut.
- Graduate Text In Mathematics, Springer.

9.11 *TERMINAL QUESTION*

Long answer type question:

1: If T is any linear operator on a vector space V, then prove that ker (T) and Range (T) are invariant subspaces of V.

2: **Example:** Let V be the vector space of all polynomials of degree less than or equal to six. Let W be the subspace of V spanned by { 1, x^2, x^4 , x^6 }. Let D be the differential operator on V i.e. D f(x) = $\frac{d}{dx}$ (f(x)).

(v) Show that W is not invariant under D.

(vi) Let T = D^2 , where D^2 f(x) = $\frac{d^2}{dx^2}$ (f(x)). Show that W is invariant under T.

(vii) Find the matrix of T_W in a suitable basis of W, where T_W is the restriction of T on W.

(viii) Find the matrix of T in a suitable basis of V.

3: Let V be a finite-dimensional vector space over **F**. Let W_1, …., W_k be subspaces of V and W = W_1 + …. + W_k . Then the following conditions are equivalent:

(iv) W_1, W_2 …., W_k are independent.

(v) $W_j \cap (W_1 + \ldots + W_{j-1}) = \{\mathbf{0}\}$ for all j ; $2 \leq j \leq k$.

(vi) If β_i is an ordered basis of W_i, $1 \leq i \leq k$, then $\beta = \{\beta_1, \ldots, \beta_k\}$ is an ordered basis of W.

4: If E is a projection on a vector space V(F), then prove that V = R ⊕ N, where R is the range space of E and N is the null space of E.

Short answer type question:

1: Let T be a linear operator on a finite-dimensional vector space V(**F**). Then prove that the characteristic and the minimal polynomials for T have the same roots, except for multiplicities.

2: Prove that the minimal polynomial of a linear operator T divides its characteristic polynomial.

3: Prove that the minimal polynomial of a linear operator T is a divisor of every polynomial that annihilates T.

4: If a diagonalizable operator has only eigenvalues 0 and 1, then prove that it is a projection.

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

Prove that the characteristic polynomial and minimal polynomial are identical for A.

6: Let T be a linear operator on R^2 defined by $T(x, y) = (x + y, x + y)$.

Find the minimal polynomial for T.

7: Show that similar matrices have the same minimal polynomial.

8: Find the characteristic and minimal polynomials for the matrix $A = \begin{bmatrix} 5 & -6 & -6 \\ -1 & 4 & 2 \\ 3 & -6 & -4 \end{bmatrix}$.

9.12 *ANSWERS*

Answers of check your progress:

1. $x(x - 1)$

2. $(x - 1)(x - 2)^2$.

Answer of long question

2(iii) $\begin{bmatrix} 0 & 2 & 0 & 0 \\ 0 & 0 & 12 & 0 \\ 0 & 0 & 0 & 30 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

(iv) $\begin{bmatrix} 0 & 2 & 0 & 0 & \vdots & 0 & 0 & 0 \\ 0 & 0 & 12 & 0 & \vdots & 0 & 0 & 0 \\ 0 & 0 & 0 & 30 & \vdots & 0 & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \vdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & \vdots & 0 & 6 & 0 \\ 0 & 0 & 0 & 0 & \vdots & 0 & 0 & 20 \\ 0 & 0 & 0 & 0 & \vdots & 0 & 0 & 0 \end{bmatrix}$

Answers of terminal questions:

5: $x (x - 2)$

8: $(x - 1) (x - 2)$.

UNIT-10: EIGEN VALUES AND EIGEN VECTORS

CONTENTS

10.1 *INTRODUCTION*

In the last unit, we focused on Inner Product Spaces. Now again, we emphasize on Vector Spaces. After the study of Linear Transformation, we have studied some properties of a linear operator. Here, we shall elaborate these concepts and matrices help us in a great deal. Basis of a matrix and its role to understand eigen values and eigen vectors will be discussed in detail. Besides this, diagonalisation process and required conditions will be discussed thoroughly.

10.2 *OBJECTIVES*

After the study of this chapter, learner shall understand:

- Linear operators and their properties.
- For finite-dimensional vector spaces, **T** can be represented as a matrix.
- How can we convert square matrix into diagonal matrix?
- Role of basis of a linear transformation in diagonalisation.

10.3 *BASICS OF LINEAR OPERATORS*

In this section, we shall discuss linear operators (T) on a finite-dimensional vector space V(**F**). We know that a linear operator T on a vector space V(**F**) is a mapping T: V → V, such that

$$T(\alpha x + \beta y) = \alpha\, T(x) + \beta\, T(y) \quad \forall\; x, y \in V \text{ and } \alpha, \beta \in \mathbf{F}.$$

We have already studied following **important properties** of such a linear operator T:

(i) T is non-singular (i.e. one to one) if and only if ker(T) = {0}.

(ii) T is invertible ⟺ T is non-singular ⟺ T is onto.

(iii) T is singular ⟺ ker T ≠ {0}

10.4 *EIGEN VALUES & EIGEN VECTORS*

Now, we shall define eigen value and eigen vectors of T as:

Eigen Values of T: Let T be linear operator on a vector space V(**F**). A scalar $\alpha \in \mathbf{F}$ is called an eigen value or characteristic value of T, if there exists some V ≠ 0, v ∈ V such that, $T(v) = \alpha v$.

Eigen Vector: If α is an eigen value of T, then v ∈ V such that $T(v) = \alpha v$ is called an eigen vector or **characteristic vector** belonging to α.

Eigen space: The set of all eigenvectors of T belonging to an eigenvalue α is called an eigenspace of T, belonging to α. It is represented as W_α. Hence

$$W_\alpha = \{ v \in V : T(v) = \alpha v \}.$$

Example 1: Let T : $\mathbf{R}^2 \to \mathbf{R}^2$ be a linear operator defined be T(x, y) = (2x + y, x + 2y). By trial and error method, we find one eigen value of T and corresponding eigen vector.

We observe that $T(1, 1) = (3, 3) = 3(1, 1)$

Or $T(2, 2) = (6, 6) = 3(2, 2)$

Here 3 is an eigenvalue of T and (1, 1) , (2, 2) ∈ $\mathbf{R}^2$ are corresponding eigenvectors.

Also T(3, -3) = (3, -3) = 1(3, -3)

So here 1 is eigenvalue of T and (3, -3) ∈ $\mathbf{R}^2$ is corresponding eigenvector.

Example 2: Let T : $\mathbf{R}^3 \to \mathbf{R}^3$ be a linear operator whose matrix with respect to the standard basis { e_1 = (1, 0, 0), e_2 = (0, 1, 0), e_3 = (0, 0, 1) } is A = $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

So, $T(e_1) = e_1 = 1e_1$

$T(e_2) = e_2 = 1e_2$

$T(e_3) = 0 = 0\, e_3.$

We observe that 1, 1 and 0 are eigenvalues of T and corresponding eigenvectors are e_1, e_2 and e_3 respectively.

Note: Now we discuss the eigenspace W_1 corresponding to eigenvalue 1. So $W_1 = \{ v \in \mathbf{R}^3 : T(v) = 1.v \}$. Let v ∈ $\mathbf{R}^3$. then there exist α, β, γ ∈ **R** such that

$$v = \alpha e_1 + \beta e_2 + \gamma e_3$$

So $v \in W_1$ iff $T(\alpha e_1 + \beta e_2 + \gamma e_3) = 1\,(\alpha e_1 + \beta e_2 + \gamma e_3)$

iff $\alpha\, T(e_1) + \beta\, T(e_2) + \gamma\, T(e_3) = \alpha e_1 + \beta e_2 + \gamma e_3$

iff $\alpha e_1 + \beta e_2 + \gamma.0\, e_3 = \alpha e_1 + \beta e_2 + \gamma e_3$

iff $\gamma e_3 = 0$ or $\gamma = 0$

So $W_1 = \{ \alpha e_1 + \beta e_2 : \alpha , \beta \in \mathbf{R} \}$. In the same way, we can show that the eigenspace W_0 , corresponding to eigenvalue '0' is

$$W_0 = \{ \gamma e_3 : \gamma \in \mathbf{R} \}$$

Theorem: Let T be a linear operator on a vector space V(**F**).

(i) If $0 \neq v \in V$ is an eigenvector of T, then α ∈ **F** satisfying T(v) = αv is **unique**.

(ii) The eigenspace W_α corresponding to an eigen value $\alpha \in \mathbf{F}$ is a subspace of V(**F**).

(iii) $W_\alpha = \ker(T - \alpha I)$.

Proof: (i) As we know, for uniqueness; we always consider two values and show that both are equal *i.e*. value is unique. Suppose, if possible, there exist α , $\beta \in \mathbf{F}$ such that $T(v) = \alpha v$ and $T(v) = \beta v$

$$\Rightarrow \quad \alpha\beta = \beta v \quad \text{or} \quad (\alpha - \beta)v = 0$$

But $v \neq 0$, so

$$\alpha - \beta = 0 \quad \text{or} \quad \alpha = \beta$$

Hence α is unique.

(ii) We know that $W_\alpha = \{ v \in V : T(v) = \alpha v \}$

Claim: W_α is a subspace of V(**F**). As $T(0) = 0 \Rightarrow T(0) = \alpha 0$. So $0 \in W_\alpha$ i.e. W_α is non-empty. Let $v_1 , v_2 \in W_\alpha$ and a, b $\in$ **F**. then

$$T(v_1) = \alpha v_1 \quad \text{and} \quad T(v_2) = \alpha v_2$$

Now, $T(av_1 + bv_2) = aT(v_1) + b\,T(v_2)$; as T is linear

$$= a(\alpha v_1) + b(\alpha v_2)$$

$$T(av_1 + bv_2) = a(\alpha v_1) + b(\alpha v_2)$$

So, $av_1 + bv_2$ is an eigenvector, corresponding to eigenvalue α.

Hence $av_1 + bv_2 \in W_\alpha$, $\forall\, v_1 , v_2 \in W_\alpha$; a, b $\in$ **F.**

Hence W_α is a subspace of V(**F**).

(iii) By definition,

$$W_\alpha = \{ v \in V : T(v) = \alpha v \}$$

So $W_\alpha = \{ v \in V : T(v) = \alpha Iv \}$, where I is identity operator

$= \{ v \in V : T(v) = (\alpha I) v \}$

$= \{ v \in V : (T - \alpha I) v = 0 \}$

Hence $W_\alpha = \ker (T - \alpha I)$.

Theorem: Let T be a linear operator on a finite-dimensional vector space V(**F**). Then $\alpha \in \mathbf{F}$ is an eigenvalue of T if and only if $T - \alpha I$ is singular.

Proof: Necessary Condition: Let α be an eigenvalue of T. Then there exists some $0 \neq v \in V$, Such that, $T(v) = \alpha v$

$\Rightarrow$ $T(v) = \alpha I(v)$ where I is identity operator.

$\Rightarrow$ $T(v) = (\alpha I)(v)$

$\Rightarrow$ $(T - \alpha I)(v) = \mathbf{0}$, where $v \neq \mathbf{0}$.

So $v \in \ker (T - \alpha I)$. We already know that $\mathbf{0} \in \ker (T - \alpha I)$. So, $\ker (T - \alpha I) \neq \{\mathbf{0}\}$.

Hence $T - \alpha I$ is singular.

Sufficient condition: Let $T - \alpha I$ be singular operator .

$\Rightarrow \ker (T - \alpha I) \neq \{\mathbf{0}\}$,

$\Rightarrow$ there exists some $\mathbf{0} \neq v \in V$, such that $(T - \alpha I)(v) = \mathbf{0}$.

$\Rightarrow T(v) - \alpha I(v) = \mathbf{0}$.

$T(v) = \alpha v$, where $I(v) = v$.

So, α is an eigenvalue of T.

Note: (1) If T is singular, then '**0**' is always an eigenvalue of T. As $T - \mathbf{0}I = T$, it can be obviously observed.

(2) Till now, we have understood that if T is a linear operator on a finite-dimensional vector space, then the following statements are equivalent:

(i) α is an eigenvalue of T.

(ii) The operator T – α I is singular or non-invertible.

(iii) det (T – α I) = 0.

Characteristic values and Characteristic polynomial of a matrix:

Suppose T be a linear operator on a finite dimensional (say dim V = n) vector space V(**F**). Let β be an ordered basis for V and let A be the matrix of T with respect to the basis β i.e. A = $[T]_\beta$. For any scalar $\alpha \in$ **F**, we have

$$[T - \alpha I]_\beta = [T]_\beta - \alpha [I]_\beta$$

$$= A - \alpha I, \text{ where I is } n \times n \text{ unit matrix.}$$

So det(T – α I) = det $[T - \alpha I]_\beta$ = det (A – α I). Hence α is a characteristic value of T **if and only if** det (A – α I) = 0.

Note: From above discussion, we **conclude** that –

(i) Let A = $[a_{ij}]_{n \times n}$; $a_{ij} \in$ **F**. A scalar $\alpha \in$ **F** is called an **eigen value** of A if det(A – αI) = 0.

(ii) Let A = $[a_{ij}]_{n \times n}$; $a_{ij} \in$ **F**. Then the polynomial f(x) = det (A – α I) is called the **characteristic polynomial** of the matrix A.

The equation f(x) = 0 is called the **characteristic equation** of the matrix A. Here we observe that $\alpha \in$ **F** is an eigen value of the matrix A if and only if f(α) = 0.

Similar Matrices: Let A = $[a_{ij}]_{n \times n}$ and B = $[b_{ij}]_{n \times n}$ and P = $[c_{ij}]_{n \times n}$ where a_{ij}, b_{ij}, $c_{ij} \in$ **F**. Then A and B are called similar matrices if, there exist a matrix P such that

A = P^{-1} B P , where P is non-singular matrix.

You might have studied that **similarity** of matrices is an **equivalence relation** i.e. it is reflexive, symmetric and transitive.

Theorem: Similar matrices have the same characteristic polynomial and hence the same characteristic values.

Proof: Let us consider two square matrices A and B of $n \times n$ order. Then A and B are similar i.e. there exists an non-singular matrix P such that

$$B = P^{-1} A P.$$

So, $B - xI = P^{-1} A P - xI = P^{-1} A P - x P^{-1} I P$

$= P^{-1} (A - xI) P.$

So, $\det (B - xI) = \det (P^{-1} (A - xI) P)$

$$= \frac{1}{\det P} \det (A - xI) \det P$$

$\det (B - xI) = \det (A - xI).$

$\Rightarrow$ A and B have the same characteristic polynomials and consequently same eigenvalues.

Note: (1) You have studied in the chapter '**Linear Transformation**' that, if T be linear operator on an n-dimensional vector space. If β , β' are two ordered bases of V such that $A = [T]_\beta$ and $B = [T]_{\beta'}$, then there exists a non-singular matrix P (over **F**) such that $B = P^{-1} A P$.

(2) Let T be a linear operator on a finite-dimensional vector space V(**F**). then the characteristic polynomial of T is $\det(A - xI)$, where A is the matrix of T in any ordered basis for V.

(3) If T is a linear operator on an n-dimensional vector space and if $A = [T]_\beta$ with respect to an ordered basis β for V, then A is an $n \times n$ matrix and so $\det (A - xI)$ is a polynomial of degree n. Hence T **cannot** have more than n distinct eigenvalues.

(4) The eigenvalues of a linear operator defined on V(F) **may not** belong to **F**. For example, let T be a linear operator an R^2(**R**), whose matrix with respect to the standard ordered basis is

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

The characteristic polynomial of A is $\det (A - xI) = 0$

$$\Rightarrow \quad x^2 + 1 = 0$$

This equation has no roots in **R** (though, its roots $x = \pm i \in$ **C**).

Cayley-Hamilton Theorem for a linear operator: Every linear operator T on an n-dimensional vector space V(**F**) satisfies its characteristic equation f(x) = 0, i.e. f(T) = 0.

Proof: Let A be the matrix of T with respect to any basis β of V. So, A = $[T]_\beta$

Hence for matrices, Cayley-Hamilton theorem states that every square matrix satisfies its characteristic equation. Hence if f(x) = det(A – xI) = $a_0 + a_1x + a_2x^2 + \ldots.. + a_n x^n = 0$, is the characteristic equation of A, then

$$f(A) = a_0I + a_1A + a_2A^2 + \ldots.. + a_n A^n = 0$$

$$\Rightarrow \quad a_0[I]_\beta + a_1[T]_\beta + a_2 [T^2]_\beta + \ldots. + a_n [T^n]_\beta = [0]_\beta$$

$$\Rightarrow \quad [f(T)]_\beta = [0]_\beta$$

Hence $f(T) = a_0 I + a_1T + a_2 T^2 + \ldots. + a_n T^n = 0$

Example 1: Find the eigen values, eigen vectors and eigen spaces of A = $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

Solution: Step-I: Characteristic equation of A is |A – xI| = 0

$$\Rightarrow \quad \begin{vmatrix} -x & 1 \\ 1 & -x \end{vmatrix} = 0 \quad \text{or} \quad x^2 - 1 = 0 \quad \text{or} \quad x = \pm 1$$

Hence eigenvalues of A are {+1, -1}.

Step-II: An eigenvector X, corresponding to the eigenvalue 1 is given by

$$AX = \alpha X \quad \text{or} \quad (A - \alpha I)X = 0.$$

Here $\alpha = 1$ and X = $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

So, $(A - \alpha I)X = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

$$\Rightarrow \quad x_1 = x_2.$$

We can take any value for solution. Let $x_1 = x_2 = 1$. Then an eigen vector corresponding to α = 1 is

$X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = [\,1 \;\; 1]^T.$

Again eigenvector for $\alpha = -1$ is

$(A - \alpha I)X = 0$ or $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

$\Rightarrow \quad x_1 + x_2 = 0$

If $x_1 = 1$, then $x_2 = -1$

So an eigenvector corresponding to $\alpha = -1$ is $X = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$

Step-III: The two eigenspaces W_1 and W_{-1} are given by

$W_1 = \{\alpha \begin{bmatrix} 1 \\ 1 \end{bmatrix} : \alpha \in \mathbf{R}\}$ and $W_{-1} = \{\beta \begin{bmatrix} 1 \\ 1 \end{bmatrix} : \beta \in \mathbf{R}\}$.

Example 2: Let T **:** $\mathbf{R}^3 \rightarrow \mathbf{R}^3$ be a linear operator, where

$T(e_1) = 5e_1 - 6e_2 - 6e_3$; $T(e_2) = -e_1 + 4e_2 + 2e_3$; $T(e_3) = 3e_1 - 6e_2 - 4e_3$

Find the characteristic values of T and compute the corresponding eigenvectors.

Solution: On the basis of given relations, the matrix of T is

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

So the characteristic equation is det (A – xI) = 0.

$$\Rightarrow \quad \begin{vmatrix} 5-x & -1 & 3 \\ -6 & 4-x & -6 \\ -6 & 2 & -4-x \end{vmatrix} = 0.$$

On solving, we get x = 1, 2, 2. So eigenvalues of T are 1, 2, 2.

Case-I: An eigenvector corresponding to $\alpha = 2$ is given by

$(A - 2I)X = 0$

$$\Rightarrow \begin{bmatrix} 3 & -1 & 3 \\ -6 & 2 & -6 \\ -6 & 2 & -6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Now $R_2 \to R_2 + 2R_1$ and $R_3 \to R_3 + 2R_1$

$$\Rightarrow \begin{bmatrix} 3 & -1 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Since rank of coefficient matrix = number of non-zero rows = 1. So, n – r or 3 – 1 = 2 variables can be given arbitrary values.

So we have $3x_1 - x_2 + 3x_3 = 0$ …..(1)

If we take $x_3 = 0$, we get one arbitrary solution $X = [\,1 \quad 3 \quad 0]^T$.

If we take $x_2 = 0$, we get $X = [\,1\,, \quad 0\,, \quad -1]^T$.

So, two eigenvectors corresponding to $\alpha = 2$ are

$$X_1 = \begin{bmatrix} 1 \\ 3 \\ 0 \end{bmatrix} \quad \text{and} \quad X_2 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}.$$

Case-II: Now eigenvector corresponding to $\alpha = 1$ is

$$(A - I)X = 0$$

$$\Rightarrow \begin{bmatrix} 4 & -1 & 3 \\ -6 & 3 & -6 \\ -6 & 2 & -5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$R_1 \leftrightarrow \frac{R_1}{4}$, we get

$$\begin{bmatrix} 1 & -\frac{1}{4} & \frac{3}{4} \\ -6 & 3 & -6 \\ -6 & 2 & -5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$R_2 \to R_2 + 6R_1$ and $R_3 \to R_3 + 6R_1$

$$\begin{bmatrix} 1 & -\frac{1}{4} & \frac{3}{4} \\ 0 & \frac{3}{2} & -\frac{3}{2} \\ 0 & \frac{1}{2} & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$R_2 \to \frac{2}{3} R_2$ and $R_3 \to 2R_3$

$$\begin{bmatrix} 1 & -\frac{1}{4} & \frac{3}{4} \\ 0 & 1 & -1 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$R_3 \to R_3 - R_2$, we get

$$\begin{bmatrix} 1 & -\frac{1}{4} & \frac{3}{4} \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Hence rank of coefficient matrix is 2. So only 3 – 2 = 1 variable can be given arbitrary value.
Now $x_1 - \frac{x_2}{4} + \frac{3x_3}{4} = 0$ and $0 + x_2 - x_3 = 0$

Let $x_3 = \lambda \in \mathbf{R}$, then $x_2 = \lambda$

So $x_1 = \frac{\lambda}{4} - \frac{3\lambda}{4} = -\frac{\lambda}{2}$

So $X = [x_1 \quad x_2 \quad x_3]^T = [-\frac{\lambda}{2} \quad \lambda \quad \lambda]^T = [-\frac{1}{2} \quad 1 \quad 1]^T = [-1 \quad 2 \quad 2]^T$.

Example 3: Show that the eigen values of a diagonal matrix are exactly the elements in the diagonal. Hence prove that if a matrix B is similar to a diagonal matrix D, then the diagonal elements of D are the eigen values of B.

Solution: Step-I: Let $A = \begin{bmatrix} a_{11} & 0 & 0 & \cdots & \cdots & 0 \\ 0 & a_{22} & 0 & \cdots & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & \cdots & a_{nn} \end{bmatrix}$

Characteristic equation of A is det (A – xI) = 0. So

$$\begin{vmatrix} a_{11}-x & 0 & 0 & \dots & \dots & 0 \\ 0 & a_{22}-x & 0 & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \dots & a_{nn}-x \end{vmatrix} = 0$$

$\Rightarrow \quad (a_{11}-x)(a_{22}-x) \ldots\ldots (a_{nn}-x) = 0$

$\Rightarrow x = a_{11}, a_{22}, \ldots, a_{nn}.$

Hence the eigenvalues of A are its diagonal entries.

Step-II: We have already proved that similar matrices have identical eigenvalues. So both matrices have same eigen values.

Example 4: Let V be the vector space of all real-valued continuous functions. Prove that the linear operator T: V→V defined as $(Tf)x = \int_0^x f(t)dt$ has no eigenvalues.

Solution: Suppose α is an eigenvalue of T. Then there exists some $0 \neq f \in V$ such that $Tf = \alpha f$.

$\Rightarrow \quad (Tf)(x) = (\alpha f)(x)$

$\Rightarrow \quad \int_0^x f(t)dt = \alpha f(x)$, by given condition ….(1)

Differentiating with respect to x, we get

$f(x) = \alpha f'(x)$, or $\frac{f'(x)}{f(x)} = \frac{1}{\alpha}$, considering $\alpha \neq 0$

On integration, $\log_e f(x) = \frac{x}{\alpha} + \log_e a$ or $f(x) = a\, e^{x/\alpha}$ ….(2)

Putting x = 0 in equation (2), we get

$f(0) = ae^0$ or $a = f(0)$

So $f(x) = f(0)\, e^{x/\alpha}$ ….(3)

For equation (3) , we have

$\int_0^x f(0)e^{t/\alpha}\, dt = \int_0^x f(t)dt$

$f(0)\,(\alpha\, e^{t/\alpha})_0^x = \alpha\, f(x)$, using equation (1)

$\Rightarrow \quad f(0)\,\alpha\,(e^{x/\alpha} - 1) = \alpha\, f(0)\, e^{x/\alpha}$; using (3)

$\Rightarrow \quad \alpha\, e^{x/\alpha} - \alpha = \alpha\, e^{x/\alpha}$

$\Rightarrow \quad \alpha = 0$, contradiction.

So initial assumption was wrong. Hence T has no eigenvalue.

Note: We observed that diagonal matrices are easiest to find eigen values. So it is a natural question, whether we can transform every square matrix into diagonal matrix?

The answer is **NO**. Then there is a need of condition for that. Let us study these basics:

10.5 *DIAGONALIZABLE OPERATOR*

A linear operator T on a finite-dimensional vector space V(F) is called diagonalizable, if there exists an ordered basis β of V such that the matrix of T with respect to the basis β is a diagonal matrix, so

$$[T]_\beta = \begin{bmatrix} \alpha_1 & 0 & 0 & \cdots & \cdots & 0 \\ 0 & \alpha_2 & 0 & \cdots & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & \cdots & \alpha_n \end{bmatrix} = \mathrm{diag}(\alpha_1, \ldots, \alpha_n).$$

Diagonalizable matrix: An $n \times n$ matrix A over a field **F** is said to be diagonalizable, if it is **similar** to a diagonal matrix. Also A is diagonalizable if there exists an invertible matrix P such that $P^{-1} A P = D$, where D is a diagonal matrix. The matrix P is our *actual need*.

10.6 *BASIS OF DIAGONALIZABLE OPERATORS*

Theorem: A linear operator T on a finite-dimensional vector space V(**F**) is diagonalizable if and only if there exists a basis of V consisting of eigenvectors of T.

Proof: If Part: Let T be diagonalizable, Then ∃ an ordered basis $\beta = \{v_1, \ldots, v_n\}$ of V such that the matrix of T relative to β is $[T]_\beta = \begin{bmatrix} \alpha_1 & 0 & 0 & \ldots & \ldots & 0 \\ 0 & \alpha_2 & 0 & \ldots & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & \ldots & \alpha_n \end{bmatrix}$. From above expression, we get,

$T(V_1) = \alpha_1 v_1 + 0v_2 + \ldots + 0v_n$

$T(V_2) = 0v_1 + \alpha_2 v_2 + \ldots + 0v_n$

- -

$T(V_n) = 0 + 0 + \ldots + \alpha_n v_n$

Or , we can write $T(V_i) = \alpha_i v_i$; i = 1, 2, …, n. Hence $v_1, v_2, \ldots, v_n$ are eigenvectors of T i.e. the basis β consists of eigenvectors of T.

Only if part: Let $\beta = \{v_1, \ldots, v_n\}$ be a basis of V consisting of eigenvectors of T. Then, $\exists\ \alpha_i \in$ **F** such that

$T(V_i) = \alpha_i v_i$; i = 1, 2, …, n.

So, $[T]_\beta = \begin{bmatrix} \alpha_1 & 0 & 0 & \ldots & \ldots & 0 \\ 0 & \alpha_2 & 0 & \ldots & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & \ldots & \alpha_n \end{bmatrix}$

Hence T is diagonalizable.

Theorem: Let T be a linear operator on a finite-dimensional vector space V(**F**). Then the eigenvectors corresponding to distinct eigenvalues of T are linearly independent.

Proof: Let $\alpha_1, \ldots, \alpha_m$ be m distinct eigen values of T and let $v_1, \ldots, v_m$ be the corresponding eigen vectors of T. Then

$T(V_i) = \alpha_i v_i$; i = 1, 2, …, m ….(1)

Claim: $S = \{v_1, \ldots, v_m\}$ is linearly independent. Here we use the **principle of mathematical induction**. If m = 1, then $S = \{v_1\}$ where $v_1 \neq 0$. We know that a single non-zero vector is always linearly independent. So result is true for = 1. Suppose the set { v_1, ..., v_k} is linearly independent, where k < m. We shall prove that the set { v_1, ..., v_k , v_{k+1}} is also linearly independent .

Let $\beta_1 v_1 + \ldots + \beta_k v_k + \beta_{k+1} v_{k+1} = 0$; $\beta_i \in \mathbf{F}$ ….(2)

$\Rightarrow \quad T(\beta_1 v_1 + \ldots + \beta_k v_k + \beta_{k+1} v_{k+1}) = T(0)$

$\Rightarrow \quad \beta_1 T(v_1) + \ldots + \beta_k T(v_k) + \beta_{k+1} T(v_{k+1}) = 0$

$\Rightarrow \quad \beta_1 (\alpha_1 v_1) + \ldots + \beta_k (\alpha_k v_k) + \beta_{k+1} (\alpha_{k+1} v_{k+1}) = 0$ ….(3)

Multiplying equation (2) by α_{k+1} and then subtracting from equation (3), we get

$$\beta_1(\alpha_1 - \alpha_{k+1})v_1 + \ldots + \beta_k(\alpha_k - \alpha_{k+1})v_k = 0.$$

But $v_1, .., v_n$ are linearly independent.

So $\beta_1(\alpha_1 - \alpha_{k+1}) = 0 = \ldots = \beta_k(\alpha_k - \alpha_{k+1})$

$\Rightarrow \quad \beta_1 = 0 \ldots = \beta_k$ as $\alpha_1, \ldots, \alpha_{k+1}$ are all distinct.

Putting these values in equation (2), we get

$$\beta_{k+1} v_{k+1} = 0 \quad \Rightarrow \quad \beta_{k+1} = 0, \text{ as } v_{k+1} \neq 0.$$

So {v_1, ..., v_{k+1}} are also linearly independent if { v_1, ..., v_k} are linearly independent. But we have already proved that the result is true for m = 1. Hence by principle of mathematical induction, S = { v_1, ..., v_m}is linearly independent.

Corollary 1: If T is a linear operator on an n-dimensional vector space V(**F**), then T can not have more than n distinct eigenvalues.

Proof: Let us consider that T has m distinct eigenvalues where m > n. From this theorem, the corresponding m eigen vectors of T are linearly independent. But dim V = n, so maximum number of linearly independent vectors in V(F) is n. **Contradiction!**

So T can’t have more than n distinct eigen values.

Corollary 2: Let T be a linear operator on an n-dimensional vector space V(F) and suppose that T has n distinct eigenvalues. Then T is diagonalizable.

Proof: Suppose T has n distinct eigenvalues, say $c_1, \ldots, c_n$. Let $v_1, \ldots, v_n$ be the corresponding eigenvectors. By using this theorem, $v_1, \ldots, v_n$ are linearly independent over **F**. Since dim V = n, so $\beta = \{v_1, \ldots, v_n\}$ is an ordered basis of V which consists of eigenvectors of T. Hence by this theorem, T is diagonalizable.

Corollary 3: Let T be a linear operator on a finite-dimensional vector space. Let $c_1, \ldots, c_m$ be distinct eigenvalues of T and W_i be the eigenspace of T corresponding to the eigenvalue c_i ; $1 \leq i \leq m$.

So $W = W_1 + W_2 + \ldots + W_m$

If β_i is an ordered basis for W_i, then $\beta = \{ \beta_1, \ldots, \beta_m \}$ is an ordered basis for W. Further dim W = dim W_1 + …. + dim W_m.

Proof: Let $W_1 + W_2 + \ldots + W_m = 0$; where $w_i \in W_i$; ; $1 \leq i \leq m$.

Claim: $w_i = 0$ for each i. Suppose there are some non-zero w_i. If we ignore zero w_i, then,

$w_{i_1} + w_{i_2} + \ldots + w_{i_k} = 0$, each w_{i_k} is non-zero.

$\Rightarrow$ All these vectors are linearly dependent.

But corresponding eigenvalues $c_{i_1}, \ldots, c_{i_k}$ are all distinct.

Contradiction!

So by this theorem, all $w_i = 0$.

Step II: As β_i is an ordered basis for W_i.

$\Rightarrow$ β_i spans W_i.

$\Rightarrow$ $\beta = \{ \beta_1, \ldots, \beta_m \}$ spans the subspace $W = W_1 + W_2 + \ldots + W_m$

Claim: β is a linearly independent set. Let $x_1 + \ldots + x_m = 0$, where $x_i \in W_i$ is some linear combination of the vectors in β_i. So as proved in Step-I, $x_i = 0$ for each i. As each β_i is linearly independent.

$\Rightarrow$ all the scalars in x_i must be zero.

$\Rightarrow$ β is a linearly independent set.

Hence β is a basis of $W = W_1 + \ldots + W_m$

$\Rightarrow$ $\dim W = \dim W_1 + \ldots + \dim W_m$.

Theorem: Let $c_1, \ldots, c_n$ be n distinct eigenvalues of an $n \times n$ matrix A and let $X_1, \ldots, X_n$ be the corresponding eigenvectors of A. If $P = [X_1, \ldots, X_n]$ be $n \times n$ matrix, then A is diagonalizable and $P^{-1} A P = \text{diag}(c_1, \ldots, c_n)$.

Proof: By corollary (3) of previous theorem, it is obvious that A is diagonalizable. Since we know that eigenvectors associated with different eigenvalues are linearly Independent.

$\Rightarrow$ $X_1, \ldots, X_n$ are linearly independent.

So all X_i are non-zero vectors also.

$\Rightarrow$ $\det(P) \neq 0$ i.e. P is invertible.

Given that $AX_i = c_i X_i$, $i = 1, 2, \ldots, n$ ....(1)

Now $AP = A[X_1, \ldots, X_n] = [AX_1, \ldots, AX_n]$

$= [c_1 X_1, \ldots, c_n X_n]$ using(1)

$$= [X_1, \ldots, X_n]\begin{bmatrix} c_1 & 0 & \ldots & \ldots & 0 \\ 0 & c_2 & \ldots & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \ldots & c_n \end{bmatrix}$$

So, $AP = P\,\text{diag}(c_1, \ldots, c_n)$

$\Rightarrow$ $P^{-1} A P = \text{diag}(c_1, \ldots, c_n)$.

Example 1: Let A = $\begin{bmatrix} 5 & -6 & -6 \\ -1 & 4 & 2 \\ 3 & -6 & -4 \end{bmatrix}$ then ,

(i) Find eigenvalues of A, corresponding eigenvectors and eigenspaces of A.

(ii) Is a diagonalizable ?

(iii) Find a non-singular matrix P such that P^{-1} A P is a diagonal matrix.

Solution: (i) Characteristic equation of A is

$$|A - xI| = \begin{vmatrix} 5-x & -6 & -6 \\ -1 & 4-x & 2 \\ 3 & -6 & -4-x \end{vmatrix} = 0$$

On solving we get x = 1, 2, 2.

Case (i): Eigenvector corresponding to x = 1 is given by (A – I)x = 0

$$\Rightarrow \quad \begin{bmatrix} 4 & -6 & -6 \\ -1 & 4 & 2 \\ 3 & -6 & -5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$R_1 \leftrightarrow R_2$, we get

$$\begin{bmatrix} -1 & 3 & 2 \\ 4 & -6 & -6 \\ 3 & -6 & -5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$R_2 \to R_2 + 4R_1$ and $R_3 \to R_3 + 3R_1$

$$\begin{bmatrix} -1 & 3 & 2 \\ 0 & 6 & 2 \\ 0 & 3 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$R_2 \to \frac{1}{2} R_2$

$$\begin{bmatrix} -1 & 3 & 2 \\ 0 & 3 & 1 \\ 0 & 3 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$R_3 \to R_3 - R_2$

$$\begin{bmatrix} -1 & 3 & 2 \\ 0 & 3 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$-x_1 + 3x_2 + 2x_3 = 0$, and $3x_2 + x_3 = 0$

Since rank of coefficient matrix = 2. So only 3 – 2 = 1 variable will take arbitrary value. Let x_3 = 3 , then $x_2 = -1$ and $x_1 = 3$

So $X_1 = \begin{bmatrix} 3 \\ -1 \\ 3 \end{bmatrix}$

Case (ii): Eigen vector, corresponding to x = 2 is (A – 2I)X = 0

$$\begin{bmatrix} 3 & -6 & -6 \\ -1 & 2 & 2 \\ 3 & -6 & -6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$R_1 \leftrightarrow R_2$, we get

$$\begin{bmatrix} -1 & 2 & 2 \\ 3 & -6 & -6 \\ 3 & -6 & -6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$R_2 \to R_2 + 3R_1$ and $R_3 \to R_3 + 3R_1$

$$\begin{bmatrix} -1 & 2 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$\Rightarrow \quad -x_1 + 2x_2 + 2x_3 = 0$

Here rank of coefficient matrix is 1. So 3 – 1 = 2 variables can take arbitrary value. By taking $x_2 = 0$, we get $x_1 = 2$, $x_3 = 1$. By taking $x_3 = 0$, we get $x_1 = 2$, $x_2 = 1$. So two linearly independent eigenvectors corresponding to x = 2 are $X_2 = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$ and $X_3 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$

Case (iii): $W_1 = \{ aX_1 : a \in \mathbf{R} \} = \{ a(3, -1, 3) : a \in \mathbf{R} \}$

$W_2 = \{bX_2 + cX_3 : b, c \in \mathbf{R} \} = \{ b(2, 0, 1) + c(2, 1, 0) : b , c \in \mathbf{R}\}$

(iii) First we show that X_1, X_2, X_3 are linearly independent over **R**. Let a, b $\in$ **R** such that $aX_1 + bX_2 + c X_3 = 0$. Then

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

3a + 2b + 2c = 0

-a + 0b + 0c = 0 $\Rightarrow$ a = 0

$3a + b + 0c = 0$

So we have $b + c = 0$ and $b = 0$

$\Rightarrow c = 0$

So X_1, X_2 , X_3 are linearly independent. Hence A is diagonalizable.

(iii) Let $P = [X_1 \quad X_2 \quad X_3] = \begin{bmatrix} 3 & 2 & 2 \\ -1 & 0 & 1 \\ 3 & 1 & 0 \end{bmatrix}$

Now using elementary properties of matrices, we can get P^{-1} . Then it can be easily verified that

$$P^{-1}A\,P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

Example: For the matrix, $A = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 1 & -6 \\ 2 & -2 & 3 \end{bmatrix}$, find a matrix P, such that $P^{-1}A\,P$ is a diagonal matrix.

Solution: For given matrix, characteristic equation is $|A - xI| = \begin{vmatrix} 1-x & 2 & 0 \\ 2 & 1-x & -6 \\ 2 & -2 & 3-x \end{vmatrix} = 0$

On solving, we get x = 5, 3, -3. As A is 3×3 matrix having three different eigenvalues. So A is diagonalizable.

Case I: Eigenvector, corresponding to x = 5 is given by $(A - 5I)X = 0$

$$\Rightarrow \quad \begin{bmatrix} -4 & 2 & 0 \\ 2 & -4 & -6 \\ 2 & -2 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$R_1 \leftrightarrow \frac{1}{2} R_1$, we get

$$\begin{bmatrix} -2 & 1 & 0 \\ 2 & -4 & -6 \\ 2 & -2 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$R_2 \rightarrow R_1 + R_2$ and $R_3 \rightarrow R_3 + R_1$, we get

$$\begin{bmatrix} -2 & 1 & 0 \\ 0 & -3 & -6 \\ 0 & -1 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$R_2 \leftrightarrow \frac{1}{3} R_2$

$$\begin{bmatrix} -2 & 1 & 0 \\ 0 & -1 & -2 \\ 0 & -1 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$R_3 \rightarrow R_3 - R_2$

$$\begin{bmatrix} -2 & 1 & 0 \\ 0 & -1 & -2 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$\Rightarrow$ $-2x_1 + x_2 + 0\, x_3 = 0.$

$x_2 + 2x_3 = 0.$

If we take, $x_3 = -1$, then $x_2 = 2$, $x_1 = 1$.

So eigenvector corresponding to x = 5 is, $X_1 = [1 \quad 2 \quad -1]^T$.

Case II: Now eigenvector corresponding to x = 3 is (A – 3I)X = 0

$$\begin{bmatrix} -2 & 2 & 0 \\ 2 & -2 & -6 \\ 2 & -2 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$R_2 \rightarrow R_1 + R_2$ and $R_3 \rightarrow R_3 + R_1$, we get

$$\begin{bmatrix} -2 & 2 & 0 \\ 0 & 0 & -6 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$\Rightarrow$ $-x_1 + x_2 = 0$ and $x_3 = 0$

$\Rightarrow$ $x_1 = x_2$ and $x_3 = 0$

So eigenvector corresponding to x = 3 is , $X_2 = [1 \quad 0 \quad 0]^T$.

Case III: eigenvector corresponding to x = -3 is

$(A + 3I)X = 0$

$$\Rightarrow \quad \begin{bmatrix} 4 & 2 & 0 \\ 2 & 4 & -6 \\ 2 & -2 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$R_1 \leftrightarrow \frac{1}{2} R_1$, we get

$$\begin{bmatrix} 2 & 1 & 0 \\ 2 & 4 & -6 \\ 2 & -2 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$R_2 \rightarrow R_2 - R_1$ and $R_3 \rightarrow R_3 - R_1$, we get

$$\begin{bmatrix} 2 & 1 & 0 \\ 0 & 3 & -6 \\ 0 & -3 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$R_3 \rightarrow R_3 + R_2$

$$\begin{bmatrix} 2 & 1 & 0 \\ 0 & 3 & -6 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow \quad 2x_1 + x_2 = 0 \text{ and } x_2 - 2x_3 = 0$$

If we take $x_3 = 1$, then $x_2 = 2$ and $x_1 = -1$. So eigenvector corresponding to x = -3 is $X_3 = [-1 \quad 2 \quad 1]^T$. Here eigen vectors corresponding to distinct eigen values are linearly independent.

So $P = [X_1 \quad X_2 \quad X_3] = \begin{bmatrix} 1 & 1 & -1 \\ 2 & 1 & 2 \\ -1 & 0 & 1 \end{bmatrix}$.

Now, we can get P^{-1} such that

$$P^{-1} A P = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -3 \end{bmatrix}.$$

Example: Find the eigenvalues and bases of the corresponding characteristic spaces of the matrix A $= \begin{bmatrix} 2 & 1 & 0 \\ 0 & 1 & -1 \\ 0 & 2 & 4 \end{bmatrix}$

Is **A** diagonalizable? Justify.

Solution: The characteristic equation of A is $\begin{vmatrix} 2-x & 1 & 0 \\ 0 & 1-x & -1 \\ 0 & 2 & 4-x \end{vmatrix} = 0$

On solving, we get x = 2, 2, 3.

Case (i): Eigenvector, corresponding to x = 2 is given by (A – 2I)X = 0

$$\Rightarrow \quad \begin{bmatrix} 0 & 1 & 0 \\ 0 & -1 & -1 \\ 0 & 2 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$R_3 \rightarrow R_3 + 2R_2$

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & -1 & -1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$\Rightarrow \quad x_2 = 0$, $x_2 + x_3 = 0 \quad \Rightarrow x_3 = 0$

x_1 can take any real value . Let $x_1 = 1$

So $\quad X_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$

Case (ii): Eigen vector, corresponding to x = 3 is (A – 3I)X = 0

$$\Rightarrow \quad \begin{bmatrix} -1 & 1 & 0 \\ 0 & -2 & -1 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$R_3 \rightarrow R_3 + R_2$

$$\begin{bmatrix} -1 & 1 & 0 \\ 0 & -2 & -1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$\Rightarrow \quad -x_1 + x_2 = 0\ ,\ \ 2x_2 + x_3 = 0.$

If we take $x_3 = -2$, then $x_2 = 1$, $x_1 = 1$. So eigenvector corresponding to x = 3 is

$$X_2 = \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}.$$

Bases: The characteristic space W_2, corresponding to the eigenvalue x = 2 is spanned by X_1. Hence $\{X_1\}$is a basis of W_2. Similarly $\{X_2\}$ is a basis of W_3. Thus we have obtained two linearly independent eigen vectors X_1 and X_2, corresponding to eigen values 2, 2, 3 of A. So we can't get a 3×3 invertible matrix P such that

$$P^{-1}AP = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

Hence A is **not** diagonalizable.

Theorem: Let T be a linear operator on a finite-dimensional vector space V(**F**). If c_1, …., c_k are k distinct eigenvalues of T and W_i be the eigenspace of T corresponding to the eigenvalue c_i ($1 \leq i \leq k$), then the following conditions are equivalent –

(i) T is diagonalizable.

(ii) The characteristic polynomial of T is
$f(x) = (x - c_1)^{d_1} \ldots\ldots (x - c_k)^{d_k}$, where
$d_k = \dim W_i$ ($1 \leq i \leq k$) and $d_1 + \ldots. + d_k = \dim V = n$.

(iii) $\dim V = \dim W_1 + \ldots.. + \dim W_k$.

Proof: Since we know that $W_i = \{ v_i : T(v_i) = c_i v_i \}$

$$\Rightarrow \quad W_i = \{ v_i : (T - c_i I)(v_i) = 0 \}$$

Claim: We shall prove **(i) ⇒ (ii)**

Suppose T is diagonalizable. Then there exists an ordered basis $\beta = \{ v_1, \ldots., v_n \}$ of V such that the matrix of T relative to β is

$$[T]_\beta = \begin{bmatrix} c_1 & 0 & \dots & \dots & 0 \\ 0 & c_2 & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \dots & c_n \end{bmatrix}.$$

Suppose c_1 is repeated d_1 times, …, c_k is repeated d_k times. Then

$$[T]_\beta = \text{diag}\,[c_1 \ \dots \ c_1 \ \dots \ \dots \ c_k \ \dots \ c_k].$$

So, characteristic polynomial of T is given by

$$f(x) = (x - c_1)^{d_1} \dots\dots (x - c_k)^{d_k}, \text{ where}$$

$$d_1 + d_2 + \dots + d_k = n = \dim V.$$

Thus $[T - c_iI]_\beta$ has only d_i zeros on the main diagonal for all i = 1, 2, …., k and

$$\text{rank}\,(T - c_iI) = n - d_i\ ;\ \forall\ i = 1, 2, \dots, k \qquad \dots(1)$$

Then by **rank-nullity theorem**,

$$\text{Rank}(T - c_iI) + \text{Nullity}(T - c_iI) = \dim V = n \qquad \dots(2)$$

Using equation (1), we have

$$\text{Nullity}\,(T - c_iI) = d_i$$

$$\Rightarrow \quad \dim \ker(T - c_iI) = d_i$$

$$\Rightarrow \quad \dim W_i = d_i \text{ for } i = 1, 2, \dots, k.$$

Claim: Now we shall prove **(ii) ⇒ (iii)**

Here given that, $\dim V = d_1 + d_2 + \dots + d_k$

$$\Rightarrow \quad \dim V = \dim W_1 + \dots + \dim W_k.$$

Claim: Now we shall show **(iii) ⇒ (i)** .

$$\text{Let } \dim V = \dim W_1 + \dots + \dim W_k \qquad \dots(3)$$

Let $W = W_1 + W_2 + \ldots + W_k$.

Since $c_1, \ldots, c_k$ are distinct eigenvalues of T and $W_1, \ldots, W_k$ are the corresponding eigenspaces of T, so

$$\dim W = \dim W_1 + \ldots + \dim W_k \text{ (we have proved this in theorem)} \quad \ldots(4)$$

Further, if β_i is a basis of W_i, for i = 1, 2, …., k ; where $W_i = \ker(T - c_i I)$,

Then $\beta = \{ \beta_1, \ldots, \beta_k \}$ is a basis of W. From equations (3) and (4), we conclude that

$$\dim V = \dim W \text{ and so } V = W = W_1 + W_2 + \ldots + W_k \text{, since W is a subspace of V.}$$

Hence $\beta = \{ \beta_1, \ldots, \beta_k \}$ is a basis of V consisting of eigenvectors of T and so T is diagonalizable.

Check your progress

Problem 1: Find the characteristic polynomials for the identity operator and zero operator on an n-dimensional vector space.

Problem 2: If $c \neq 0$, is an eigenvalue of an invertible operator T, then prove that c^{-1} is an eigenvalue of T^{-1}.

Problem 3: Let T be a linear operator on R^3 which is represented in the standard ordered basis by the matrix $A = \begin{bmatrix} -9 & 4 & 4 \\ -8 & 3 & 4 \\ -16 & 8 & 7 \end{bmatrix}$. Prove that T is diagonalizable by exhibiting a basis for $\mathbf{R}^3$, each vector of which is eigen vector of T.

Problem 4: Find the eigenvalues, eigenvectors of the matrix

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

Problem 5: Find the eigenvalues, eigenvectors and eigenspaces of the matrix $A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

Problem 6: Find the eigenvalues, eigenvectors and eigenspaces of the matrix A=$\begin{bmatrix} 4 & 2 & -2 \\ -5 & 3 & 2 \\ -2 & 4 & 1 \end{bmatrix}$.

Also prove that **A** is diagonalizable.

10.7 *SUMMARY*

In this unit, we understood the concept of linear operators and their different applications. One of such applications is invertibility of T. Then we elaborated the role of bases of T and their representations. At last, we ensured some conditions of diagonalisation of square matrices.

10.8 *GLOSSARY*

Eigen Values of T: Let T be linear operator on a vector space V(**F**). A scalar $\alpha \in$ **F** is called an eigen value or characteristic value of T, if there exists some V $\neq$ 0,

$$v \in V \text{ such that, } T(v) = \alpha\, v.$$

Eigen Vector: If α is an eigen value of T, then v $\in$ V such that T(v) = α v is called an eigen vector or characteristic vector belonging to α.

Eigen space: The set of all eigenvectors of T belonging to an eigenvalue α is called an eigenspace of T, belonging to α. It is represented as W_α. Hence,

$$W_\alpha = \{ v \in V : T(v) = \alpha\, v \}.$$

Similar Matrices: Let A = $[a_{ij}]_{n \times n}$ and B = $[b_{ij}]_{n \times n}$ and P = $[c_{ij}]_{n \times n}$ where a_{ij}, b_{ij}, $c_{ij} \in$ **F**. Then A and B are called similar matrices if, there exist a matrix P such that $A = P^{-1} B P$, where P is non-singular matrix.

10.9 *REFERENCES*

- S Kumaresan; Linear Algebra-A Geometric Approach; PHI-2016.
- K Hoffman & Ray Kunze; Linear Algebra; PHI-1971.
- P Halmos; Finite-Dimensional Vector Spaces;
- G Strang; Introduction To Linear Algebra;

10.10 *SUGGESTED READING*

- NPTEL videos.
- Schaum series.
- A R Vashishtha, Krishna Prakashan; Meerut.
- Graduate Text In Mathematics, Springer.

10.11 *TERMINAL QUESTION*

Long answer type question

1: Let T be a linear operator on a vector space V(**F**). Then prove the following:

(iv) If $0 \neq v \in V$ is an eigenvector of T, then $\alpha \in$ **F** satisfying T(v) = αv is **unique**.

(v) The eigenspace W_α corresponding to an eigen value $\alpha \in$ **F** is a subspace of V(**F**).

(vi) W_α = ker (T – α I) .

2: State and prove the Cayley-Hamilton Theorem for a linear operator.

3: Find the eigen values, eigen vectors and eigen spaces of 2×2 identity matrix.

4: Show that the eigen values of a diagonal matrix are exactly the elements in the diagonal. Hence prove that if a matrix B is similar to a diagonal matrix D, then the diagonal elements of D are the eigen values of B.

5: Let *V* be the vector space of all real-valued continuous functions. Then prove that the linear operator T: V→V defined as (Tf)x = $\int_0^x f(t)dt$ has no eigenvalues.

6: For the matrix, A = $\begin{bmatrix} 1 & 2 & 0 \\ 2 & 1 & -6 \\ 2 & -2 & 3 \end{bmatrix}$, find a matrix P, such that P^{-1}A P is a diagonal matrix.

Short answer type question

1: Let T be a linear operator on a finite-dimensional vector space V(**F**). Then prove that $\alpha \in$ **F** is an eigenvalue of T if and only if T– α I is singular.

2: Prove that similar matrices have the same characteristic polynomial and hence the same characteristic values.

3: Prove that A linear operator T on a finite-dimensional vector space V(**F**) is diagonalizable if and only if there exists a basis of V consisting of eigenvectors of T.

4: Let T be a linear operator on a finite-dimensional vector space V(**F**). Then prove that the eigenvectors corresponding to distinct eigenvalues of T are linearly independent.

1: For the matrix A = $\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 6 & -11 & 6 \end{bmatrix}$, prove that there exists a matrix P such that

$$P^{-1}AP = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}.$$

2: Let A = $\begin{bmatrix} 5 & -6 & -6 \\ -1 & 4 & 2 \\ 3 & -6 & -4 \end{bmatrix}$ then

(i) Find eigenvalues of A, corresponding eigenvectors and eigen spaces of A.
(ii) Is a diagonalizable?
(iii) Find a non-singular matrix P such that $P^{-1}A\,P$ is a diagonal matrix.

10.12 *ANSWERS*

Answers of check your progress:

1. $\{(1 - x)^n , (-1)^n x^n\}$.

3. (eigen values are 3, -1, -1, and P = $\begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 2 \\ 2 & 2 & 0 \end{bmatrix}$).

4. $\{1, k(1,0,0) : k \in \mathbf{R} \}$

5. [1, -1; $X_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, $X_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$, $X_3 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$ $W_1 = L\{ X_1 , X_2 \}$, $W_{-1} = L \{ X_3\}$] .

6. $\{1, 2, 5; \begin{bmatrix}2\\1\\4\end{bmatrix}, \begin{bmatrix}1\\1\\2\end{bmatrix}, \begin{bmatrix}0\\1\\1\end{bmatrix}\}$

Answer of long question:

3: Eigenvalues of A are { +1, -1}. Eigen vector corresponding to $\alpha = 1$ is

$X = \begin{bmatrix}x_1\\x_2\end{bmatrix} = [1 \;\; 1]^T$ and eigenvector corresponding to $\alpha = -1$ is $X = \begin{bmatrix}1\\-1\end{bmatrix}$

6: $\begin{bmatrix}1 & 1 & -1\\2 & 1 & 2\\-1 & 0 & 1\end{bmatrix}$

UNIT-11: JORDAN CANONICAL FORM

CONTENTS

11.1 *INTRODUCTION*

An upper triangular matrix of a specific shape known as a Jordan matrix encoding a linear operator on a finite-dimensional vector space with regard to some basis is called a Jordan normal form, or Jordan canonical form (JCF) in linear algebra. In such a matrix, the diagonal entries to the left and bottom of any non-zero off-diagonal entry equal to 1 are identical, and they are located immediately above the main diagonal (on the superdiagonal).

A vector space V over a field K is defined. If and only if all of the matrix's eigenvalues fall inside K, or, to put it another way, if the operator's characteristic polynomial divides into linear factors over K, there will be a basis with regard to which the matrix has the necessary form. If K

is algebraically closed (that is, if it is the field of complex numbers), then this condition is always met. The eigenvalues (of the operator) are the diagonal entries of the normal form, and the algebraic multiplicity of the eigenvalue is the number of times each eigenvalue appears.

The Jordan normal form of an operator is sometimes known as the Jordan normal form of *M* if the operator was initially given by a square matrix *M*. Any square matrix that has its field of coefficients expanded to include all of the matrix's eigenvalues has a Jordan normal form. While it is customary to group blocks for the same eigenvalue together, no ordering is imposed among the eigenvalues or among the blocks for a given eigenvalue, though the latter could be ordered by weakly decreasing size. Despite its name, the normal form for a given *M* is not entirely unique because it is a block diagonal matrix formed of Jordan blocks, the order of which is not fixed.

In particular, the Jordan–Chevalley decomposition is straightforward when applied to a basis where the operator adopts its Jordan normal form. The Jordan normal form is a specific case of the diagonal form for diagonalizable matrices, such as normal matrices.

The Jordan decomposition theorem was initially proposed by Camille Jordan in 1870, and the Jordan normal form bears his name.

$$\begin{pmatrix} \boxed{\begin{matrix}\lambda_1 & 1 & \\ & \lambda_1 & 1 \\ & & \lambda_1\end{matrix}} & & & & \\ & \boxed{\begin{matrix}\lambda_2 & 1 \\ & \lambda_2\end{matrix}} & & & \\ & & \boxed{\lambda_3} & & \\ & & & \ddots & \\ & & & & \boxed{\begin{matrix}\lambda_n & 1 \\ & \lambda_n\end{matrix}} \end{pmatrix}$$

A matrix example in Jordan normal form. Every matrix entry that isn't visible is zero. The squares that are delineated are called "Jordan blocks". One number lambda is present on the main diagonal of each Jordan block, whereas ones are present above it. The eigenvalues of the matrix are called lambdas, and they don't have to be unique.

11.2 *OBJECTIVES*

After reading this unit learners will be able to

- Understand the concept of Jordan blocks.
- Implement the application of Jordan canonical form.
- Understand the concept of Jordan decomposition theorem.
- Visualized and understand the concept of nilpotent operator.

11.3 *JORDAN BLOCKS*

Let V denote a finite dimensional vector space over a field F.

Suppose that the characteristic polynomial of T splits in F and $\lambda_1, \ldots, \lambda_k$ are the distinct eigenvalues of T in F. Let $N_{\lambda_1}, N_{\lambda_2}, \ldots, N_{\lambda_k}$ be the distinct eigenspaces of T.

We know that the diagonalizability of T means the following direct sum decomposition of V in terms of distinct eigenspaces of T given by

$$V = N_{\lambda_1} \oplus N_{\lambda_2} \oplus \ldots \oplus N_{\lambda_k}.$$

Naively, diagonalizability fails if some N_{λ_i} is "small".

Definition 1: Let $\lambda \in F$. We define a Jordan block J_λ to be the matrix

$$J_\lambda = \begin{pmatrix} \lambda & 1 & 0 & \ldots & 0 & 0 \\ 0 & \lambda & 1 & \ldots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \ldots & \lambda & 1 \\ 0 & 0 & 0 & \ldots & 0 & \lambda \end{pmatrix}$$

Note that the principal diagonal entries are all λ and the upper diagonal entries are all 1. Every other entry is 0. We often omit 0 from the expression.

Our aim is to select an ordered basis B of V such that

$$[T]_B = \begin{pmatrix} A_1 & 0 & \ldots & 0 \\ 0 & A_2 & \ldots & 0 \\ \cdot & \cdot & & \cdot \\ & & \ldots & \\ \cdot & \cdot & & \cdot \\ 0 & 0 & \ldots & A_k \end{pmatrix}$$

where each 0 is a zero matrix, and each A_i is a square matrix of the form (λ) or a Jordan block J_λ defined above, such that λ is an eigenvalue of T.

Definition 2: The matrix $[T]_B$ is called a Jordan canonical form of T. We say that the ordered basis B is a Jordan canonical basis for T.

Jordan block A_i is almost a diagonal matrix. $[T]_B$ is a diagonal matrix if and only if each A_i is of the form (λ).

Example 1: Suppose that T is a linear operator on C^8 , and $B = \{v_1, \cdot\cdot\cdot, v_8\}$ is an ordered basis for C^8 such that

$$J = [T]_B = \begin{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} & & & \\ & (1) & & \\ & & \begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix} & \\ & & & \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \end{pmatrix}$$

is a Jordan canonical form of T.

The characteristic polynomial of T is

$$\det(tI - J) = (t-1)^4 (t-3)^3 t^2,$$

and hence the multiplicity of each eigenvalue is the number of times the eigenvalue appears on the diagonal of J.

Also observe that v_1, v_4, v_5 and v_7 are the only vectors in B that are eigenvectors of T. These are the vectors corresponding to the columns of J with no 1 above the diagonal entry. Note that,

$T(v_2) = v_1 + v_2$ and therefore $(T - I)(v_2) = v_1$ and $(T - I)(v_3) = v_2$, since v_1 and v_4 are eigenvectors of T corresponding to $\lambda = 2$. It follows that $(T - I)^3 (v_i) = 0$ for $i = 1, 2, 3, 4$.

Similarly, $(T - 3I)^2 (v_i) = 0$ for $i = 5, 6$ and $(T - 0I)^2 (v_i) = 0$ for $i = 7, 8$

In view of these observations, we can say that:

If v lies in a Jordan canonical basis for a linear operator T and is associated with a Jordan block with diagonal entry λ, then $(T - \lambda I)^p (v) = 0$ for some large enough p. Eigenvectors satisfy this condition for $p = 1$.

Our aim is to prove that every linear operator whose characteristic polynomial splits has a Jordan canonical form that is unique upto the order of the Jordan blocks. It is not true that Jordan canonical form is completely determined by the characteristic polynomial of the operator.

Example 2: Let $T^{'}$ be the linear operator on C^8 such that $[T^{'}]_B = J^{'}$, , where B is the ordered basis of the previous example and

$$J^{'} = \begin{pmatrix} 1 & & & & & & & \\ & 1 & & & & & & \\ & & 1 & & & & & \\ & & & 1 & & & & \\ & & & & 3 & & & \\ & & & & & 3 & & \\ & & & & & & 0 & \\ & & & & & & & 0 \end{pmatrix}$$

Then the characterstic polynomial of $T^{'}$ is also $(t-1)^4(t-3)^2t^2$, which is the same as that of T of the previous example, but the Jordan canonical forms are different.

11.4 *GENERALIZED EIGENSPACES*

We now extend the definition of eigenspace to generalized eigenspace of an operator *T*. Our aim is to select ordered bases for these subspaces such that their union form an ordered basis for *V* and the Jordan canonical form is achieved.

Definition 3: Let *T* be a linear operator on a vector space *V*, and let $\lambda \in F$. A nonzero vector *v* in *V* is called a generalized eigenvector of *T* corresponding to *λ* if and only if $(T - \lambda I)^p(v) = 0$ for some positive integer *p*.

Note that if *v* is a generalized eigenvector of *T* corresponding to *λ*, and if *p* is the smallest positive integer for which $(T - \lambda I)^p(v) = 0$, then $(T - \lambda I)^{p-1}(v)$ is an eigenvector of *T* corresponding to *λ*. Therefore, *λ* is an eigenvalue of *T*

Definition 4: Let *T* be a linear operator on *V*, and let $\lambda \in F$ be an eigenvalue of *T*. The generalized eigenspace of *T* corresponding to *λ*, denoted by K_λ, is the subset of *V* defined by

$$K_\lambda = \{v \in V \mid (T - \lambda I)^p(v) = 0,\ p \in N\}.$$

K_λ consists of the zero vector and all generalized eigenvectors corresponding to *λ*.

Theorem 1: Let T be a linear operator on V, and let λ be an eigenvalue of T. Then

(i) K_λ is a T-invariant subspace of V containing the eigenspace

$N_\lambda (= \ker(T - \lambda I))$.

(ii) For any scalar $\mu \neq \lambda$, the restriction of $T - \mu I$ to K_λ is one-one.

Proof (i): It is easy to verify.

(ii) Let $v \in K_\lambda$ and $(T - \mu I)(v) = 0$. Suppose that $v \neq 0$. Let p be the smallest integer for which

$(T - \lambda I)^p(v) = 0$, and let $w = (T - \lambda I)^{p-1}(v) \neq 0$. Then $(T - \lambda I)(w) = (T - \lambda I)^p(v) = 0$, and hence $w \in N_\lambda$. Furthermore,

$(T - \mu I)(w) = (T - \mu I)(T - \lambda I)^{p-1}(v) = (T - \lambda I)^{p-1} (T - \mu I)(v) = 0,$

so that $w \in N_\mu$. But $N_\lambda \cap N_\mu = \{0\}$, and thus $w = 0$, contrary to the hypothesis. So $v = 0$ and $(T - \mu I)_{|K_\lambda}$ is one-one.

Theorem 2: Let T be a linear operator on a finite dimensional vector space V such that the characteristic polynomial of T splits in V. Suppose that λ is an eigenvalue of T with multiplicity m. Then

(i) $\dim(K\lambda) \leq m$.

(ii) $K_\lambda = \ker((T - \lambda I)^m)$.

Proof (i): Let $W = K_\lambda$, and let $p(t)$ be the characteristic polynomial of $T_W = T_{|W}$. Then $p(t)$ divides the characteristic polynomial of T, and therefore it follows that λ is the only eigenvalue of T_W. Hence $p(t) = (t - \lambda)^d$, where $d = \dim(W)$ and $d \leq m$.

(ii) Clearly $\ker((T - \lambda I)^m) \subset K_\lambda$. Now let W and $p(t)$ be as in (i). Then $p(T_W)$ is 0 by the Cayley-Hamilton theorem. Therefore, $(T - \lambda I)^d (v) = 0$ for all $v \in W$. Since $d \leq m$, we have $K_\lambda \subset \ker((T - \lambda I)^m)$.

11.5 *JORDAN CANONICAL FORM*

Theorem 3: Let T be a linear operator on a finite dimensional vector space V, such that the characteristic polynomial of T splits in F. Let $\lambda_1, \lambda_2, ..., \lambda_k$ be the distinct eigenvalues of T. Then, for every $v \in V$, there exist vectors
$v_1 \in K_{\lambda_1}, v_2 \in K_{\lambda_2}, ..., v_k \in K_{\lambda_k}$;
Such that $v = v_1 + v_2 + ... + v_k$

Proof: The natural number k denotes the number of distinct eigenvalues of T. The proof in by induction on the number k.

Let $k = 1$, and let m be the multiplicity of λ_1. Then, $(T - \lambda_1)^m$ is the characteristic polynomial of T, and hence $(T - \lambda_1 I)^m = 0$ by the Cayley-Hamilton theorem. Thus $V = k\lambda_1$, and the result follows.

Now suppose that for some integer $k > 1$, the result is true whenever T has less than k distinct eigenvalues. We assume that T has k distinct eigenvalues. Let m_k denote the multiplicity of λ_k and $p(t)$ the characteristic polynomial of T. Then $p(t) = (t - \lambda_k)^{mk}\, q(t)$, for some polynomial $q(t)$ not divisible by $(t - \lambda_k)$. Let $W_k = \text{range}(T - \lambda_k I)^{mk}$. Then, W_k is T-invariant.

Observe that $(T - \lambda_k I)^{mk}$ maps $k\lambda_i$ onto itself for $i < k$. For suppose that $i < k$. Since $(T - \lambda_k I)^{mk}$ maps $k\lambda_i$ into itself and since $\lambda_k \neq \lambda_i$,
it follows from a previous theorem that the restriction of $T - \lambda_k I$ to $k\lambda_i$ is one-to-one and hence onto.

One consequence of this observation is that for $i < k$, $k\lambda_i$ is contained in W_K; and hence λ_i is an eigenvalue of T_{W_K} for $i < k$. Next, observe that λ_k is not an eigenvalue of T_{W_K}. For, suppose that $T(v) = \lambda_k v$ for some v $\in W_k$. Then $v = (T - \lambda_k I)^{mk}\,(w)$ for some $w \in V$, and it follows that

$$0 = (T - \lambda_k I)(v) = (T - \lambda_k I)^{m_k+1}(w).$$

Therefore, $W \in K_{\lambda_k}$ and by a previous theorem we get $v = (T - \lambda_k I)^{mk}\,(w) = 0$. This shows that v can not be an eigenvector, hence λ_k is not an eigenvalue of T_{W_K}.

We observe that every eigenvalue of T_{W_K} is an eigenvalue of T and the distinct eigenvalues of T_{W_K} are $\lambda_1, \cdot\cdot\cdot, \lambda_{k-1}$. Now let $v \in V$. Then $(T - \lambda_k I)^{mk}\,(v) \in W_K$. Since T_{W_K} has $k - 1$ distinct eigenvalues $\lambda_1, \cdot\cdot\cdot, \lambda_{k-1}$, the induction hypothesis applies.

Let K'_{λ_i} be the generalized eigenspace for the operator T_{W_K} with respect to the eigenvalue λ_i, for $i = 1, 2, . . . , k - 1$. Hence, by the induction hypothesis, there exist vectors

$$w_1 \in K'_{\lambda_1}, w_2 \in K'_{\lambda_2}, ..., w_{k-1} \in K'_{\lambda_{k-1}},$$

such that
$(T - \lambda I)^{m_k}(v) = w_1 + w_2 + ... + w_{k-1}$.
We note that

(a) $K'_{\lambda_i} \subset K_{\lambda_i}$ for $i < k$

(b) $(T - \lambda_k I)^{m_k}$ maps K_{λ_i} onto itself for $i < k$

Therefore, it follows that there exist vectors $v_i \in K_{\lambda_i}$ for $i < k$, such that $(T - \lambda_k I)^{m_k}(v_i) = w_i$. Hence, $(T - \lambda_k I)^{m_k}(v) = (T - \lambda_k I)^{m_k}(v_1) + ... + (T - \lambda_k I)^{m_k}(v_{k-1})$,

and it follows that

$v - (v_1 + v_2 + ... + v_{k-1}) \in K_{\lambda_k}$. Therefore, there exists a vector $v_k \in K_{\lambda_k}$ such that $v = v_1 + v_2 + ... + v_k$.

Theorem 4: Let T be a linear operator on a finite dimensional vector space V, such that the characteristic polynomial of T splits in F. Let $\lambda_1 + \lambda_2 + ... + \lambda_k$ be the distinct eigenvalues of T with multiplicities $m_1 + m_2 + ... + m_k$ respectively. For $1 \le i \le k$, let B_i denote an ordered basis for K_{λ_i}. Then, the following statements are true.

(i) $B_i \cap B_j = \phi$ for $i \neq j$.

(ii) $B = B_1 \cup ... \cup B_k$ is an ordered basis for V.

(iii) $\dim(K\lambda_i) = m_i,\ for\, i = 1, ..., k$.

Proof (*i*): Let $v \in B_i \cap B_j \subset K\lambda_i \cap K\lambda_j$, where $i \neq j$. By a previous theorem, $T - \lambda_i I$ is one-one on $K\lambda_j$, and therefore $(T - \lambda_i I)^p(v) \neq 0$ for every positive integer p. This contradicts the fact that $v \in K\lambda_i$, and the result follows.

(ii) Let $v \in V$. We know by the previous theorem that, for $1 \le i \le k$, there exist vectors $v_i \in K\lambda_i$ such that $v = v_1 + \cdot \cdot \cdot + v_k$. Therefore B spans V, since each v_i is a linear combination of the vectors of B_i. Let q be the cardinality of B. Then dim $V \le q$. For each i, let d_i = dim($K\lambda_i$). Then, $q = \sum_{i=1}^{k} d_i \le \sum_{i=1}^{k} m_i = \dim(V)$. Hence, $q = \dim(V)$; consequently B is a basis for V.

(iii) Using (ii) we see that $\sum_{i=1}^{k} d_i = \sum_{i=1}^{k} m_i$. But $d_i \le m_i$, and therefore $d_i = m_i$ for all i.

Corollary 1: Let T be a linear operator on a finite dimensional vector space V, such that the characteristic polynomial of T splits in F. Then T is diagonalizable if and only if $N_\lambda = K_\lambda$ for every eigenvalue λ of T.

Proof: T is diagonalizable over F if and only if $\dim(N_\lambda) = \dim(K_\lambda)$ for each eigenvalue λ of T. But $\dim(N_\lambda) \leq \dim(K_\lambda)$, and hence these subspaces have same dimension if and only if they are equal.

Our aim is to select suitable bases for the generalized eigenspaces of the linear operator T, so that we may use the previous theorem and obtain a Jordan canonical form. We will find the following definition useful.

Definition 5: Let T be a linear operator on a vector space V. Let v be a generalized eigenvector of T corresponding to the eigenvalue λ. Suppose that p is the smallest positive integer for which $(T-\lambda I)^p(v)=0$. Then, the ordered set

$$C=\{(T-\lambda I)^{p-1}(v),(T-\lambda I)^{p-2}(v),\ldots,(T-\lambda I)(v),v\}$$

is called a cycle of length p of generalized eigenvectors of T corresponding to λ. The vectors $(T-\lambda I)^{p-1}(v)$ and v are called the initial vector and the end vector of the cycle, respectively.

Remark: Notice that the initial vector of a cycle of generalized eigenvectors of T is the only eigenvector of T in the cycle. Also observe that if v is an eigenvector of T corresponding to the eigenvalue λ, then the set $\{v\}$ is a cycle of generalized eigenvectors of T corresponding to λ of length 1.

Let us recall some of the main observations of the first example that we discussed. Suppose that T is a linear operator on C^8, and $B = \{v_1, \cdot \cdot \cdot , v_8\}$ is an ordered basis for C^8 such that

$$J=[T]_B=\begin{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} & & & \\ & (1) & & \\ & & \begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix} & \\ & & & \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \end{pmatrix}$$

is a Jordan canonical form of T.

(1) The first four vectors of B lie in K_1.

(2) The vectors in B that determine the first Jordan block of J are of the form

$$\{v_1, v_2, v_3\} = \{(T - I)^2(v_3), (T - I)(v_3), v_3\}.$$

(3) $(T - I)^3 (v_3) = 0$.

The relation between these vectors is the key to finding Jordan canonical form. We observe that the subset $C_1 = \{v_1, v_2, v_3\}$, $C_2 = \{v_4\}$, $C_3 = \{v_5, v_6\}$, $C_4 = \{v_7, v_8\}$ are the cycles of generalized eigenvectors of T that occur in B. Notice that B is a disjoint union of these cycles. Moreover, if $W_i = \text{span}(C_i)$, for $1 \leq i \leq 4$, we see that C_i is a basis for W_i and $[T_{W_i}]c_i$ is the i-th Jordan block of the Jordan canonical form of T.

Theorem 5: Let T be a linear operator on a finite dimensional vector space V whose characteristic polynomial splits in F. Suppose that B is a basis for V such that B is a disjoint union of cycles of generalized eigenvectors of T. Then the following statements are true:

(i) For each cycle C of generalized eigenvectors contained in B, the subspace $W = \text{span}(C)$ is T-invariant, and $[T_W]_C$ is a Jordan block.

(ii) B is a Jordan canonical basis for V.

Proof: Suppose that the cycle C corresponding to λ has length p, and v is the end vector of C. Then, $C = \{v_1, \cdot \cdot \cdot , v_p\}$, where $v_i = (T - \lambda I)^{p-i} (v)$ for $i < p$ and $v_p = v$. We have $(T - \lambda I)(v_1) = (T - \lambda I)^{p-(i-1)}(v) = v_{i-1}$. Therefore, T maps W into itself, and we see that $[T_W]_C$ is a Jordan block.

We can repeat the arguments of (i) for each cycle in B and finally obtain $[T]_B$.

With the help of following theorems we will see that a Jordan canonical basis is nothing but union of disjoint cycles of generalized eigen vectors corresponding to the eigen values of the operator.

Properties 1: Let T be a linear operator on a vector space V, and let λ be an eigenvalue of T. Suppose that $C_1, \cdot \cdot \cdot, C_r$ are cycles of generalized eigenvectors of T corresponding to λ, such that the initial vectors of the C_i s are distinct and form a linearly independent set. Then the C_i's are disjoint and $C = \bigcup_{i=1}^{r} C_i$ is linearly independent.

2: Every cycle of generalized eigenvectors of a linear operator is linearly independent.

3: Let T be a linear operator on a finite dimensional vector space V, and let λ be an eigenvalue of T. Then K_λ has an ordered basis consisting of a union of disjoint cycles of generalized eigenvectors corresponding to λ.

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

The characteristic polynomial of A is $p(t) = (t - 3)(t - 2)^2$, hence $\lambda_1 = 3$, $\lambda_2 = 2$ are the distinct eigenvalues with multipilcites 1 and 2 respectively. Then dim(K_{λ_1}) = 1 and dim(K_{λ_2}) = 2. Clearly,

$N_{\lambda_1} = \ker(T - 3I) = K_{\lambda_1}$ and $(-1, 2, 1) N_{\lambda_1}$. Therefore, $B_1 = \left\{ \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix} \right\}$ is a basis for K_{λ_1}.

Since dim(K_{λ_2}) = 2, therefore a generalized eigenspace has a basis consisting of union of cycles of length 1 or a single cycle of length 2. The first case is impossible because the vectors in this case would be eigenvectors contradicting the fact that dim(N_{λ_2}) = 1. Therefore, the desired basis is a cycle of length 2. A vector v is the end vector of such a cycle if and only if $(A - 2I)(v) \neq 0$, but $(A - 2I)^2 (v) = 0$. Simple calculation shows that

$$\left\{ \begin{pmatrix} -1 \\ -3 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix} \right\}$$

is a basis for the solution space of $(A - 2I)^2 x = 0$. Now choose a vector v in this set so that $(A - 2I)v \neq 0$. The vector $v = (-1, 2, 0)$ is a candidate for v. Since $(A - 2I)(v) = (1, -3, -1)$ we obtained the cycle of generalized eigenvectors $B_2 = \{(A - 2I)v, v\} = \left\{ \begin{pmatrix} -1 \\ -3 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix} \right\}$

Then, $B = \left\{ \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -3 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix} \right\}$

is a Jordan canonical basis and

$$J = [T]_B = \begin{pmatrix} (3) & \\ & \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix} \end{pmatrix}$$

Is a Jordan canonical form for A.

11.6 *JORDAN DECOMPOSITION THEOREM*

Definition 6: An operator $T: V \to V$ is called nilpotent if $T^k = 0$ for some positive integer k.

Theorem 6 (Jordan Decomposition): Let T be a linear operator on a finite dimensional vector space V such that the characteristic polynomial of T splits in F. Then $T = S + Z$, where S is a diagonalizable operator, Z is a nilpotent operator and $SZ = ZS$.

Proof: We divide the proof into the following steps.

Step 1: T has only one distinct eigenvalue λ, of multiplicity $n = \dim V$. Then, $V = K_\lambda$. If we take $Z = T - \lambda I$, $S = \lambda I$, then $T = Z + S$ and $ZS = SZ$. Moreover, S is diagonal in every basis and Z is nilpotent, for $V = K_\lambda = \ker(Z^n)$.

Step 2: In the general case, let $\lambda_1, \ldots, \lambda_k$ be the distinct eigenvalues of T with multiplicities $n_1, \ldots, n_k$. Let $T_i = T_{|K_{\lambda_i}}$. Then $T = T_1 \oplus \cdots \oplus T_k$. Since each T_i has only one eigenvalue λ_i, we can apply the previous result.

Thus $T_i = S_i + Z_i$; such that $S_i = \lambda_i I$ is diagonal on K_{λ_i} and $N_i = T_i - S_i$ is nilpotent of order n_i on K_{λ_i}. Then $T = S + N$, where $S = S_1 \oplus \cdots \oplus S_k$ and $Z = Z_1 \oplus \cdots \oplus Z_k$. Clearly $SZ = ZS$. Moreover, Z is nilpotent and S is diagonalizable. For, if $m = \max(n_1, \cdots, n_k)$,

then $Z_m = (Z_1)^m \oplus \cdots \oplus (Z_k)^m = 0$; and S is diagonalized by a basis for V which is made up of bases for the generalized eigenspaces. Hence the proof.

Definition 7 (Uniqueness of *S* and *Z*): Under the hypothesis of the Jordan decomposition theorem, there is only one way of expressing T as $S + Z$, where S is diagnalizable, Z is nilpotent and $SZ = ZS$.

Proof: Let $K\lambda_1, \ldots, K\lambda_k$ be the generalized eigenspaces of T corresponding to the distinct eigenvalues $\lambda_1, \ldots, \lambda_k$. Then, $V = K\lambda_1 \oplus \ldots \oplus K\lambda_k$ and $T = T_1 \oplus \ldots \oplus T_k$, where $T_i = T_{|K\lambda_i}$.

Note that $K\lambda_i$ is invariant under every operator that commute with *T*. Since *S* and *Z* both commute with *T*, therefore $K\lambda_i$ is invariant under *S* and *Z*. Put $S_i = \lambda_i I$ and $Z_i = T_i - S_i$. It suffices to show that $S_{|K\lambda_i} = S_i$, for this $Z_{|K\lambda_i} = Z_i$, proving the uniqueness of *S* and *Z*.

Since *S* is diagonalizable, so is $S_{|K\lambda_i}$. Therefore $S_{|K\lambda_i} - \lambda_i I = S_{|K\lambda_i} - S_i$ is diagonalizable. This operator is the same as $Z_i - Z_{|K\lambda_i}$. Since $Z_{|K\lambda_i}$ commutes with $\lambda_i I$ and with T_i, it also commutes with Z_i . We can use binomial theorem to prove that $Z_i - Z_{|N_i}$ is nilpotent.

Hence, the matrix representation of $S_{|N_i} - S_i$ is nilpotent diagonal matrix, and therefore the zero matrix. Hence the proof.

Computation:

By a previous theorem, each generalized eigenspace $K\lambda_i$ contains an ordered basis B_i consisting of a union of disjoint cycles of generalized eigenvectors corresponding to *i* . Then $B = \bigcup_{i=1}^{k} B_i$ is a Jordan canonical basis for *T*. For each *i*, let $T_i = T_{|k\lambda_i}$, and let $A_i = [T_i]_{B_i}$. Then A_i is the Jordan canonical form for T_i, and

$$J = [T]_B = \begin{pmatrix} A_1 & 0 & ... & 0 \\ 0 & A_2 & ... & 0 \\ : & : & : & : \\ 0 & 0 & 0 & A_k \end{pmatrix}$$

is the Jordan canonical form for *T*. We now follow the book by Friedberg et.al. to describe the technique of dot diagrams, followed by some illustrative examples.

The Dot Diagram of $T_i = T_{|K\lambda_i}$ **:** Suppose that B_i is a disjoint union of cycles of generalized eigen vectors $C_1, \cdot\cdot\cdot, C_{n_i}$ with length $p_1 \geq p_2 \geq \cdot\cdot\cdot \geq p_{n_i}$ respectively. The dot diagram of T_i contains one dot for each vector in B_i, and the dots are configured according to the following rules.

- The array consists of n_i columns (one column for each cycle).
- Counting from left to right, the j-th column consists of the p_j dots that correspond to the vectors of C_j starting with the initial vector at the top and continuing down to the end vector.

$$\begin{array}{llll}
.(T-\lambda_i I)^{p_1-1}(v_1) & .(T-\lambda_i I)^{p_2-1}(v_2) & \cdots & .(T-\lambda_i I)^{p_{n_i}-1}(v_{n_i}) \\
.(T-\lambda_i I)^{p_1-2}(v_1) & .(T-\lambda_i I)^{p_2-2}(v_2) & \cdots & .(T-\lambda_i I)^{p_{n_i}-2}(v_{n_i}) \\
\vdots & \vdots & \vdots & \vdots \\
.(T-\lambda_i I)(v_1) & .(T-\lambda_i I)(v_2) & \cdots & .(v_{n_i}) \\
.v_1 & .v_2 & &
\end{array}$$

- The dot diagram of T_i has n_i columns (one for each cycle) and p_1 rows. Since $p_1 \geq p_2 \geq \cdots \geq p_{n_i}$, the columns of the dot diagram either become shorter in length or remain the same in length as we move from left to right

(i) $n_i = \dim(N_{\lambda_i})$

(ii) r_i is the number of dots in the i-th row, given by

$$r_1 = \dim V - \text{rank}(T - \lambda_1 I);$$

$$r_j = \text{rank}((T - \lambda_i I))^{j-1} - \text{rank}((T - \lambda_i I)^j) \text{ if } j > 1.$$

Example 3: Let $A = \begin{pmatrix} 2 & -1 & 0 & 1 \\ 0 & 3 & -1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & -1 & 0 & 3 \end{pmatrix}$

Then, $p(t) = (t-2)^3(t-3)$ is the characteristic polynomial. The distinct eigenvalues are $\lambda_1 = 2$, $\lambda_2 = 3$ with multiplicities 3 and 1 respectively. Therefore, $\dim(K\lambda_1) = 3$ and $\dim(K\lambda_2) = 1$. Let $T_1 = T_{|K\lambda_1}, T_2 = T_{|K\lambda_2}$.

The dot diagram of T_1: It has 3 dots. The possibilities are

••• • • •
         •

.   .

We now calculate $r_1 = 4 - \text{rank}(A-2I) = 4-2 = 2$. Therefore, $r_2 = 1$ and the dot diagram is

. .

.

Therefore, the Jordan canonical form for T_1 is $\begin{pmatrix} \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix} & \\ & (2) \end{pmatrix}$ and the Jordan canonical form for T is

$$J = \begin{pmatrix} \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix} & & \\ & (2) & \\ & & (3) \end{pmatrix}$$

We now find a Jordan canonical basis for T. We first find a Jordan canonical basis for T_1.

$$.(T-2I)v_1 \quad .v_2$$
$$.v_1$$

Therefore $v_1 \in \ker((T-2I)^2)$ but $v_1 \notin \ker((T-2I))$. Now

$$(A-2I) = \begin{pmatrix} 0 & -1 & 0 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 0 & 1 \end{pmatrix}; \ (A-2I)^2 = \begin{pmatrix} 0 & -2 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -2 & 1 & 1 \end{pmatrix}$$

It is easy to see that a basis for $\ker((T-2I)^2) = K\lambda_1$ is

$$\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 2 \end{pmatrix} \right\}.$$

Note that $\begin{pmatrix}0\\1\\2\\0\end{pmatrix}$ and $\begin{pmatrix}0\\1\\2\\0\end{pmatrix}$ do not belong to N_{λ_1}. Choose $v_1 = \begin{pmatrix}0\\1\\2\\0\end{pmatrix}$

And consider $(T - 2I)(v_1) = (A - 2I)(v_1) = \begin{pmatrix}-1\\-1\\-1\\-1\end{pmatrix}$

Now choose $v_2 = \begin{pmatrix}1\\0\\0\\0\end{pmatrix}$ which belongs to N_{λ_1} and which is linearly independent of $\begin{pmatrix}-1\\-1\\-1\\-1\end{pmatrix}$. Then

$B_1 = \left\{\begin{pmatrix}1\\0\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\2\\0\end{pmatrix}, \begin{pmatrix}-1\\-1\\-1\\-1\end{pmatrix}\right\}$ is linearly independent and hence a basis for K_{λ_1}.

Therefore, the Jordan canonical basis $B_1 = \left\{\begin{pmatrix}1\\0\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\2\\0\end{pmatrix}, \begin{pmatrix}-1\\-1\\-1\\-1\end{pmatrix}\right\}$ is associated to the diagram as

$$\cdot\begin{pmatrix}-1\\-1\\-1\\-1\end{pmatrix} \quad \cdot\begin{pmatrix}1\\0\\0\\0\end{pmatrix}$$

$$\cdot\begin{pmatrix}0\\1\\2\\0\end{pmatrix}$$

Since $\lambda_2 = 3$ has multiplicity 1, we have $\dim(K_{\lambda_2}) = \dim(N_{\lambda_2}) = 1$. Hence, any eigenvector constitute a basis B_2. Therefore, we may consider

$$B_2 = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

Thus,

$$B = B_1 \cup B_2 = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \\ -1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\}$$

is a Jordan canonical basis for A. If we take $Q = \begin{pmatrix} -1 & 0 & 1 & 1 \\ -1 & 1 & 0 & 0 \\ -1 & 2 & 0 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix}$.

then $QJQ^{-1} = A$

Example 4: Let $A = \begin{pmatrix} 2 & -4 & 2 & 2 \\ -2 & 0 & 1 & 3 \\ -2 & -2 & 3 & 3 \\ -2 & -6 & 3 & 7 \end{pmatrix}$.

The characteristic polynomial is $p(t) = (t-2)^2(t-4)^2$ and the eigenvalues are $\lambda_1 = 2$, $\lambda_2 = 4$. Let $T_1 = K_{\lambda_1}$, $T_2 = K_{\lambda_2}$.

Dot diagram of T_1 :

$$\cdot\ \cdot \qquad :$$

Now $r_1 = 4 - \text{rank}(A - 2I) = 4 - 2 = 2$. Therefore, the correct dot diagram is

$$\cdot\ \cdot$$

Hence $A_1 = [T_1]_{B_1} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$. In this case B_1 is any basis of N_{λ_1}

e.g., $B_1 = \left\{ \begin{pmatrix} 2 \\ 1 \\ 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 2 \\ 0 \end{pmatrix} \right\}$.

Dot diagram of T_2: We have $r_1 = 4 - \text{rank}(A - 4I) = 4 - 3 = 1$, therefore the correct dot diagram is

:

and the Jordan block $A_2 = [T_2]_{B_2} = \begin{pmatrix} 4 & 1 \\ 0 & 4 \end{pmatrix}$, where B_2 is any basis for K_{λ_2} corresponding to the dots. In this case B_2 is a cycle of length 2. The end vector of this cycle is a vector $v \in K_{\lambda_2} = \ker((T-4I)^2)$, such that $v \notin N_{\lambda_1} = \ker((T-4I))$. It is easy to see that a basis for N_{λ_1} is

$$\left\{ \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix} \right\}$$

Choose v to be any solution of

$$(A-4I)x = \left\{ \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix} \right\}$$

for example, $v = (A-4I)x = \begin{pmatrix} 1 \\ -1 \\ -1 \\ 0 \end{pmatrix}$

Thus $B_2 = \{(A - 4I)v, v\} = \left\{ \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ -1 \\ 0 \end{pmatrix} \right\}$. Therefore,

$$B = B_1 \cup B_2 = \left\{ \begin{pmatrix} 2 \\ 1 \\ 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ -1 \\ 0 \end{pmatrix} \right\}$$

is a Jordan canonical basis for *A*. The corresponding Jordan canonical form is

$$J = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix},$$

Where $A_1 = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$ and $A_2 = \begin{pmatrix} 4 & 1 \\ 0 & 4 \end{pmatrix}$

Check your progress

Problem 1: For the characteristic polynomial $(t-1)^4 (t-3)^2 t^2$ find the Jordan canonical form.

Problem 2: Check the characteristic polynomial for the matrix $A = \begin{pmatrix} 2 & -4 & 2 & 2 \\ -2 & 0 & 1 & 3 \\ -2 & -2 & 3 & 3 \\ -2 & -6 & 3 & 7 \end{pmatrix}$.

11.7 *SUMMARY*

In this unit, we have learned about the important concept of Jordan blocks, Jordan canonical forms, Jordan decomposition theorem, generalized eigenspaces and nilpotent operator. After completion of this unit learners will be able to:

- Formation of Jordan Canonical form on the basis of characteristic polynomial of any matrix.
- Find out any matrix is nilpotent or not.

- Visualized the concept of Jordan decomposition theorem.

11.8 *GLOSSARY*

- Jordan Blocks
- Jordan canonical form
- Jordan decomposition theorem
- Generalized eigenspaces

11.9 *REFERENCES*

- Gel′fand I.M. Lectures on linear algebra (1989), Courier Corporation.
- Kenneth Hoffman & Ray Kunze, Linear Algebra (2nd edition)(2015), Prentice-Hall.
- David C. Lay, Linear Algebra and its Application (3rd Edition) (2007) Pearson Education Asia, India Reprint.
- Seymour Lipshutz and Marc Lipson, Schaum's outlines "Linear Algebra" (3rd Edition)(2012), Mc Graw Hill Education.
- J. N. Sharma and A. R. Vasistha, Linear Algebra (29th Edition) (1999), Krishna Prakashan.

11.10 *SUGGESTED READING*

- Minking Eie & Shou-Te Chang (2020), **A First Course In Linear Algebra, World Scientific.**
- Axler, Sheldon (2015), Linear algebra done right. Springer.
- https://nptel.ac.in/courses/111106051
- https://archive.nptel.ac.in/courses/111/104/111104137
- https://epgp.inflibnet.ac.in/

11.11 *TERMINAL QUESTION*

Long Answer Type Question:

1. Let T be a linear operator on a finite dimensional vector space V, such that the characteristic polynomial of T splits in F. Let $\lambda_1, \lambda_2, ..., \lambda_k$ be the distinct eigenvalues of T. Then show that , for every $v \in V$, there exist vectors

$v_1 \in K_{\lambda_1}, v_2 \in K_{\lambda_2}, ..., v_k \in K_{\lambda_k}$; Such that $v = v_1 + v_2 + ... + v_k$

2. Let T be a linear operator on a finite dimensional vector space V, such that the characteristic polynomial of T splits in F. Let $\lambda_1 + \lambda_2 + ... + \lambda_k$ be the distinct eigenvalues of T with multiplicities $m_1 + m_2 + ... + m_k$ respectively. For $1 \le i \le k$, let B_i denote an ordered basis for K_{λ_i}. Then prove that the following statements are true.

(a) $B_i \cap B_j = \phi$ for $i \neq j$.

(b) $B = B_1 \cup ... \cup B_k$ is an ordered basis for V.

(c) $\dim(K\lambda_i) = m_i, \; for\, i = 1, ..., k$

3. Let T be a linear operator on a finite dimensional vector space V such that the characteristic polynomial of T splits in F. Then prove that $T = S + Z$, where S is a diagonalizable operator, Z is a nilpotent operator and $SZ = ZS$.

Short answer type question:

1. Let T be a linear operator on V, and let λ be an eigenvalue of T. Then prove that

(i) K_λ is a T-invariant subspace of V containing the eigenspace $N_\lambda(= \ker(T - \lambda I))$.

(ii) For any scalar $\mu \neq \lambda$, the restriction of $T - \mu I$ to K_λ is one-one.

2. Let T be a linear operator on a finite dimensional vector space V such that the characteristic polynomial of T splits in V. Suppose that λ is an eigenvalue of T with multiplicity m. Then

(i) $\dim(K\lambda) \le m$.

(ii) $K_\lambda = \ker((T - \lambda I)^m)$

3. Under the hypothesis of the Jordan decomposition theorem prove that, there is only one way of expressing T as $S + Z$, where S is diagnalizable, Z is nilpotent and $SZ = ZS$.

Fill in the blanks:

1. Every cycle of generalized eigenvectors of a linear operator is ……………

2. An operator $T: V \to V$ is called nilpotent if …………………for some positive integer k

11.12 ***ANSWERS***

Answers of check your progress:

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

2: $p(t) = (t-2)^2(t-4)^2$

Answer of fill in the blanks questions:

1. linearly independent **2.** $T^k = 0$

BLOCK- IV

INNER PRODUCT SPACE AND OPERATORS

UNIT-12: INNER PRODUCT SPACES

CONTENTS

12.1 *INTRODUCTION*

Till now, we have studied about vector spaces **without** any metric (distance) concept viz length, angle and distance. But to *visualize* a vector space, we need metric concepts. Without any metric concept; we can't imagine or visualize the geometry of a space. Here we shall study a special class of vector spaces through which we can understand a model of Euclidean Geometry.

There is a significant difference in Mathematical and Physical interpretation when we focus on vectors. In physical world, a vector is a straight arrow-headed line, while in a vector space, besides real life vectors, we study convergent sequences, continuous functions, differentiable functions, integrable functions as vectors. Such type of vectors may not be arrow-headed lines. So we have to **generalize** the concept of angle between two vectors. For this purpose we study inner product.

12.2 *OBJECTIVES*

After the study of this chapter, we shall understand:

- Inner product and its relation with norm and metric.
- Orthogonalisation and Gram-Schmidt process.
- Cauchy Schwarz and Bessel inequalities.
- Riesz representation theorem.

12.3 *INNER PRODUCT SPACES*

In this chapter, we shall consider vector spaces over the field or real numbers (**R**) or complex numbers (**C**) only. In R^3, we define dot product (or scalar product) as follows:

Let $\vec{a} = (x_1, x_2, x_3)$, $\vec{b} = (y_1, y_2, y_3)$ in R^3 where all x_i, $y_j \in \mathbf{R}$

Now $\vec{a} \cdot \vec{b} = x_1y_1 + x_2y_2 + x_3y_3 = \vec{b} \cdot \vec{a}$

We **observe** that dot product satisfies the following properties:

(i) $\vec{a} \cdot \vec{a} \geq 0$ i.e. $x_1^2 + x_2^2 + x_3^2 \geq 0$

Also if $x_1^2 + x_2^2 + x_3^2 = 0$

$\Rightarrow$ $x_1 = x_2 = x_3 = 0$

i.e. $\vec{a} = (0, 0, 0) = \vec{0}$

(ii) $\vec{a} \cdot \vec{b} = \vec{b} \cdot \vec{a}$, as we already know.

(iii) $\vec{a} \cdot (\lambda \vec{b} + \mu\vec{c}) = \lambda (\vec{a} \cdot \vec{b}) + \mu (\vec{a} \cdot \vec{c}) \quad \forall \ \lambda, \mu \in \mathbf{R}$

Here (ii) and (iii) properties can easily be verified. Similarly we can define dot product on $\mathbf{R^n}$. Sometime $\vec{a} \cdot \vec{b}$ is represented as $< \vec{a} , \vec{b} >$. Now we generalize the concept of dot product as inner product in a vector space.

Inner Product: An inner product on a vector space V is a map $\langle \, , \, \rangle : V \times V \to \mathbf{R}$ satisfying the following properties :

(i) $\langle x , x \rangle \geq 0$ and $\langle x , x \rangle = 0$ if and only if $x = 0$.

(ii) $\langle x , y \rangle = \langle y , x \rangle$

(iii) $\langle x + z , y \rangle = \langle x , y \rangle + \langle z , y \rangle$ and $\langle x , y + z \rangle = \langle x , y \rangle + \langle x , z \rangle$

(iv) $\langle ax , y \rangle = a \langle x , y \rangle \;\; \forall \, x , y , z \in V$ and $a \in \mathbf{R}$

Generally function in analysis is represented by ***f*** ; but here, we represent it by **$\langle \, , \, \rangle$.**

So **$(V, \langle \, , \, \rangle)$** is called an inner product space. For brevity, we say V is an inner product space without explicitly mentioning the inner product $\langle \, , \, \rangle$.

Example 1: The dot product defined above on $\mathbf{R}^n$ (in particular $\mathbf{R}^2$) is an inner product. It can be easily verified. Sometimes it is called **standard inner product**.

Example 2: If we consider inner product on **V(C),** where C represent field of complex numbers, then following properties must be satisfied :

(i) $\langle x , y \rangle = \langle \overline{y , x} \rangle$, where $\langle \overline{y , x} \rangle$ is complex conjugate of $\langle x , y \rangle$.

(ii) $\langle x , x \rangle \geq 0$ and $\langle x , x \rangle = 0 \Leftrightarrow x = 0$.

(iii) $\langle \alpha x + \beta y , z \rangle = \alpha \langle x , z \rangle + \beta \langle y , z \rangle$ where $\alpha , \beta \in \mathbf{C}$

(iv) $\langle x , \alpha y + \beta z \rangle = \bar{\alpha} \langle x , y \rangle + \bar{\beta} \langle x , z \rangle$

Example 3: Prove that the vector space $\mathbf{C}^n(\mathbf{C}) = \{ (\alpha_1, \ldots\ldots \alpha_n) : \alpha_I \in \mathbf{C} \}$ is an inner product space with respect to the inner product : $\langle u , v \rangle = \alpha_1\overline{\beta_1} + \alpha_2\overline{\beta_2} + \ldots\ldots\ldots + \alpha_n\overline{\beta_n}$, where $u = (\alpha_1, \ldots\ldots, \alpha_n)$, $v = (\beta_1, \ldots\ldots, \beta_n) \in \mathbf{C}^n$

Solution: Given that : $\langle u , v \rangle = \alpha_1\overline{\beta_1} + \ldots\ldots\ldots + \alpha_n\overline{\beta_n}$ ………(1)

So we have

(i) $\langle v , u \rangle = \beta_1\overline{\alpha_1} + \ldots\ldots\ldots + \beta_n\overline{\alpha_n}$

$\Rightarrow \langle \overline{v , u} \rangle = \overline{\beta_1\overline{\alpha_1} + \ldots\ldots\ldots + \beta_n\overline{\alpha_n}} = (\overline{\beta_1\overline{\alpha_1}}) + \ldots\ldots\ldots + (\overline{\beta_n\overline{\alpha_n}})$

$= \overline{\beta_1}\, \alpha_1 + \ldots\ldots\ldots + \overline{\beta_n}\, \alpha_n$ (as $\overline{(\overline{\alpha_n})} = \alpha_n \;\forall\, n$)

$= \alpha_1\overline{\beta_1} + \ldots\ldots\ldots + \alpha_n\overline{\beta_n}$

So, $\langle \overline{v , u} \rangle = \langle u , v \rangle$

(ii) $\langle u , u \rangle = \alpha_1\overline{\alpha_1} + \ldots\ldots\ldots + \alpha_n\overline{\alpha_n}$

$= |\alpha_1|^2 + \ldots\ldots\ldots + |\alpha_n|^2 \geq 0$

Also $< u , u > = 0 \Leftrightarrow |\alpha_1|^2 + \dots\dots\dots + |\alpha_n|^2 = 0$

$\Leftrightarrow \alpha_1 = 0 = \alpha_2 = \dots\dots\dots = \alpha_n$

$\Leftrightarrow u = (\alpha_1, \dots\dots\dots \alpha_n), = (0, \dots\dots\dots, 0) = \bar{0}$

(iii) Let $\alpha, \beta \in$ **C** and $w = (\gamma_1, \dots\dots\dots, \gamma_n) \in C^n$, then

$< \alpha u + \beta v , w > = < \alpha(\alpha_1, \dots\dots\dots, \alpha_n) + \beta(\beta_1, \dots\dots\dots, \beta_n), (\gamma_1, \dots\dots\dots, \gamma_n) >$

$= < (\alpha\alpha_1 + \beta\beta_1, \dots\dots, \alpha\alpha_n + \beta\beta_n), (\gamma_1, \dots\dots\dots, \gamma_n) >$

$= (\alpha\alpha_1 + \beta\beta_1)\overline{\gamma_1} + \dots\dots\dots + (\alpha\alpha_n + \beta\beta_n)\overline{\gamma_n}$

$= (\alpha\alpha_1\overline{\gamma_1} + \beta\beta_1\overline{\gamma_1}) + \dots\dots + (\alpha\alpha_n\overline{\gamma_n} + \beta\beta_n\overline{\gamma_n})$

$= \alpha(\alpha_1\overline{\gamma_1} + \dots\dots\dots + \alpha_n\overline{\gamma_n}) + \beta(\beta_1\overline{\gamma_1} + \dots\dots\dots + \beta_n\overline{\gamma_n})$

$= \alpha < u , w > + \beta < v , w >$

Hence $\mathbf{C^n}$ is an inner product space.

Note: The inner product given by equation (1) is called the **standard inner product on $\mathbf{C^n}$.**

Example 4: Prove that the following is an inner product on $\mathbf{R^2}$,

$< u , v > = \alpha_1\beta_1 - 2\alpha_1\beta_2 - 2\alpha_2\beta_1 + 5\alpha_2\beta_2$, where $u = (\alpha_1, \alpha_2)$ and $v = (\beta_1, \beta_2) \in \mathbf{R^2}$.

Solution: Here < u , v > will be a real number, so

(i) $< u , v > = < \overline{v , u} >$, obviously.

(ii) $< u , u > = \alpha_1\alpha_1 - 2\alpha_1\alpha_2 - 2\alpha_2\alpha_1 + 5\alpha_2\alpha_2$

$= \alpha_1^2 - 4\alpha_1\alpha_2 + 5\alpha_2^2$

$= \alpha_1^2 - 4\alpha_1\alpha_2 + 4\alpha_2^2 + \alpha_2^2$

$= (\alpha_1 - 2\alpha_2)^2 \geq 0$

Now, $< u , u > = 0$,

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

$\Leftrightarrow \alpha_1 - 2\alpha_2 = 0$ and $\alpha_2 = 0$.

So $< u , u > = 0 \Leftrightarrow u = (\alpha_1, \alpha_2) = (0 , 0)$

(iii) Let $\alpha, \beta \in$ **R** and $w = (\gamma_1, \gamma_2) \in \mathbf{R^2}$, then

$\alpha u + \beta v = \alpha(\alpha_1, \alpha_2) + \beta(\beta_1, \beta_2) = (\alpha\alpha_1 + \beta\beta_1, \alpha\alpha_2 + \beta\beta_2)$

Now, $< \alpha u + \beta v , w > = < (\alpha\alpha_1 + \beta\beta_1, \alpha\alpha_2 + \beta\beta_2), (\gamma_1, \gamma_n) >$

$= (\alpha\alpha_1 + \beta\beta_1)\gamma_1 - 2(\alpha\alpha_1 + \beta\beta_1)\gamma_2 - 2(\alpha\alpha_2 + \beta\beta_2)\gamma_1 + 5(\alpha\alpha_1 + \beta\beta_1)\gamma_1$

$= \alpha(\alpha_1\gamma_1 - 2\alpha_1\gamma_2 - 2\alpha_2\gamma_1 + 5\alpha_2\gamma_2)$

$+ \beta(\beta_1\gamma_1 - 2\beta_1\gamma_2 - 2\beta_2\gamma_1 + 5\beta_2\gamma_2)$

$= \alpha <u , w > + \beta < v , w >$, (using (1))

Hence < u , v > , defined by equation (1), is an inner product on $\mathbf{R^2}$.

Example 5: Let V be the vector space of all real polynomials of degree ≤ 2. Prove that

< f(x) , g(x) > = $\int_0^1 f(x)g(x)dx$, ∀ f(x) , g(x) ∈ V, is an inner product on V.

Solution: (i) Since f(x) and g(x) are real polynomials, so < f(x) , g(x) > ∈ **R**

Hence < f(x) , g(x) > = < $\overline{f(x), g(x)}$ > = < $\overline{g(x), f(x)}$ >

(ii) Now < f(x) , f(x) > = $\int_0^1 f(x)f(x)dx$ = $\int_0^1 f(x)^2 dx$ ≥ 0

Also, < f(x) , f(x) > = 0 , if and only if

$\int_0^1 f(x)^2 dx$ = 0 , if and only if

⇒ f(x) = 0 ,

So, < f(x) , f(x) > = 0 ⟺ f(x) = 0,

(iii) Let α, β ∈ **R** and f(x), g(x), h(x) ∈ V. Then

< α f(x) + β g(x) , h(x) > = $\int_0^1 (\alpha f(x) + \beta g(x))h(x)dx$

= $\alpha \int_0^1 f(x)h(x)dx + \beta \int_0^1 g(x)h(x)dx$

= α < f(x) , h(x) > + β < g(x), h(x) >

Hence < f(x) , g(x) > , defined by equation (1) is an inner product on V.

Example 6: Given α_1 = (1, 3), α_2= (2, 1) ∈ **R**. Find an α ∈ $\mathbf{R}^2$ such that < α , α_1 > = 3,

< , α_2 > = -1. Here < , > is the standard inner product on $\mathbf{R}^2$.

Solution: We know that the standard inner product on $\mathbf{R}^2$ is

< (a_1 , a_2) ,(b_1 , b_2) > = $a_1b_1 + a_2b_2$ …….(1)

Let α = (x, y) ∈ $\mathbf{R}^2$.

So, < , α_1 > = < (x, y), (1, 3) > = x + 3y = 3 ………(2)

$< \alpha , \alpha_2 > \; = < (x, y), (2, 1) > \; = 2x + y \; = -1$ ……….(3)

On solving equations (2) and (3), we get $x = -6/5$, $y = 7/5$

So, $\alpha = \left(\frac{-6}{5} , \frac{7}{5} \right)$

Example 7: Let W_1 and W_2 be two subspaces of a vector space V . If W_1 and W_2 are both inner product spaces, then prove that $W_1 + W_2$ is also an inner product space.

Solution: Let $x, y \in W_1 + W_2$, then

$x = x_1 + x_2$, $y = y_1 + y_2$ where $x_1 , y_1 \in W_1$ and $x_2 , y_2 \in W_2$

We define, $< x , y > = < x_1 , y_1 > + < x_2 , y_2 >$ ……..(1)

Here, $<x_1 , y_1>$ is the inner product on W_1 and $<x_2 , y_2>$ is the inner product on W_2.

Now from equation (1), we have

(i) $< \overline{y , x} > = \overline{< y_1 , x_1 > + < y_2 , x_2 >} = \overline{< y_1 , x_1 >} + \overline{< y_2 , x_2 >}$
$= < x_1 , y_1 > + < x_2 , y_2 >$ (as W_1 and W_2 are I.P.S.)
$= < x , y >$

(ii) $< x , x > \; = \; < x_1 , x_1 > + < x_2 , x_2 >$

Since, $< x_1 , x_1 > \; \geq 0$ and $< x_2 , x_2 > \; \geq 0$

$\Rightarrow < x_1 , x_1 > + < x_2 , x_2 > \; \geq 0$

$\Rightarrow < x , x > \; \geq 0$

Also, $< x , x > \; = 0$

$\Leftrightarrow \; < x_1 , x_1 > \; = 0$ and $< x_2 , x_2 > \; = 0$

$\Leftrightarrow x_1 = 0$ and $x_2 = 0$

$\Leftrightarrow \; x = x_1 + x_2 \; = 0$

(iii) Let $\alpha , \beta \in \mathbf{F}$ and $z = z_1 + z_2 \in W_1 + W_2$

Now, $\alpha x + \beta y \; = \alpha (x_1 + x_2) + \beta (y_1 + y_2) \; = (\alpha x_1 + \beta y_1) + (\alpha x_2 + \beta y_2)$

So, $< \alpha x + \beta y , z > \; = < (\alpha x_1 + \beta y_1) + (\alpha x_2 + \beta y_2) , z_1 + z_2 >$
$= < (\alpha x_1 + \beta y_1 , z_1 > + < \alpha x_2 + \beta y_2 \; , z_2 >$
$= \alpha < x_1 , z_1 > \; + \beta < y_1 , z_1 > \; + \alpha < x_2 , z_2 > \; + \beta < y_2 , z_2 >$
$= \alpha (< x_1 , z_1 > + \; < x_2 , z_2 > \;) + \beta (< y_1 , z_1 > \; + < y_2 , z_2 > \;)$
$= \alpha < x , z > + \beta < y , z >$, (using eqn. (1))

Hence, $W_1 + W_2$ is also an inner product space.

Theorem 1: Let V be an inner product space and u, v, w ∈ V ; α , β ∈ **F** (where F = **R** or **C**) then,

(i) $< u, \alpha v > = \bar{\alpha} < u, v >$

(ii) $< 0, v > = < u, 0 > = 0$

(iii) $< u, v > = 0, \ \forall u \in V \Rightarrow v = 0$, and

(iv) $< u, v > = 0, \ \forall v \in V \Rightarrow u = 0$,

(v) $< u, w > = < v, w >, \ \forall w \in V \Leftrightarrow u = v$

Proof: (i) By definition of inner product

$< u, \alpha v > = < \overline{\alpha v , u} > = \bar{\alpha} < \overline{v , u} > = \bar{\alpha} < u, v >$

(ii) We know for any u ∈ V and 0 ∈ **F** , 0u = 0 ∈ V

So, $< 0, v > = < 0u, v > = 0 < u, v > = 0$

Similarly , $< u, 0 > = < u, 0v > = \bar{0} < u, v > = 0 < u, v > = 0$

(iii) It is given that $< u, v > = 0, \ \forall u \in V$

In particular, we can write,

$< u, v > = 0$

$\Leftrightarrow u = 0$

Similarly, we can prove other part.

(iv) Let $< u, w > = < v, w >, \ \forall w \in V$ then,

$< u – v, w > = < u, w > - < v, w > = 0$

$< u – v, w > = 0 \ \forall w \in V$

So by previous part, $u – v = 0 \Rightarrow u = v$.

Conversely, If we take u = v, then

$< u, w > - < v, w > = < u – v, w > = < 0 , w > = 0$

Hence , < u, w > = < v, w > $\forall$ w $\in$ V.

Note: If V is an inner product space with standard inner product and say V = $\mathbf{R}^3$, then for a $\in$ $\mathbf{R}^3$,

We have, < a, a > = $a_1^2 + a_2^2 + a_3^2$ where a = (a_1 , a_2 , a_3),

Here $\sqrt{a_1^2 + a_2^2 + a_3^2}$ or $\sqrt{< a, a >}$ is defined as norm of vector a. Actually, it is generalization of length of a physical vector.

Norm of a Vector: Let V be an inner product space. The norm function $\| . \|$: V $\rightarrow$ R has the following properties :

(i) $\| x \| \geq 0$ and $\| x \| = 0$ if and only if x = 0 ; x $\in$ V

(ii) $\| \alpha x \| = |\alpha| \, \| x \|$, $\alpha \in$ **F** , x $\in$ V,

Norm of a vector v $\in$ V is defined as $\| v \| = \sqrt{< v, v >}$.

A vector u in an inner product space V is said to be of unit norm or unit length if

$\| u \| = 1$ or < u, u > = 1.

Furthermore, given a non-zero vector v $\in$ V, there is a vector u $\in$ V such that

$\| u \| = 1$ and v = $\| v \|$ u.

This u is called the **unit vector along v**, because u = $\frac{v}{\| v \|}$ and $\| u \| = \frac{\| v \|}{\| v \|} = 1$

Example 8: (i) Find the norm of the vector x = (2, -3, 6) $\in$ $\mathbf{R}^3$ **.**

(ii) Prove that $\frac{x}{\| x \|}$ is of unit length.

Solution: (i) Using the concept of standard inner product of $\mathbf{R}^3$, we have

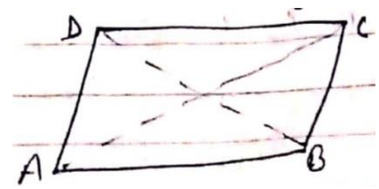
< x, x > = 2(2) + (-3) (-3) + 6(6) = 49

Hence, $\| x \| = \sqrt{< x, x >} = \sqrt{49}$ = 7 units

(ii) Let u = $\frac{x}{\| x \|} = \frac{1}{7}$ (2, -3, 6) = ($\frac{2}{7}$, $\frac{-3}{7}$, $\frac{6}{7}$)

< u, u > = $\frac{2}{7}\left(\frac{2}{7}\right) + \left(\frac{-3}{7}\right)\left(\frac{-3}{7}\right) + \left(\frac{6}{7}\right)\left(\frac{6}{7}\right) = \frac{49}{49} = 1$

$\Rightarrow \| u \| = 1 \Rightarrow u = \frac{x}{\| x \|}$ is of unit length.



Example 9: Let V be an inner product space and x, y, z $\in$ V. Prove that

$$\| x + y \|^2 + \| x - y \|^2 = 2 (\| x \|^2 + \| y \|^2).$$ Also interpret it geometrically.

Solution: Some writers say it parallelogram law.

We have $\| x + y \|^2 = < x + y , x + y > = < x , x + y > + < y , x + y >$

$= < x , x > + < x , y > + < y , x > + < y , y >$ ……….(1)

Now, $\| x - y \|^2 = < x - y , x - y > = < x , x - y > - < y , x - y >$

$= < x , x > - < x , y > - < y , x > + < y , y >$ ……….(2)

Adding equation (1) and (2) , we have

$$\| x + y \|^2 + \| x - y \|^2 = 2 (< x , x > + < y , y >) = 2 (\| x \|^2 + \| y \|^2)$$

Geometric interpretation: Let **x** and **y** be two vectors in the vector space V_2**(R)** with standard inner product defined on it. Suppose the vector **x** is represented by the side AB and the vector **y** by the side BC of a parallelogram ABCD. Then the vectors **x + y** and **x – y** represented the diagonals AC and DB of the parallelogram.

So, $AC^2 + DB^2 = 2(AB^2 + BC^2)$ i.e. the sum of the squares of the sides of a parallelogram is equal to the sum of the squares of its diagonals.

Example 10: Prove that we can **always** define an inner product on a finite-dimensional vector space V(**R**) or V(**C**).

Solution: Let V be a finite- dimensional vector space over the field F = **R** or **C**.

Let B = { α_1 , ….. , α_n } be a basis for V .

Let $\alpha , \beta \in$ V. Then we can write $\alpha = a_1 \alpha_1 + \ldots + a_n \alpha_n$ and $\beta = b_1 \alpha_1 + \ldots + b_n \alpha_n$

Where, $a_1, \ldots, a_n$ and $b_1, \ldots, b_n$ are uniquely determined elements of **F**.

Let us define $< \alpha , \beta > = a_1 \overline{b_2} + \ldots + a_n \overline{b_n}$ ……….(1)

Now it can be easily verified that above expression satisfies all the conditions of inner product. Hence, we can always define an inner product on a finite dimensional vector space V(**C**).

Example 11: If , β are vectors in an inner product space V(**F**) and a, b $\in$ **F**, then prove that

(i) $\|a\alpha + b\beta\|^2 = |a|^2\|\alpha\|^2 + a\bar{b} < \alpha, \beta > + \bar{a}b < \beta, \alpha > + |b|^2 \|\beta\|^2$

(ii) $\text{Re} < \alpha, \beta > = \frac{1}{4}(\|\alpha + \beta\|^2 - \|\alpha - \beta\|^2)$

Solution: (i) We have,

$\|a\alpha + b\beta\|^2 = < a\alpha + b\beta, a\alpha + b\beta > = < a\alpha, a\alpha + b\beta > + < b\beta, a\alpha + b\beta >$

$= a < \alpha, a\alpha + b\beta > + b < \beta, a\alpha + b\beta >$

$= a < \alpha, a\alpha > + a < \alpha, b\beta > + b < \beta, a\alpha > + b < \beta, b\beta >$

$= a\bar{a} < \alpha, \alpha > + a\bar{b} < \alpha, \beta > + b\bar{a} < \beta, \alpha > + b\bar{b} < \beta, \beta >$

$= |a|^2\|\alpha\|^2 + a\bar{b} < \alpha, \beta > + \bar{a}b < \beta, \alpha > + |b|^2 \|\beta\|^2$

(ii) Now we can write

$\|\alpha + \beta\|^2 = < \alpha + \beta, \alpha + \beta > = < \alpha, \alpha + \beta > + < \beta, \alpha + \beta >$

$= < \alpha, \alpha > + < \alpha, \beta > + < \beta, \alpha > + < \beta, \beta >$

$\|\alpha + \beta\|^2 = \|\alpha\|^2 + < \alpha, \beta > + < \beta, \alpha > + \|\beta\|^2$ …..(1)

Also, $\|\alpha - \beta\|^2 = < \alpha - \beta, \alpha - \beta > = < \alpha, \alpha - \beta > - < \beta, \alpha - \beta >$

$= < \alpha, \alpha > - < \alpha, \beta > - < \beta, \alpha > + < \beta, \beta >$

$\|\alpha - \beta\|^2 = \|\alpha\|^2 - < \alpha, \beta > - < \beta, \alpha > + \|\beta\|^2$ …..(2)

Now subtracting equation (2) from equation (1), we get

$\|\alpha + \beta\|^2 - \|\alpha - \beta\|^2 = < \alpha, \beta > + < \beta, \alpha > + < \alpha, \beta > + < \beta, \alpha >$

$= 2(< \alpha, \beta > + < \beta, \alpha >)$

$$= 2(<\alpha, \beta> + <\overline{\alpha, \beta}>)$$

$$= 2\,(2\,\mathrm{Re} <\alpha, \beta>)$$

So, $\mathrm{Re} <\alpha, \beta> = \frac{1}{4}(\|\alpha + \beta\|^2 - \|\alpha - \beta\|^2)$

Note: (1) If F = **R**, then $\mathrm{Re} <\alpha, \beta> = <\alpha, \beta>$

So, $<\alpha, \beta> = \frac{1}{4}(\|\alpha + \beta\|^2 - \|\alpha - \beta\|^2)$

(2) An inner product space V(**R**) is called **Euclidean space** while V(**C**) is called **unitary space**.

Example 12: If α and β are vectors in a unitary space, then prove that –

(i) $4<\alpha, \beta> = \|\alpha + \beta\|^2 - \|\alpha - \beta\|^2 + i\|\alpha + i\beta\|^2 - i\|\alpha - i\beta\|^2$

(ii) $<\alpha, \beta> = \mathrm{Re} <\alpha, \beta> + i\,\mathrm{Re} <\alpha, i\beta>$

Solution: (i) As in previous example, we can write

$$\|\alpha + \beta\|^2 = \|\alpha\|^2 + <\alpha, \beta> + <\beta, \alpha> + \|\beta\|^2$$

and $\|\alpha - \beta\|^2 = \|\alpha\|^2 - <\alpha, \beta> - <\beta, \alpha> + \|\beta\|^2$

So, $\|\alpha + \beta\|^2 - \|\alpha - \beta\|^2 = 2<\alpha, \beta> + 2<\beta, \alpha>$ ……..(1)

Now $\|\alpha + i\beta\|^2 = <\alpha + i\beta, \alpha + i\beta> = <\alpha, \alpha + i\beta> + <i\beta, \alpha + i\beta>$

$$= <\alpha, \alpha> + <\alpha, i\beta> + <i\beta, \alpha> + <i\beta, i\beta>$$

$$= \|\alpha\|^2 + \bar{\imath} <\alpha, \beta> + i <\beta, \alpha> + i\,\bar{\imath} <\beta, \beta>$$

$$= \|\alpha\|^2 - i <\alpha, \beta> + i <\beta, \alpha> + \|\beta\|^2$$

So $i\|\alpha + i\beta\|^2 = i \quad \|\alpha\|^2 + <\alpha, \beta> - <\beta, \alpha> + i\|\beta\|^2$ ………..(2)

Replacing i by –i , we get

$$-i\|\alpha + i\beta\|^2 = -i\|\alpha\|^2 + <\alpha, \beta> - <\beta, \alpha> + i\|\beta\|^2 \quad \text{………..(3)}$$

Hence adding equations (1), (2) and (3), we get

$$\| \alpha + \beta \|^2 - \| \alpha - \beta \|^2 + i \| \alpha + i\beta \|^2 - i \| \alpha - i\beta \|^2 = 4 < \alpha , \beta >$$

(ii) From the knowledge of complex numbers, we have

$$< \alpha , \beta > = \text{Re} < \alpha , \beta > + i\, \text{Im} < \alpha , i\beta > \qquad \text{.........(1)}$$

If z = x + iy, then y = Im z = Re { -i (x + iy) } = Re (-iz)

$\therefore$ Im $< \alpha , \beta >$ = Re { -i $< \alpha , \beta >$ } = Re { $\bar{i} < \alpha , \beta >$ } = Re { $< \alpha , i\beta >$ }

So from (1), we have

$$< \alpha , \beta > = \text{Re} < \alpha , \beta > + i\, \text{Re} < \alpha , i\beta >$$

Note: In the study of physical vectors, we define dot/scalar product as $\vec{a}.\vec{b}$ = ab cos θ, where a = $|\vec{a}|$, b = $|\vec{b}|$ and θ is the angle between $\vec{a}$ and $\vec{b}$.

Since we know that |cos θ| ≤ 1. So, ab|cos θ| ≤ ab as a ≥ 0 , b ≥ 0.

$|\vec{a}.\vec{b}| \leq$ ab or $|\vec{a}.\vec{b}| \leq |\vec{a}||\vec{b}|$

This is a particular case of Cauchy-Schwarz's inequality, which we shall study for an inner product space.

12.4 *CAUCHY SCHWARZ INEQUALITY*

Theorem 2: Let V be an inner product space. If x , y ∈ V, then

$|< x , y >| \leq \| x \| + \| y \|$. Further, equality holds if and only if x and y are linearly dependent (that is, one is a multiple of other).

Proof: Here we shall give three different proofs of Cauchy-Schwarz's inequality:

(i) It is basically geometric in nature
(ii) Here we shall use basic concepts of calculus
(iii) Here we shall use some results on quadratic equations.

Proof: Case (i): If x = 0 or y = 0,

Then $< x , y >$ = 0 and either $< x , x >$ = 0 or $< y , y >$ = 0, Hence the result is obviously true.

Case (ii): Now consider the case, when $\| x \| = \| y \| = 1$,

Consider $< x - y , x - y >$, then by definition of inner product

$$< x - y , x - y > \geq 0,$$

$$\Rightarrow < x , x > - 2 < x , y > + < y , y > \geq 0$$

$$1 - 2 < x , y > + 1 \geq 0$$

$$\Rightarrow < x , y > \leq 1 \qquad \ldots\ldots\ldots(1)$$

Similarly, $< x + y , x + y > \geq 0,$

$$\Rightarrow -< x , y > \leq 1 \qquad \ldots\ldots\ldots(2)$$

Combining both results, we get

$$|< x , y >| \leq 1 \quad \text{or } |< x , y >| \leq \| x \| \, \| y \| \text{ as } \| x \| = \| y \| = 1$$

Now, we prove the statement concerning the equality

Let $|< x , y >| = 1$, then $< x , y > = 1$ or -1

If $< x , y > = 1$, then from the above discussion of inequalities, we deduce that

$$< x - y , x - y > = 0 \quad \text{or } x = y$$

If $< x , y > = -1$, we can deduce that x = -y.

Thus equality holds if and only if either x + y = 0 or x – y = 0.

i.e. if and only if $x = \pm y$.

So x and y are **linearly dependent,** when equality holds.

Case (iii): Now suppose x and y be non-zero and not necessarily of unit length.

Then $u = \frac{x}{\| x \|}$ and $v = \frac{y}{\| y \|}$ s.t. $\| u \| = \| v \| = 1$

Then as in last case, we have $|< u , v >| \leq 1$

$$\text{So } \left|< \frac{x}{\| x \|} , \frac{y}{\| y \|} >\right| = \frac{|< x , y >|}{\| x \| \| y \|} \leq 1$$

$$\Rightarrow | < x , y > | \leq \| x \| \| y \|,$$

Now, in the case of equality, we have $| < x , y > | \leq \| x \| \| y \|$,

If x and y are non-zero, then $< x , y > = \| x \| \| y \|$ or

$$-< x , y > = \| x \| \| y \|$$

If we assume, $< x , y > = \| x \| \| y \|$

$$\Leftrightarrow < \frac{x}{\| x \|} , \frac{y}{\| y \|} > = 1$$

$$\Leftrightarrow \frac{x}{\| x \|} = \frac{y}{\| y \|}$$

$$\Leftrightarrow x = \left(\frac{\| x \|}{\| y \|} \right) y$$

$\Rightarrow$ x is a scalar multiple of y, or x and y are linearly dependent.

The other case is similar.

Proof 2: Fix x and y in V.

If y = 0, then the result is obviously true.

So, we take $y \neq 0$

Let us consider the real valued function of the real variable

$$f(t) = < x + ty, x + ty >.$$

We want to investigate the extremum points of f.

So, $f(t) = < x , x > + 2t < x , y > + t^2 < y , y >$ …….(1)

So we observe that f(t) is a polynomial in t with real coefficients.

Now $f'(t) = 2 < x , y > + 2t < y , y >$

So t_o will be an extremum point for f if $f'(t_o) = 0$,

i.e. $< x , y > + t_o < y , y > = 0$

So, $t_o = -\frac{<x ,y>}{<y ,y>}$

Now $f''(t) = 2 < y , y > = 2 \| y \|^2 > 0$ as $y \neq 0$

So f(t) is minimum at $t = t_o$

$\Rightarrow 0 \leq f(t_o) \leq f(t)$ for all t

$\Rightarrow f(t) \geq 0$ for all t

From equation (1) , we get

$$< x , x > + 2t_o < x , y > + t_o^2 < y , y > \geq 0$$

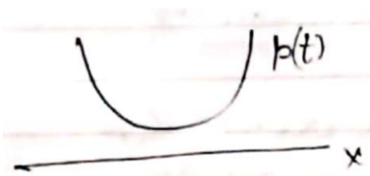
$$< x , x > - \frac{2 (<x ,y>)^2}{<y ,y>} + \frac{(<x ,y>)^2}{<y ,y>} \geq 0$$

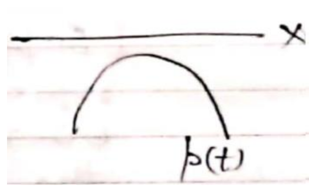
$$< x , x > - \frac{(<x ,y>)^2}{<y ,y>} \geq 0$$

$$\Rightarrow \| x \|^2 \geq \frac{(<x ,y>)^2}{\| y \|^2}$$

$$\Rightarrow |< x , y >| \leq \| x \| \| y \|$$

Proof 3: Let $p(t) = at^2 + bt + c$ be a quadratic polynomial in t with real coefficient. We know that for imaginary roots, p(t) will always remain +ve or always remain –ve.





For this to happen , $b^2 - 4ac \leq 0$.

Now f(t) as in second proof is a quadratic polynomial in t with real coefficients

$a = < y , y >$, $b = 2< x , y >$ and $c = < x , x >$

Also f(t) is always non negative. So we conclude that $b^2 - 4ac \leq 0$. From this, we shall get the required result.

Note: If we consider $\mathbf{R^n}$ with dot(scalar) product, then Cauchy-Schwarz inequality becomes

$$\left| \sum_{i=1}^{n} x_i y_i \right| \leq \left(\sum_i x_i^2\right)^{1/2} \left(\sum_i y_i^2\right)^{1/2}, \text{ for all } x_i, y_i \in \mathbf{R}.$$

This concrete inequality is quite useful in analysis.

Theorem 3: (Triangle Inequality) If , β are vectors in an inner product space V, then

$$\| \alpha + \beta \| \leq \| \alpha \| + \| \beta \|$$

Proof: We have , $\| \alpha + \beta \|^2 = < \alpha + \beta , \alpha + \beta > = < \alpha, \alpha + \beta > + < \beta , \alpha + \beta >$

$$= < \alpha, \alpha > + < \alpha, \beta > + < \beta , \alpha > + < \beta , \beta >$$

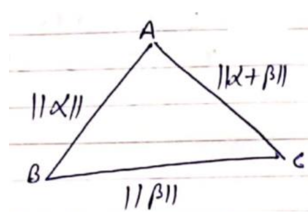
$$= \| \alpha \|^2 + \| \beta \|^2 + (< \alpha, \beta > + < \overline{\alpha, \beta} >)$$

$$\| \alpha + \beta \|^2 = \| \alpha \|^2 + \| \beta \|^2 + 2 \operatorname{Re} (< \alpha, \beta >) \qquad ……(1)$$

But Re (z) $\leq$ |z| ,

So, $\| \alpha + \beta \|^2 \leq \| \alpha \|^2 + \| \beta \|^2 + 2 \,| < \alpha, \beta > |$

$\leq \| \alpha \|^2 + \| \beta \|^2 + 2 \| \alpha \| \| \beta \|$, (by Cauchy Schwarz inequality)



$\Rightarrow \| \alpha + \beta \|^2 \leq (\| \alpha \| + \| \beta \|)^2$

So, $\| \alpha + \beta \| \leq \| \alpha \| + \| \beta \|$

Geometrical Interpretation:

Suppose the vectors α , β represent the sides AB and BC respectively of a Δ ABC in the Euclidean space.

Then $\| \alpha \|$ = AB and $\| \beta \|$ = BC.

Also the vector $\alpha + \beta$ represents the side AC of the triangle ABC and $\| \alpha + \beta \|$ = AC.

Then from above inequality we know, $\| \alpha + \beta \| \leq \| \alpha \| + \| \beta \|$

$\Rightarrow AC \leq AB + BC$

If inequality holds, i.e. AC < AB + BC is true for any triangle ABC.

If equality holds, then AC = AB + BC means points A, B, C are collinear.

Example 13: Verify Cauchy Schwarz inequality for $\alpha = (1, 2, -2)$, and $\beta = (2, 3, 6) \in \mathbf{R}^3$.

Solution: With standard inner product, we have

$$< \alpha, \beta > = 2 + 6 - 12 = -4, \quad \text{so} \ | < \alpha, \beta > | = 4$$

Now, $\| \alpha \|^2 = 1 + 4 + 4$, so $\| \alpha \| = 3$

And $\| \beta \|^2 = 4 + 9 + 36$, then $\| \beta \| = 7$

So, $\| \alpha \| \| \beta \| = 21$

Hence, $| < \alpha, \beta > | \leq \| \alpha \| \| \beta \|$ is verified.

Example 14: If in an inner product space V, $\| \alpha + \beta \| = \| \alpha \| + \| \beta \|$, then prove that α and β are linearly dependent. Show by means of an example that the converse may **NOT** be true.

Solution: Given expression is $(\| \alpha + \beta \|)^2 \leq (\| \alpha \| + \| \beta \|)^2$

$$< \alpha + \beta , \alpha + \beta > = \| \alpha \|^2 + \| \beta \|^2 + 2 \| \alpha \| \| \beta \|$$

$$< \alpha, \alpha > + < \alpha, \beta > + < \beta , \alpha > + < \beta , \beta > = \| \alpha \|^2 + \| \beta \|^2 + 2 \| \alpha \| \| \beta \|$$

$$< \alpha, \beta > + < \overline{\alpha, \beta} > = 2 \| \alpha \| \| \beta \|$$

$$2 \, \text{Re} (< \alpha, \beta >) = 2 \| \alpha \| \| \beta \| \text{ or } \text{Re} (< \alpha, \beta >) = \| \alpha \| \| \beta \| \quad \ldots\ldots(1)$$

But, $\text{Re} (< \alpha, \beta >) \leq | < \alpha, \beta > |$

So, $| < \alpha, \beta > | \geq \| \alpha \| \| \beta \|$ …….(2)

But, by Cauchy Schwarz inequality

$|<\alpha, \beta>| \leq \|\alpha\| \|\beta\|$ …….(3)

From equation (2) and (3) , we have

$|<\alpha, \beta>| = \|\alpha\| \|\beta\|$

So from the equality case of Cauchy Schwarz inequality, we conclude that α and β are linearly dependent .

Conversely, let us take,

$\alpha = (1, -2, 2)$, $\beta = (-2, 4, -4) \in R^3$

Then obviously α and β are linearly dependent as $\beta = -2\alpha$

Now, $\|\alpha\| = \sqrt{1+4+4} = 3$;

$\|\beta\| = \sqrt{4+16+16} = 6$

$\alpha + \beta = (-1, 2, -2) \Rightarrow \|\alpha + \beta\| = \sqrt{1+4+4} = 3$

So, $\|\alpha + \beta\| \neq \|\alpha\| + \|\beta\|$

but α and β are linearly dependent .

Example 15: If W is a subspace of V and v ∈ V satisfies $<v, w> + <w, v> \leq <w, w>$, for all w ∈ W , then prove that $<v, w> = 0$ for all w ∈ W, where V is an inner product space.

Solution: Since W is a subspace of V(F), therefore

$\frac{1}{n} . w = \frac{w}{n} \in W, \forall\, n \in N ; \frac{1}{n} \in \mathbf{F}.$

Given expression is

$<v, w> + <w, v> \leq <w, w>$, for all w ∈ W …..(1)

Replacing w by $\frac{w}{n}$ in equation (1) , we get

$<v, \frac{w}{n}> + <\frac{w}{n}, v> \leq <\frac{w}{n}, \frac{w}{n}>$ or $\frac{1}{n} <v, w> + \frac{1}{n} <w, v> \leq \frac{1}{n^2} <w, w>$

or $< v, w > + < w, v > \leq \frac{1}{n} < w, w >, \quad \forall\, n \in N$

Taking lim n → ∞, we get

$$< v, w > + < w, v > \leq 0$$

Thus $< v, w > + < w, v > \leq 0 , \forall\, w \in W$ …..(2)

Replacing w by –w in equation (2), we get

$$< v, -w > + < -w, v > \leq 0$$

$$-< v, w > - < w, v > \leq 0$$

or $< v, w > + < w, v > \geq 0$ …..(3)

From equations (2) and (3), we conclude that

$$< v, w > + < w, v > = 0 , \forall\, w \in W \quad …..(4)$$

Since W is a subspace of V, so $i \in \mathbf{F}$ and $w \in W \Rightarrow iw \in W$

Replacing w by iw in equation (4), we get

$$< v, iw > + < iw, v > = 0$$

$$\bar{i} < v, w > + i < w, v > = 0$$

$$-i < v, w > + i < w, v > = 0$$

$$-< v, iw > + < iw, v > = 0 \quad …..(5)$$

So subtracting equation (5) from equation (4), we get

$$2< v, w > = 0 \text{ or } < v, w > = 0 , \forall\, w \in W.$$

Definition (Metric): A metric on a set X is a function $d : X \times X \rightarrow R$ with the following properties:

(i) $d(x, y) \geq 0$ for $x, y \in X$ and $d(x, y) = 0$ if and only if $x = y$.

(ii) d(x, y) = d(y, x), for all x, y ∈ X

(iii) d(x, z) ≤ d(x, y) + d(y, z), for all x, y, z ∈ X

(iv) Property (iii) is called the triangle inequality.

Theorem 4: Let V be an inner product space. If we define d(x, y) = ‖ x − y ‖ for x, y ∈ V, Then d is a metric on V.

Proof: (i) By definition of norm, we know

‖ x − y ‖ ≥ 0

⇒ d(x, y) ≥ 0 ,

Also, d(x, y) = 0 , if and only if

‖ x − y ‖ = 0 , if and only if

x – y = 0 , if and only if

x = y

(ii) d(x, y) = ‖ x − y ‖ = ‖ $(-1)(y-x)$ ‖

= |-1| ‖ $y-x$ ‖ by ‖ αx ‖ = |α| ‖ x ‖ ; α ∈ F , x ∈ V

= ‖ $y-x$ ‖ = d(y, x)

(iii) d(x, z) = ‖ x − z ‖

= ‖ (x − y) + (y − z) ‖

≤ ‖ x − y ‖ +‖ y − z ‖ , by triangle inequality

So , d(x, z) ≤ d(x, y) + d(y, z), for all x, y, z ∈ V.

Hence d is a metric on V.

Orthogonality: Let V be an inner product space. An element u ∈ V is said to be orthogonal to v ∈ V if < u, v > = 0. Obviously, orthogonality is a symmetric relation i.e. if u is orthogonal to v, then v is also orthogonal to u.

$< u, v > = 0$, if and only if $< v , u > = 0$

Note: (1) Zero vector is orthogonal to each $v \in V$ as $< 0, v > = < v , 0 > = 0$

(**2**) If $u \in V$ is orthogonal to $v \in V$, then every scalar multiple of u is orthogonal to v. Let $k \in \mathbf{F}$ and $< u , v > = 0$ then $< ku , v > = k < u, v > = 0$. So ku is also orthogonal to v , , $\forall\, k \in \mathbf{F}$.

(**3**) Zero vector is the **only** vector which is orthogonal to itself. If u is orthogonal to u, then

$< u, u > = 0 \Rightarrow u = 0$

(**4**) A vector $u \in V$ is said to be orthogonal to set S if it is orthogonal to each vector in S. That is $< u , v > = 0$, for every $v \in V$.

(**5**) Two subspaces W_1 and W_2 of V(F) are called orthogonal if every vector in each subspace is orthogonal to every vector in the other.

(**6**) Let **S** be a set of vectors in an inner product space V. Then **S** is said to be an orthogonal set provided that any two distinct vectors in **S** are orthogonal. So, $< u , v > = 0$, for every distinct u, $v \in$ **S.**

(**7**) Let **S** be a set of vectors in an inner product space V. The **S** is said to be an orthonormal set if:

(a) $u \in S \Rightarrow \| u \| = 1$

(b) $u, v \in S$ and $u \neq v$, then $< u , v > = 0$

Thus an orthonormal set is an orthogonal set with the additional property that norm of each vector is 1. So a set **S** consisting of mutually orthogonal unit vectors is called an orthonormal set.

A finite set $S = \{ \alpha_1 , \ldots , \alpha_n \}$ is orthonormal if

$$< \alpha_i , \alpha_j > = S_{ij} = \begin{cases} 1 , i = j \\ 0 , i \neq j \end{cases} ,$$

(**8**) If an orthonormal set **S** is a basis of an inner product space V, then the set **S** is called an orthonormal basis of V.

e.g. the set **S** = { (1, 0, 0), (0, 1, 0), (0, 0, 1) } is an orthonormal basis of $\mathbf{R}^3$.

Also, it can be easily verified that the set

$S' = \{ (\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}), (\frac{1}{\sqrt{2}}, 0, \frac{-1}{\sqrt{2}}), (0, 1, 0) \}$ is another orthonormal basis of $\mathbf{R^3}$.

Example 16: (Pythagoras Theorem) Prove that vectors **x** and **y** in a real inner product space (Euclidean space) V are orthogonal if and only if

$$\| x + y \|^2 = \| x \|^2 + \| y \|^2$$

Solution: We have,

$$\| x + y \|^2 = < x + y, x + y > = < x, x + y > + < y, x + y >$$

$$= < x, x > + < x, y > + < y, x > + < y, y >$$

$$= \| x \|^2 + < x, y > + < \overline{x, y} > + \| y \|^2$$

$$\| x + y \|^2 = \| x \|^2 + 2 < x, y > + \| y \|^2 \text{ as V is real I.P.S.} \qquad \ldots..(1)$$

But given that , $\| x + y \|^2 = \| x \|^2 + \| y \|^2$ …..(2)

So we have, $\| x \|^2 + \| y \|^2 = \| x \|^2 + 2 < x, y > + \| y \|^2$

$\Rightarrow \quad < x, y > = 0$

$\Rightarrow$ x and y are orthogonal.

Conversely, let x and y be orthogonal

$\Rightarrow \quad < x, y > = 0$

then as done above, it can be observed,

$$\| x + y \|^2 = \| x \|^2 + 2 < x, y > + \| y \|^2$$

By using, $< x, y > = 0$, we get

$$\| x + y \|^2 = \| x \|^2 + \| y \|^2$$

Example 17: Prove that in a complex inner product space (or unitary space) V, if x is **orthogonal** to y, then $\| x + y \|^2 = \| x \|^2 + \| y \|^2$

However, the converse may **NOT** be true. Justify.

Solution: If x is orthogonal to y, then $< x, y > = 0$

$$\Rightarrow \quad < \overline{x, y} > = 0$$

$$\Rightarrow \quad < y , x > = 0,$$

Now,

$$\| x + y \|^2 = < x, x > + < x, y > + < y, x > + < y, y > \text{, (by previous example)}$$

So, $\| x + y \|^2 = \| x \|^2 + \| y \|^2$

Conversely, let V = $\mathbf{C}^2(\mathbf{C})$ with standard inner product

Let x = (0, i) and y = (0, 1) $\in$ V. Then

$$< x , y > = 0 + i = i \neq 0$$

So x is not orthogonal to y.

Also, $\| x \|^2 = 0(0) + i(\bar{i}) = i(-i) = 1$

$$\| y \|^2 = 0 + 1 = 1$$

Now, x + y = (0, 1 + i)

$$\| x + y \|^2 = 0 + (1 + i)(1 – i) = 2$$

Hence, $\| x + y \|^2 = \| x \|^2 + \| y \|^2$, though x is not orthogonal to y.

Example 18: Find a vector of unit length which is orthogonal to the vector (3, -2, 2) of $\mathbf{R}^3(\mathbf{R})$ relative to the standard inner product .

Solution: Let x = (3, -2, 2) and y = (a, b, c) $\in \mathbf{R}^3$ be orthogonal vectors.

Then $< x, y > = 0$

$$\Rightarrow 3a – 2b + 2c = 0$$

This system has infinite (actually uncountable) solutions. Let us take one solution by taking

a = 2, b = -3, c = -6

So, y = (2, -3, -6) is orthogonal to x = (3, -2, 2)

Now, $\| y \|^2 = 4 + 9 + 36 = 49 \Rightarrow \| y \| = 7$

So, $u = \frac{y}{\| y \|} = \frac{1}{7}(2, -3, -6) \Rightarrow u = \left(\frac{2}{7}, \frac{-3}{7}, \frac{-6}{7}\right)$

Theorem 5: An orthogonal set of non-zero vectors in an inner product space V is linearly independent.

Proof: Let **S** be an orthogonal set of non-zero vectors of V. In order to show that **S** is linearly independent, we shall prove that every finite subset of **S** is linearly independent.

Let $\{v_1, v_2, \ldots, v_n\}$ be any finite subset of **S**.

By orthogonality of **S**, we have

$$< v_i , v_j > = 0 \text{ , for } i \neq j \qquad \ldots(1)$$

Let us assume $\alpha_1 v_1 + \ldots + \alpha_n v_n = 0$; where $\alpha_i \in$ **F**,

So, $< \alpha_1 v_1 + \ldots + \alpha_n v_n , \alpha_1 v_1 + \ldots + \alpha_n v_n > = 0$

$\Rightarrow \quad < \alpha_1 v_1 , \alpha_1 v_1 + \ldots + \alpha_n v_n > + \ldots + < \alpha_n v_n , \alpha_1 v_1 + \ldots + \alpha_n v_n > = 0$

$(< \alpha_1 v_1 , \alpha_1 v_1 > + \ldots + < \alpha_1 v_1 , \alpha_n v_n >) + \ldots + (< \alpha_n v_n , \alpha_1 v_1 > + \ldots + < \alpha_n v_n , \alpha_n v_n >) = 0$

$\alpha_1 \overline{\alpha_1} < v_1 , v_1 > + \alpha_2 \overline{\alpha_2} < v_2 , v > + \ldots + \alpha_n \overline{\alpha_n} < v_n , v_n >$; using equation (1)

$$\| \alpha_1 \|^2 \| v_1 \|^2 + \| \alpha_2 \|^2 \| v_2 \|^2 + \ldots + \| \alpha_n \|^2 \| v_n \|^2 = 0$$

But every term is non-negative and sum is zero.

So, $\| \alpha_i \|^2 \| v_i \|^2 = 0 \ \forall \ i$

But each $v_i \neq 0$, by statement.

So, $| \alpha_i |^2 = 0 \ \forall \ i$

$\Rightarrow \quad \alpha_i = 0 \ \forall \ I = 1, 2, 3, \ldots, n.$

So, $\{v_1, v_2, \ldots, v_n \}$ is linearly independent subset of **S**.

⇒ **every** finite subset of **S** is linearly independent.

⇒ **S** in linearly independent.

Note: In the same way, it can be proved that an orthonormal set S in an inner product space V is linearly independent.

Example 4: If $\{v_1, v_2, \ldots, v_n\}$ is an orthonormal set in V and if $w \in V$, then prove that ,

$$u = w - \sum_{i=1}^{n} < w, v_i > v_i \text{ ; is orthogonal to each of } v_1, v_2, \ldots, v_n .$$

Solution: For any i = 1, 2, ….., n. we have

$$< u, v_i > = < w - \sum_{i=1}^{n} \alpha_i v_i , v_i > \text{, where } \alpha_i = < w, v_i >$$

$$= < w - \alpha_1 v_1 - \alpha_2 v_2 - \ldots - \alpha_n v_n , v_i >$$

$$= < w, v_i > - \alpha_1 < v_1 , v_i > - \ldots - \alpha_i < v_i , v_i > - \ldots - \alpha_n < v_n , v_i >$$

$$= < w, v_i > - 0 - \ldots - \alpha_i - 0 - \ldots - 0$$

$$< u, v_i > = < w, v_i > - < w, v_i > = 0$$

So, $< u, v_i > = 0$, for i = 1, 2, ….., n

Hence u is orthogonal to v_i , for i = 1, 2, ….., n.

Complete Orthonormal Set: An orthonormal set is said to be complete if it is not contained in any larger orthonormal set.

Orthonormal dimension: Let V be a finite-dimensional inner product space of dimension n. If S is any orthonormal set in V then S is linearly independent. So S cannot contain more than n distinct vectors. The orthonormal dimension of V is defined as the largest number of vectors an orthonormal set in V can contain.

For finite dimensional inner product spaces, orthonormal dimension is same as linear dimension.

Note: Now we recall some **basics** of vectors in $\mathbf{R^2}$. It will help us to '*visualize*' the geometry behind **Gram-Schmidt orthogonalisation process**.

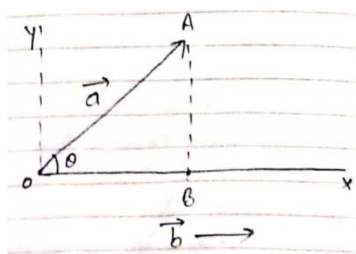
(1) Let us consider two vectors $\vec{a}$ and $\vec{b}$ in R^2. Then

$|\vec{a}| = a$ and $|\vec{b}| = b$

We have to find:

(i) projection of $\vec{a}$ on $\vec{b}$
(ii) component of $\vec{a}$ along $\vec{b}$.
(iii) component of $\vec{a}$ perpendicular to $\vec{b}$.

Let us realize these vectors as shown –



So, $OA = |\vec{a}| = a$, $\angle AOB = \theta$

(i) Projection of $\vec{a}$ on $\vec{b}$

$$= OB = OA \cos\theta$$

$$= \frac{a(\vec{a}.\vec{b})}{a\,b} \text{ as } \vec{a}\,.\,\vec{b} = ab\cos\theta$$

Projection of $\vec{a}$ on $\vec{b}$ = $\frac{(\vec{a}.\vec{b})}{b}$.

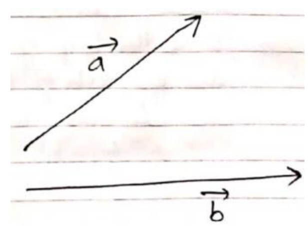
(ii) Component of of $\vec{a}$ along $\vec{b}$ = (Projection of $\vec{a}$ on $\vec{b}$)

$\hat{b} = (\frac{(\vec{a}.\vec{b})}{b})\frac{\vec{b}}{b} = \frac{(\vec{a}.\vec{b})\,\vec{b}}{b^2}$,

(iii) From vector law of addition, we have

$$\overrightarrow{OA} = \overrightarrow{OB} + \overrightarrow{BA}$$

$$\vec{a} = \frac{(\vec{a}.\vec{b})\,\vec{b}}{b^2} + \overrightarrow{BA}$$



So, $\overrightarrow{BA}$ = component of $\vec{a}$ perpendicular to $\vec{b} = \vec{a} - \frac{(\vec{a}.\vec{b})\,\vec{b}}{b^2}$

These fundamental concepts will help you to understand the next theorem.

12.5 *GRAM-SCHMIDT ORTHOGONALISATION PROCESS*

Theorem 6: Every finite-dimensional inner product space has an orthonormal basis.

Proof: Let V(F) be an n-dimensional inner product space and let S = {v_1, ….., v_n }be a basis of V. Firstly, we shall **construct** an orthogonal set in V with the help of elements of S. Since S is a basis, so all elements of S are non-zero.

Let us take,

$$w_1 = v_1,\ w_2 = v_2 - \frac{< v_2, w_1 > w_1}{\|w_1\|^2} \text{ or } w_2 = v_2 - \frac{< v_2, v_1 > v_1}{\|v_1\|^2} \qquad \text{…..(1)}$$

Since $v_1 \neq 0$, so $\| v_1 \| \neq 0$,

We have, $< w_2, w_1 > = < v_2 - \alpha v_1, v_1 >$ where $\alpha = \frac{< v_2, v_1 >}{\|v_1\|^2}$

So, $< w_2, w_1 > = < v_2, v_1 > - \alpha < v_1, v_1 >$

$$= < v_2, v_1 > - \frac{< v_2, v_1 >}{\|v_1\|^2} \| v_1 \|^2 = < v_2, v_1 > - < v_2, v_1 > = 0$$

$\therefore < w_2, w_1 > = 0$ and $v_2 = \alpha v_1 + w_1 = \alpha w_1 + w_2$,

We observe that $w_2 \neq 0$, for otherwise, $v_2 = \alpha v_1$

$\Rightarrow$ v_1 , v_2 are linearly dependent.

This is contradictory, as S is a basis , so every subset of S will be linearly independent.

$$\text{Let } w_3 = v_3 - \frac{< v_3, w_2 > w_2}{\|w_2\|^2} - \frac{< v_3, w_1 > w_1}{\|w_1\|^2} \qquad \text{…..(2)}$$

where $\| w_2 \| \neq 0$, $\| w_1 \| \neq 0$

We can write, $w_3 = v_3 - \alpha_1 w_1 - \alpha_2 w_2$, where

$$\alpha_1 = \frac{< v_3, w_1 >}{\|w_1\|^2} \quad \text{and} \quad \alpha_2 = \frac{< v_3, w_2 >}{\|w_2\|^2} \qquad \text{…..(3)}$$

Now, $< w_3, w_2 > = < v_3 - \alpha_1 w_1 - \alpha_2 w_2 , w_2 >$

$$= < v_3, w_2 > - \alpha_1 < w_1, w_2 > - \alpha_2 < w_2, w_2 >$$

$$= < v_3, w_2 > - \frac{< v_3, w_1 >}{\|w_1\|^2} < w_1, w_2 > - \frac{< v_3, w_2 >}{\|w_2\|^2} \| w_2 \|^2$$

$$= < v_3, w_2 > - 0 - < v_3, w_2 > \quad (\text{as } < w_1, w_2 > = 0)$$

$< w_3, w_2 > = 0$

Similarly, $< w_3, w_1 > = 0$,

Also, $v_3 = \alpha_1 w_1 + \alpha_2 w_2 + w_3$

If follows that { w_1, w_2, w_3 } is an orthogonal set. Further $w_3 \neq 0$, for otherwise,{w_1, w_2, w_3 } is linearly dependent, which is again a contradiction. Here you should note that { w_1, w_2, v_3 }= { $v_1, v_2 - \alpha_1 v_1,$, v_3 } is linearly independent as { v_1, v_2, v_3 }are linearly independent. Proceeding in a similar manner, if we take

$w_n = v_n - \frac{< v_n, w_{n-1} > w_{n-1}}{\| w_{n-1} \|^2} - \ldots - \frac{< v_n, w_1 > w_1}{\| w_1 \|^2}$, then it can be verified that {$w_1, \ldots, w_n$} is an orthogonal set . Consequently, T = { $\frac{w_1}{\| w_1 \|}$, …. , $\frac{w_n}{\| w_n \|}$ } is an orthogonal set. Since an orthonormal set is linearly independent and so T forms basis of V as dim V = n.

Hence T is an orthonormal basis of v.

Note: (1) To obtain an orthonormal basis of V, where V = $\mathbf{R}^3$ i.e. dim V = 3, we proceed as follows:

(i) Let {v_1, v_2, v_3} be a basis of V.

(ii) Find {w_1, w_2, w_3} where $w_1 = v_1$

$$w_2 = v_2 - \frac{< v_2, w_1 > w_1}{\| w_1 \|^2}$$

$$w_3 = v_3 - \frac{< v_3, w_2 > w_2}{\| w_2 \|^2} - \frac{< v_3, w_1 > w_1}{\| w_1 \|^2}$$

(iii) { $\frac{w_1}{\| w_1 \|}$, $\frac{w_2}{\| w_2 \|}$, $\frac{w_3}{\| w_3 \|}$ } is an orthogonal basis of V.

(2) Generally existence theorem in analysis are non-constructive i.e. you prove the theorem, but there is no formula or general method to solve numerical questions. But Gram-Schmidt process is constructive in nature. It provides a method to solve numerical.

Example 19: Apply the Gram-Schmidt process to the vectors given below to obtain an orthonormal basis for R^3**(R)** with the standard inner product:

(i) S_1 = { (1, 1, 0) , (1, 0, 1) , (0, 1, 1) }

(ii) S_2 = { (1, 1, 0) , (1, 0, -1) , (0, 3, 4) }

Solution: (i) Let v_1 = (1, 1, 0) , v_2 = (1, 0, 1) , v_3 = (0, 1, 1)

Let $w_1 = v_1$ = (1, 1, 0) , $\Rightarrow$ $\| w_1 \|^2 = < w_1, w_1 > = 1^2 + 1^2 + 0 = 2$

$$\therefore \frac{w_1}{\|w_1\|} = \frac{1}{\sqrt{2}}(1, 1, 0) = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right)$$

Let $w_2 = v_2 - \frac{<v_2, w_1> w_1}{\|w_1\|^2}$ …..(1)

$$< v_2, w_1 > = < v_2, v_1 > = 1^2 + 0 + 0 = 1$$

So, $w_2 = (1, 0, 1) - \frac{1}{2}(1, 1, 0) = \left(\frac{1}{2}, -\frac{1}{2}, 1\right)$.

$$\| w_2 \|^2 = < w_2 , w_2 > = \frac{3}{2}$$

So, $\frac{w_2}{\|w_2\|} = \sqrt{\frac{2}{3}}\left(\frac{1}{2}, -\frac{1}{2}, 1\right) = = \left(\frac{1}{\sqrt{6}}, -\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}\right)$

Again, let $w_3 = v_3 - \frac{<v_3, w_1> w_1}{\|w_1\|^2} - \frac{<v_3, w_2> w_2}{\|w_2\|^2}$ …..(2)

So we obtain, $< v_3, w_1 > = < v_3, v_1 > = 0 + 1 + 0 = 1$ and $< v_3, w_2 > = \frac{1}{2}$

$$\| w_1 \|^2 = 2 , \quad \| w_2 \|^2 = \frac{3}{2}$$

So form equation (2) , we have

$$w_3 = (0, 1, 1) - \frac{1}{2}(1, 1, 0) - \frac{2}{3}\left(\frac{1}{2}, -\frac{1}{2}, 1\right)\frac{1}{2} = \left(-\frac{2}{3}, \frac{2}{3}, \frac{2}{3}\right)$$

$$\| w_3 \|^2 = \frac{4}{3} \Rightarrow \frac{w_3}{\|w_3\|} = \left(-\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right)$$

Hence orthonormal basis is $\left\{ \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right), \left(\frac{1}{\sqrt{6}}, -\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}\right), \left(-\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right) \right\}$

(ii) Do it yourself.

$$S_1 = \left\{ \left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right), \left(\frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}}\right), (0, 1, 0) \right\}$$

Example 20: Let V be a set of real functions satisfying $\frac{d^2y}{dx^2} + 9y = 0$,

(i) Prove that V is a two-dimensional real vector space.

(ii) In V, inner product is defined by

$$< y, z > = \int_0^{\pi} yz\, dx$$

Find an orthonormal basis for V.

Solution: (i) Suppose V is a collection of solutions of

$$\frac{d^2y}{dx^2} + 9y = 0$$

Let $\frac{d}{dx} \equiv D$

$\Rightarrow$ $(D^2 + 9)\, y = 0$

Auxiliary equation is $m^2 + 9 = 0$ or $m = \pm 3i$

So, solution is $y = c_1 \cos 3x + c_2 \sin 3x$

Let $V = \{c_1 \cos 3x + c_2 \sin 3x : c_1, c_2 \in R\}$ .....(1)

Let $S = \{\cos 3x, \sin 3x\}$

The Wronskian of $v_1 = \cos 3x$ and $v_2 = \sin 3x$ is

$$W(x) = \begin{vmatrix} v_1 & v_2 \\ \frac{dv_1}{dx} & \frac{dv_2}{dx} \end{vmatrix} = \begin{vmatrix} \cos 3x & \sin 3x \\ -3\sin 3x & 3\cos 3x \end{vmatrix} = 3 \neq 0$$

So S is linearly independent subset of V and by equation (1) , L(S)=V.

Hence S is a basis of V.

Thus, $\dim V = 2$

(ii) Let $v_1 = \cos 3x$, $v_2 = \sin 3x$

Now $w_1 = v_1$, So $\| w_1 \|^2 = < w_1 , w_1 > = \int_0^{\pi} \cos^2(3x)\, dx$

$$= \int_0^{\pi} \frac{\cos 6x + 1}{2}\, dx = \frac{\pi}{2} \text{, on solving}$$

$$\therefore \frac{w_1}{\| w_1 \|} = \sqrt{\frac{2}{\pi}} \, . \cos 3x$$

Let $w_2 = v_2 - \frac{<v_2, w_1> w_1}{\|w_1\|^2}$ …..(2)

$\therefore \quad <v_2, w_1> = <v_2, v_1> = \int_0^{\pi} \sin 3x \cos x \; dx = \frac{1}{2}\int_0^{\pi} \sin 6x \; dx = 0,$

$\therefore \quad w_2 = v_2 = \sin 3x$

Now, $\| w_2 \|^2 = <w_2, w_2> = \int_0^{\pi} \sin^2(3x)\, dx = \int_0^{\pi} \left(\frac{1-\cos 6x}{2}\right) dx = \frac{\pi}{2}$

$\therefore \quad \frac{w_2}{\|w_2\|} = \sqrt{\frac{2}{\pi}} \sin 3x$

Hence an orthonormal basis of V is $\left\{ \sqrt{\frac{2}{\pi}} \cos 3x, \sqrt{\frac{2}{\pi}} \sin 3x \right\}$

Example 21: Obtain an orthonormal basis for V, the space of all real polynomials of degree at most 2, the inner product being defined by

$$<f, g> = \int_0^1 f(x)g(x)\, dx$$

Solution: We have, $V = \{ a_0 + a_1 x + a_2 x^2 ; a_i \in \mathbf{R} \}$

Let $S = \{1, x, x^2\}$. Then obviously, S is a basis of V

Let $v_1 = 1$, $v_2 = x$ and $v_3 = x^2$

So, $w_1 = v_1 = 1$

Now $\|w_1\|^2 = <w_1, w_1> = \int_0^1 1.1 \,.\, dx = 1$

Let $\quad w_2 = v_2 - \frac{<v_2, w_1> w_1}{\|w_1\|^2}$ …..(1)

Now $<v_2, w_1> = <v_2, v_1> = \int_0^1 x\, dx = \frac{1}{2}$

$\therefore w_2 = x - \frac{1}{2}$

Hence, $\|w_2\|^2 = <w_2, w_2> = \int_0^1 (x - \frac{1}{2})^2\, dx = \frac{1}{12}$

So, $\frac{w_2}{\|w_2\|} = \sqrt{12}\,(x - \frac{1}{2}) = 2\sqrt{3}\,(x - \frac{1}{2})$

Let $w_3 = v_3 - \frac{< v_3, w_1 > w_1}{\|w_1\|^2} - \frac{< v_3, w_2 > w_2}{\|w_2\|^2}$ …..(2)

Since, $< v_{3,} w_1 > = \int_0^1 x^2\, dx = \frac{1}{3}$

$$< v_{3,} w_2 > = \int_0^1 x^2 \left(x - \frac{1}{2}\right) dx = \frac{1}{12}$$

$$w_3 = x^2 - \frac{1}{3} . 1 - \left(x - \frac{1}{2}\right) = x^2 - x + \frac{1}{6}$$

$$\|w_3\|^2 = < w_3 , w_3 > = \int_0^1 (x^2 - x + \frac{1}{6})^2\, dx = \frac{1}{180}$$

$$\frac{w_3}{\|w_3\|} = \sqrt{180}\,(x^2 - x + \frac{1}{6}) = 6\sqrt{5}\,(x^2 - x + \frac{1}{6})$$

Hence an orthonormal basis of V is

$$\{ 1, 2\sqrt{3}\,(x - \frac{1}{2}) , 6\sqrt{5}\,(x^2 - x + \frac{1}{6}) \}$$

12.6 *BESSEL'S INEQUALITY*

Theorem 7: If V is an inner product space and if $\{w_1 , …. ,w_n\}$ is an orthonormal set in V , then

$\sum_{i=1}^{n} | w_i , v |^2 \leq \|v\|^2$, for all $v \in V$

Furthermore, equality holds if and only if V is in subspace spanned by $w_1 , …. ,w_n$.

Proof: Let $v \in V$ be arbitrary.

Consider the vector

$$x = v - \sum_{i=1}^{n} \alpha_i w_i \text{ ; where } \alpha_i = < v , w_i > \qquad …..(1)$$

Then, $< x, x > = < v - \sum_{i=1}^{n} \alpha_i w_i , v - \sum_{j=1}^{n} \alpha_j w_j$

$= < v, v > - < v , \sum_{j=1}^{n} \alpha_j w_j > - < \sum_{i=1}^{n} \alpha_i w_i , v > + < \sum_{i=1}^{n} \alpha_i w_i , \sum_{j=1}^{n} \alpha_j w_j >$

$= \|v\|^2 - \sum_{j=1}^{n} \overline{\alpha_j} < v , w_j > - \sum_{i=1}^{n} \alpha_i < w_i , v > + \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j < w_i , w_j >$

$= \|v\|^2 - \sum_{j=1}^{n} < \overline{v , w_j} > < v , w_j > - \sum_{i=1}^{n} < v, w_i > < \overline{v , w_i} > + \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \overline{\alpha_i}. 1$ (as $< w_i , w_j > = 1$ only if i = j)

So, $< x, x > = ||v||^2 - \sum_{i=1}^{n} |< v, w_i >|^2 - \sum_{i=1}^{n} |< v, w_i >|^2 + \sum_{i=1}^{n} |< v, w_i >|^2$

$< x, x > = ||v||^2 - \sum_{i=1}^{n} |< v, w_i >|^2 = ||v||^2 - \sum_{i=1}^{n} |< \overline{v, w_i} >|^2$ as $|z| = |\bar{z}|$

$\therefore \quad || x ||^2 = ||v||^2 - \sum_{i=1}^{n} |< w_i , v >|^2$ …..(2)

Since $|| x ||^2 \geq 0$, so by equation (2), we have

$$||v||^2 - \sum_{i=1}^{n} |< w_i , v >|^2 \geq 0 \quad \text{or} \quad \sum_{i=1}^{n} |< w_i , v >|^2 \leq ||v||^2 \text{ for each } v \in V$$

If the equality holds i.e. if $\sum_{i=1}^{n} |< w_i , v >|^2 = ||v||^2$, then from equation (2), we have

$$|| x ||^2 = 0 \text{ or } || x || = 0$$

$$\Rightarrow \quad x = 0$$

So, $v = \sum_{i=1}^{n} \alpha_i w_i = \sum_{i=1}^{n} < v, w_i > w_i$

Thus, if the equality holds, then v is linear combination of $\{ w_1 , …. , w_n \}$.

Conversely, if v is a linear combination of $\{ w_1 , …. , w_n \}$, then we can write

$v = \sum_{i=1}^{n} \alpha_i w_i$ where $\alpha_i = < v , w_i >$

So, $x = 0 \Rightarrow || x ||^2 = 0$

Hence from equation (2), we have

$$||v||^2 = \sum_{i=1}^{n} |< w_i , v >|^2 \quad \text{i.e. equality holds.}$$

12.7 *ORTHOGONAL COMPLEMENT*

Let V be an inner product space, and let S be any set of vectors in V. The orthogonal complement of S (written as $S^{\perp}$ and read as S perpendicular or S perp.) is defined by

$$\mathbf{S^{\perp} = \{ v \in V : < u, v > = 0 \;\; \forall \, u \in S \}}$$

Thus $S^{\perp}$ is the set of all those vectors in V which are orthogonal to every vector in S.

Theorem 8: Let S be any set of vectors in an inner product space V. Then $S^{\perp}$ is a subspace of V.

Proof: By definition, $S^\perp = \{ v \in V : < u, v > = 0 \ \forall u \in S\}$

Since $< 0, u > = 0 \ \forall u \in S$

So, $0 \in S^\perp$ and thus $S^\perp$ is not empty.

Let $x, y \in F$ and $w_1, w_2 \in S^\perp$

Then $< w_1, u > = 0 \ \forall u \in S$ and

$< w_2, u > = 0 \ \forall u \in S$

So, $< xw_1 + yw_2, u > = x < w_1, u > + y < w_2, u >$

$= x.0 + y.0 = 0 \ \forall u \in S$

So, $xw_1 + yw_2 \in S^\perp \ \forall w_1, w_2 \in S^\perp$ and $x, y \in \mathbf{F}$

Hence $S^\perp$ is a subspace of V.

Note: (1) Here we should note that S MAY NOT be a subspace of V while $S^\perp$ is always a subspace of V.

(2) Obviously, it can be observed that $V^\perp = \{ \bar{0} \}$ and $\{ \bar{0} \}^\perp = V$.

Orthogonal Complement of an orthogonal complement: Let S be any subset of an inner product space V. the $S^\perp$ is a subset of B.

We define $(S^\perp)^\perp$, written as $S^{\perp\perp}$, by

$$S^{\perp\perp} = \{ v \in V : < v, u > = 0 , \forall u \in S^\perp \}$$

Obviously $S^{\perp\perp}$ is a subspace of V.

Note: It is very easy to show that $S \subset S^{\perp\perp}$

Let $u \in S$, then $< u, v > = 0 \ \forall v \in S^\perp$.

So by definition of $S^{\perp\perp}$, we conclude that $u \in S^{\perp\perp}$. So $\subset S^{\perp\perp}$

Theorem 9: (Projection Theorem) Let W be any subspace of a finite dimensional inner product space V. Then **(i)** $V = W \oplus W^{\perp}$ **(ii)** $W^{\perp\perp} = W$

Proof: (**i**) By definition , $W^{\perp} = \{ v \in V : < v, u > = 0 , \forall u \in W \}$, and $W^{\perp}$ is a subspace of V. By the given hypothesis, W is also a finite dimensional inner product space and so W has an orthonormal basis.

Let $S = \{ w_1 , \ldots ,w_m \}$ be an orthonormal basis of W.

$$\therefore < w_i , w_j > = \begin{cases} 0, & \text{if } i \neq j \\ 1 , & \text{if } i = j \end{cases} \qquad \ldots..(1)$$

Let $v \in V$ be arbitrary,

Let $w = \sum_{i=1}^{m} \alpha_i w_i$, where $\alpha_i = < v, w_i >$ …..(2)

Now we assume $x = v - w$ …..(3)

Then,

$$< x, w_i > = < v - w, w_i > = < v, w_i > - < w, w_i >$$

$$= < v, w_i > - < \alpha_1 w_1 + \ldots + \alpha_m w_m , w_i >$$

$$= < v, w_i > - \alpha_1 < w_1 , w_i > - \ldots - \alpha_i < w_i , w_i > - \ldots - \alpha_m < w_m , w_i >$$

$$= < v, w_i > - 0 - \ldots - \alpha_i$$

$$= < v, w_i > - < v, w_i >$$

So, $< x, w_i > = 0$, for $i = 1, 2, \ldots , m$. …..(4)

Since S is a basis of W, each $u \in W$ is expressible as

$$u = \beta_1 w_1 + \beta_2 w_2 + \ldots + \beta_m w_m ; \quad \beta_i \in F$$

We have , $< x, u > = < x, \beta_1 w_1 + \ldots + \beta_m w_m >$

$$= \overline{\beta_1} < x, w_1 > + \ldots + \overline{\beta_m} < x, w_m >$$

$$= \overline{\beta_1} .0 + \ldots + \overline{\beta_m} . 0 = 0, \qquad \text{(using eqn. 4)}$$

So $< x, u > = 0$, $\forall u \in W$

$\Rightarrow x \in W^{\perp}$.

From equation (3) , $v = w + x$ where $w \in W$ and $x \in W^{\perp}$

$\therefore V = W + W^{\perp}$ …..(5)

Now we shall prove that $W \cap W^{\perp} = \{0\}$

Let $y \in W \cap W^{\perp}$ be arbitrary,

$\Rightarrow y \in W$ and $y \in W^{\perp}$

Now $y \in W^{\perp} \Rightarrow < y, u > = 0 \; \forall u \in W$

In particular, $< y, y > = 0$ as $y \in W$

$\Rightarrow y = 0$ and $W \cap W^{\perp} = \{0\}$ …..(6)

From equation (5) and (6) , we get

$V = W \oplus W^{\perp}$

(ii) From part (i) , we have

$V = W \oplus W^{\perp}$ …..(7)

Since $W^{\perp}$ is a subspace of V, on replacing W by $W^{\perp}$ in eq^n (7), we get,

$V = W^{\perp} \oplus W^{\perp\perp}$ …..(8)

As V is finite-dimensional, so from eqns (7) & (8), we get

$\dim V = \dim W + \dim W^{\perp}$ …..(9)

and $\dim V = \dim W^{\perp} + \dim W^{\perp\perp}$

$\Rightarrow \dim W = \dim W^{\perp\perp}$ …..(10)

But we already know that $W \subset W^{\perp\perp}$.

So from equation (10), we have

$$W = W^{\perp\perp}$$

Example 22: If S_1 and S_2 are subsets of an inner product space V, then show that

$$S_1 \subset S_2 \Rightarrow S_2^{\perp} \subset S_1^{\perp}$$

Solution: Let $x \in S_2^{\perp}$, then $< x, y > = 0$, for each $y \in S_2$.

In particular, $< x, z > = 0, \forall z \in S_1$ as $S_1 \subset S_2$

$$\Rightarrow x \in S_1^{\perp}$$

Hence $S_2^{\perp} \subset S_1^{\perp}$

Example 23: If W_1 and W_2 are subspaces of a finite-dimensional inner product space V, then prove that –

(i) $(W_1 + W_2)^{\perp} = W_1^{\perp} \cap W_2^{\perp}$
(ii) $(W_1 \cap W_2)^{\perp} = W_1^{\perp} + W_2^{\perp}$

Solution: Since we know that

$$W_1 \subset W_1 + W_2 \text{ and } W_2 \subset W_1 + W_2$$

So by previous example, we have

$$(W_1 + W_2)^{\perp} \subset W_1^{\perp} \text{ and } (W_1 + W_2)^{\perp} \subset W_2^{\perp}$$

So, $(W_1 + W_2)^{\perp} = W_1^{\perp} \cap W_2^{\perp}$ …..(1)

Now, suppose $z \in W_1^{\perp} \cap W_2^{\perp}$ be arbitrary

$$\Rightarrow z \in W_1^{\perp} \text{ and } z \in W_2^{\perp}$$

$$\Rightarrow < z, x > = 0, \forall x \in W_1 \text{ and } < z, y > = 0, \forall y \in W_2 \quad …..(2)$$

Now any $t \in W_1^{\perp} \cap W_2^{\perp}$ can be written as

$$t = x + y \text{ for some } x \in W_1, y \in W_2$$

So $< z, t > = < z , x + y > = < z , x > + < z , y >$

$= 0$, (using eq^n (2))

So, $z \in (W_1 + W_2)^{\perp}$ and hence

$$W_1^{\perp} \cap W_2^{\perp} \subset (W_1 + W_2)^{\perp} \qquad ....(3)$$

From equation (1) and (3) , we get

$$(W_1 + W_2)^{\perp} = W_1^{\perp} \cap W_2^{\perp} \qquad ....(4)$$

(**ii**) Since $W_1^{\perp}$ and $W_2^{\perp}$ are subspaces of V, so on taking $W_1^{\perp}$ in place of W_1 and $W_2^{\perp}$ in place of W_2 in eq^n (4), we get

$$(W_1^{\perp} + W_2^{\perp})^{\perp} = (W_1^{\perp})^{\perp} \cap (W_2^{\perp})^{\perp}$$

So $(W_1^{\perp} + W_2^{\perp})^{\perp} = W_1^{\perp\perp} \cap W_2^{\perp\perp}$

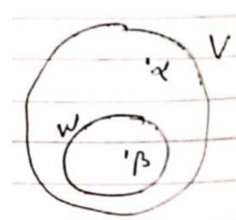
$= W_1 \cap W_2$ as $W^{\perp\perp} = W$

$$\Rightarrow (W_1^{\perp} + W_2^{\perp})^{\perp\perp} = (W_1 \cap W_2)^{\perp}$$

$$\Rightarrow W_1^{\perp} + W_2^{\perp} = (W_1 \cap W_2)^{\perp}$$

Example 24: Let W be a finite-dimensional proper subspace of an inner product space V. Let $\alpha \in V$ and $\alpha \notin W$. Show that there is a vector $\beta \in V$ such that $\alpha - \beta$ is orthogonal to W.

Solution: We know that every finite-dimensional inner product space has an orthonormal basis. Let $\{ \alpha_1 , ..., \alpha_n \}$ be an orthonormal basis of W.



Let $\beta = \sum_{i=1}^{n} < \alpha , \alpha_i > \alpha_i$ where $< \alpha , \alpha_i > \in \mathbf{F}$

Then $\beta \in W$, For each j, $1 \leq j \leq n$ we have

$$< \alpha - \beta , \alpha_j > = < \alpha - \sum_{i=1}^{n} < \alpha , \alpha_i > \alpha_i , \alpha_j >$$

$$= < \alpha , \alpha_j > - \sum_{i=1}^{n} < \alpha , \alpha_i > < \alpha_i , \alpha_j >$$

$$= < \alpha , \alpha_j > - < \alpha , \alpha_j > \quad \text{as } < \alpha_i , \alpha_j > = \delta_{ij}$$

$= 0$

$< \alpha - \beta , \alpha_j > = 0$, for all j =1, 2, …., n. …..(1)

Let $w \in W$ be arbitrary, we can write

$$w = \sum_{i=1}^{n} a_i \alpha_i \quad \text{where } a_i \in \mathbf{F}$$

We have $< \alpha - \beta , w > = < \alpha - \beta , \sum_{i=1}^{n} a_i \alpha_i >$

$= \sum_{i=1}^{n} \overline{a_i} < \alpha - \beta , \alpha_i > = 0$, by eq^n (1)

$\therefore < \alpha - \beta , w > = 0$, for each $w \in W$

Hence $\alpha - \beta$ is orthogonal to W.

12.8 *RIESZ REPRESENTATION THEOREM*

Theorem 10: Let V(R) be a finite-dimensional linear functional $f : V \to \mathbf{R}$. Then there exists a unique $y \in V$ such that $f(x) = < x , y >$, $\forall x \in V$.

Proof: Suppose there exists $y \in V$ such that

$$f(x) = < x , y >, \text{ for all } x \in V.$$

Let us choose an orthonormal basis $\{ e_1 , …., e_n \}$ of V

Then $y = \sum_{i=1}^{n} \alpha_i e_i$ for some $\alpha_i \in \mathbf{R}$

Now $f \in L(V, \mathbf{R})$ and f is completely determined if we know $f(e_i)$ for $1 \leq i \leq n$

Now $f(e_i) = < e_i , y > = \alpha_i$ for $1 \leq i \leq n$

This suggest that we take $y = \sum_{i=1}^{n} f(e_i) e_i$

It is easy to check that $f(x) = < x , y >$ for all $x \in V$

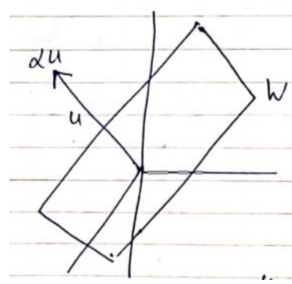
For if $x = \sum \alpha_i e_i$, then $f(x) = \sum \alpha_i f(e_i)$ …..(1)

Also $< x , y > = < x , \sum f(e_i)\ e_i >$

$= < \sum \alpha_j e_j , \sum f(e_i)\ e_i >$

$= \sum_{i,j} f(e_i)\ \alpha_j < e_i , e_j >$ …..(2)

$= \sum f(e_i)\ \alpha_i$ as $< e_i , e_j > = \delta_{ij} = \begin{cases} 1 , i = j \\ 0 , i \neq j \end{cases}$



From equations (1) and (2) , we conclude that

$f(x) = < x , y >$ for all $x \in R^n$

Uniqueness: Now, suppose z is such that,

$f(x) = < x , z >$ for all $x \in V$

then, $f(x) = < x , z > = < x , y >$

$\Rightarrow < x , z - y > = 0$ for all x.

In particular, for $x = z - y$, we obtain

$< z - y , z - y > = 0$

$z - y = 0$

$z = y$

So y is unique.

Geometric Interpretation:

If f = 0, then the obvious choice is y = 0.

If $f \neq 0$, then f is a linear form and W = ker f is of

dimension n – 1, where n = dim V.

Thus there is a unit vector u perpendicular to W, for

$V = W \oplus W^{\perp}$ (that is , u is a unit normal to the "plane"). y must therefore be a multiple αu of u. The choice of α is determined by the equation

$f(u) = < u , y > = < u , \alpha u > = \alpha$

Thus we take $y = \alpha u$ where $\alpha = f(u)$

For $x \in V$, we have $x = w + tu$, where $w \in W$ and $t \in \mathbf{R}$

Then $f(x) = f(w + tu) = f(w) + t f(u) = t f(u)$

Also $< x, y > = < w + tu, \alpha u > = \alpha < w, u > + t \alpha < u , u > = t \alpha = t f(u)$

Hence the result.

Theorem 11: For any linear operator T on a finite-dimensional inner product space V, there exists a unique linear operator T^* on V such that

$$< T\alpha , \beta > = < , T^* \beta > \text{ for all } \alpha , \beta \in V.$$

Proof: Let T be a linear operator on a finite dimensional inner product space V over the field **F**. Let $\beta \in V$ and f be a functional from V into **F** defined by

$$f(\alpha) = < T\alpha , \beta > \quad \forall \alpha \in V \qquad \ldots..(1)$$

Here $T\alpha$ stands for $T(\alpha)$

Claim: f is a linear functional on V.

Let a , b $\in$ **F** and $\alpha_1 , \alpha_2 \in V$, then

$$f (a \alpha_1 + b \alpha_2) = < T(a \alpha_1 + b \alpha_2) , \beta >$$

$$= < (aT\alpha_1 + bT\alpha_2) , \beta > \quad \text{as T is linear}$$

$$= a < T\alpha_1 , \beta > + b < T\alpha_2 , \beta >$$

$$= a f (\alpha_1) + b f (\alpha_2) , \text{ using equation (1)}$$

Hence f is a linear functional on V.

So by **Riesz representation theorem**, there exists a unique $\beta' \in V$ such that

$$f(\alpha) = <\alpha, \beta'> \quad \forall \alpha \in V \qquad \text{.....(2)}$$

From equations (1) and (2), we observe that if T is a linear operator on V, then corresponding to every vector β in V, there is a uniquely determined vector β' in V such that

$$< T\alpha, \beta > = < \alpha, \beta' > \quad \forall \alpha \in V$$

Let us denote by T^* the rule which associates β with β' i.e. let $T^* \beta = \beta'$

Then T^* is a function from V in to V and is such that

$$< T\alpha, \beta > = < \alpha, T^* \beta > \quad \forall \alpha, \beta \in V \qquad \text{.....(3)}$$

Claim: T^* is a linear operator on V.

Let a, b $\in$ **F** and $\beta_1, \beta_2 \in V$. Then $\forall \alpha \in V$, we have

$$< \alpha, T^*(a\beta_1 + b\beta_2) > = < T\alpha, a\beta_1 + b\beta_2 > \qquad \text{using equation (3)}$$

$$= \bar{a} < T\alpha, \beta_1 > + \bar{b} < T\alpha, \beta_2 >$$

$$= \bar{a} < \alpha, T^*\beta_1 > + \bar{b} < \alpha, T^*\beta_2 > \qquad \text{again by (3)}$$

$$= < \alpha, aT^*\beta_1 > + < \alpha, bT^*\beta_2 >$$

$$= < \alpha, aT^*\beta_1 + bT^*\beta_2 >$$

Hence $T^*(a\beta_1 + b\beta_2) = aT^*\beta_1 + bT^*\beta_2$

Thus T^* is a linear operator on V

Hence corresponding to a linear operator T on V, there exists a linear operator T^* on V

such that, $$< T\alpha, \beta > = < \alpha, T^* \beta > \quad \forall \alpha, \beta \in V$$

Uniqueness: Let S be a linear operator on V such that

$$< T\alpha, \beta > = < \alpha, S\beta > \quad \forall \alpha, \beta \in V$$

Then $< \alpha , T^* \beta > = < \alpha , S \beta > \quad \forall \alpha, \beta \in V$

$\Rightarrow \qquad T^* \beta = S \beta \qquad \forall \beta \in V$

$\Rightarrow \qquad T^* = S$

So T^* is unique.

Check your progress

Problem 1: Let V be a vector space of all real polynomials of degree ≤ 2, with inner product

$$< f(x) , g(x) > = \int_{-1}^{1} f(x)g(x)dx , \forall f(x) , g(x) \in V$$

If $f(x) = x^2 + x – 4$ and $g(x) = x – 1$, then find

(i) $< f(x) , g(x) >$ and

(ii) $< g(x) , g(x) >$

Problem 2: Prove that $\| \alpha v \| = |\alpha| \| v \|$, for all $\alpha \in \mathbf{F}$, $x \in V$

Problem 3: If $\{v_1, v_2, \ldots, v_n\}$ is an orthonormal set and if $w = \sum_{i=1}^{n} \alpha_i v_i \in V$, Then prove that $\alpha_i = < w, v_i >$ for $i = 1, 2, \ldots, n$.

12.9 *SUMMARY*

In this chapter we understood the process of generalization from ordinary vectors to vector spaces. So other basic concepts viz angle, length, distance were also generalized respectively as inner product, norm, and metric. As we have studied orthogonal compotent of ordinary vectors, we studied here Gram-Schmidt orthogonalisation process. Besides this, we learned various concepts and applications of inner product.

12.10 *GLOSSARY*

➢ **Inner Product:** An inner product on a vector space V is a map $< \, , \, > : V \times V \rightarrow R$ satisfying the following properties :

(i) $< x , x > \geq 0$ and $< x , x > = 0$ if and only if $x = 0$.

(ii) $< x , y > = < y , x >$

(iii) $\langle x + z , y \rangle = \langle x , y \rangle + \langle z , y \rangle$ and $\langle x , y + z \rangle = \langle x , y \rangle + \langle x , z \rangle$

(iv) $\langle ax , y \rangle = a \langle x , y \rangle \ \forall x , y , z \in V$ and $a \in R$.

- **Norm of a Vector:** Let V be an inner product space. The norm function $\| . \| : V \to R$ has the following properties :

(i) $\| x \| \geq 0$ and $\| x \| = 0$ if and only if $x = 0$; $x \in V$

(ii) $\| \alpha x \| = |\alpha| \, \| x \|$, $\alpha \in \mathbf{F}$, $x \in V$,

Norm of a vector $v \in V$ is defined as $\| v \| = \sqrt{< v, v >}$.

- **Complete Orthonormal Set:** An orthonormal set is said to be complete if it is not contained in any larger orthonormal set.
- **Gram-Schmidt orthogonlisation Process:** Every finite-dimensional inner product space has an orthonormal basis.

12.11 *REFERENCES*

- S Kumaresan; Linear Algebra-A Geometric Approach; PHI-2016.
- K Hoffman & Ray Kunze; Linear Algebra; PHI-1971.
- P Halmos; Finite-Dimensional Vector Spaces;
- G Strang; Introduction To Linear Algebra;

12.12 *SUGGESTED READING*

- NPTEL videos.
- Schaum series.
- A R Vashishtha, Krishna Prakashan; Meerut.
- Graduate Text In Mathematics, Springer.

12.13 *TERMINAL QUESTION*

1: Prove that for any $\alpha \in \mathbf{R}^2$, we can write $\alpha = \langle \alpha , e_1 \rangle e_1 + \langle \alpha , e_2 \rangle e_2$ where $e_1 = (1, 0)$, $e_2 = (0, 1)$

2: Let V be a vector space over a field **F**. Let W_1 and W_2 be twp subspaces of **V(F)** such that W_1 and W_2 are two inner product spaces also. Then prove that –

(i) A positive multiple of an inner product is also an inner product.

(ii) Difference of two inner products may **not** be an inner product.

3: Let V (R) be a vector space of polynomials with inner product defined by

$$< f , g > = \int_0^1 f(x)g(x)dx$$

If $f(x) = x^2 + 1$ and $g(x) = x – 1$, then find < f, g > and ‖ g ‖ .

12.14 *ANSWERS*

Answers of check your progress:

1. (i) 8 (ii) 8/3

3. Given that $\{v_1, v_2, \ldots, v_n\}$ is an orthonormal set. So

$$< v_i , v_j > = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases} \qquad \ldots..(1)$$

We have , $< w, v_i > = (\alpha_1 v_1 + \ldots.. + \alpha_n v_n , v_i)$

$= \alpha_1 < v_1 , v_i > + \ldots. + \alpha_i < v_i , v_i > + \ldots.+ \alpha_n < v_n , v_i >$

$= 0 +\ldots.+ \alpha_i + 0 + \ldots. + 0$

$< w, v_i > = \alpha_i$, for i = 1, 2, …., n.

Answers of terminal question:

2. (**i**) Let < u, v > be an inner product and $\lambda > 0, \lambda \in$ **R**. Then it can be easily verified that λ< u, v > is also an inner product.

(ii) Difference of two inner products may not be positive. Now do it yourself.

3. $< f , g > = \frac{-7}{12}$ and $\| g \| = \frac{1}{\sqrt{3}}$.

UNIT-13: OPERATORS

CONTENTS

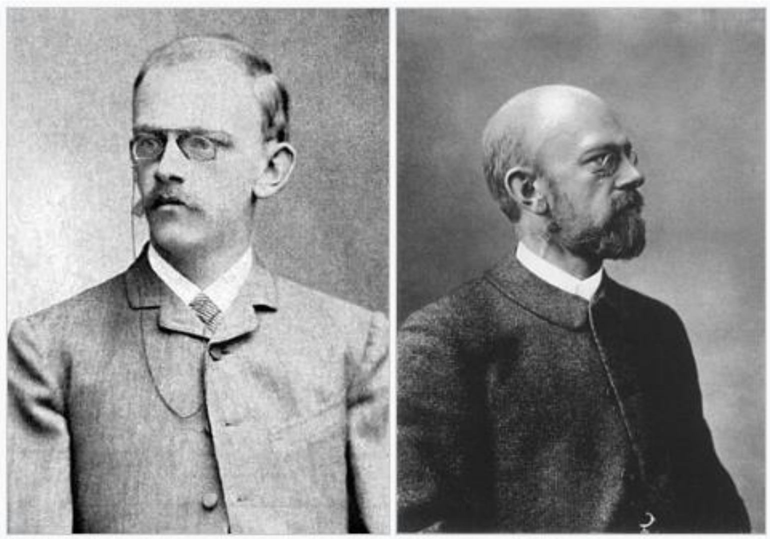
13.1 *INTRODUCTION*

German mathematician David Hilbert, who lived from January 23, 1862, to February 14, 1943, was a very influential mathematician of the late 19th and early 20th centuries. The foundations of geometry, the spectral theory of operators and its application to integral equations, the calculus of variations, commutative algebra, algebraic number theory, mathematical physics, and the foundations of mathematics (especially proof theory) are just a few of the many fundamental concepts that Hilbert discovered and developed.

Hilbert embraced and upheld the transfinite numbers and set theory of Georg Cantor. He introduced a set of issues in 1900 that paved the way for 20th-century mathematical research.

Important tools utilized in modern mathematical physics were invented by Hilbert and his pupils, who also helped to establish rigor in the field. Hilbert was a pioneer in the fields of mathematical logic and proof theory.

An inner product structure on a C-vector spaces induces a "mirrored" twin for every linear transformation, called the adjoint. Linear operators equal their own adjoints have many important properties.



Hilbert in 1886 Hilbert in 1907

13.2 *OBJECTIVES*

After reading this unit learners will be able to

- Understand the basic concept of unitary operator and normal operator.
- Understand the basic concept of adjoint operator and self-adjoint operator.
- Understand the concept of skew-symmetric and skew-Hermitian operator.
- Understand the concept of positive and non-negative operator.

13.3 *ADJOINT OPERATORS*

Let T be a linear operator on an inner product space V (here V **need not** be finite dimensional). We say that T has an adjoint T^* if there exists a linear operation T^* in V

such that $< T\alpha , \beta > = < \alpha , T^* \beta > \quad \forall\, \alpha, \beta \in V$

Note: In previous unit, we have proved that every linear operator on a finite-dimensional inner product space posses an adjoint. But it should be noted that if V is **not** finite-dimensional, then some linear operator on V may possess an adjoint while the other may not. In any case if T possesses an adjoint T^*, then it must be unique. Also observe that the adjoint of T depends not only upon T, but also on the inner product on V.

Theorem 1: Let V be a finite-dimensional inner product space and let B = { α_1 , …., α_n } be an ordered orthonormal basis for V . Let T be a linear operator on V and let A = $[a_{ij}]_{m \times n}$ be the matrix of T with respect to the ordered basis B. Then $a_{ij} = < T\alpha_j, \alpha_i >$

Proof: As B is an orthonormal basis for V, so for any $\beta \in V$,

$$\beta = \sum_{i=1}^{n} < \beta, \alpha_i > \alpha_i$$

Replacing β by T α_j , we get

$$T\alpha_j = \sum_{i=1}^{n} < T\alpha_j , \alpha_i > \alpha_i \; ; \quad j = 1, 2, \ldots, n \qquad \ldots(1)$$

Now if A = $[a_{ij}]_{m \times n}$ be the matrix of T in the ordered basis B, then we have

$$T\alpha_j = \sum_{i=1}^{n} a_{ij}\, \alpha_i \; ; \quad j = 1, 2, \ldots, n \qquad \ldots(2)$$

Since the expression for $T\alpha_j$ as a linear combination of vectors in B is unique, so from equations (1) and (2), we have

$$a_{ij} = < T\alpha_j , \alpha_i > \quad ; 1 \le i \le n, 1 \le j \le n$$

Corollary 1: Let V be a finite dimensional inner product space and let T be a linear operator on V. In any orthonormal basis for V, the matrix of T^* is the conjugate transpose of the matrix of T.

Proof: Let B = { α_1 , …., α_n } be an orthonormal basis for V. Let A = $[a_{ij}]_{m \times n}$ be the matrix of T in ordered basis B.

Then $\quad a_{ij} = < T\alpha_j , \alpha_i > \qquad \ldots(1)$

Now T^* is also a linear operator on V.

Let $C = [c_{ij}]_{n \times n}$ be the matrix of T^* in the ordered basis B.

Then $c_{ij} = < T^* \alpha_j, \alpha_i >$ …..(2)

We have $c_{ij} = < T^* \alpha_j, \alpha_i > = < \overline{\alpha_i, T^*\alpha_j} >$

$= < \overline{T\alpha_i, \alpha_j} >$ by definition of T^*

$= \overline{a_{ji}}$

So $C = [\overline{a_{ji}}]_{n \times n}$ and hence $C = A^*$, where A^* is the conjugate transpose of A.

Note: It should be remembered that in this corollary the basis B is an orthonormal basis and **not** an ordinary basis.

Theorem 2: Let S and T be linear operators on an inner product space V and c ∈ **F**. If S and T possess adjoints, the operators S + T, cT, ST, T^* will possess adjoints.

Also (**i**) $(S + T)^* = S^* + T^*$

(**ii**) $(cT)^* = \bar{c}\, T^*$

(**iii**) $(S\,T)^* = T^*\, S^*$

(**iv**) $(T^*)^* = T$

Proof: (**i**) As S and T are linear operators on V, so S + T is also a linear operator on V.

Now for every $\alpha, \beta \in V$, we have

$< (S + T)\,\alpha , \beta > = < S\alpha + T\alpha , \beta > = < S\alpha , \beta > + < T\alpha , \beta >$

$= < \alpha , S^*\beta > + < \alpha , T^* \beta >$, by definition of adjoint

$= < \alpha , S^*\beta + T^* \beta >$

$= < \alpha , (S^* + T^*)\, \beta >$

Thus for the linear operator S + T on V there exists a linear operator $S^* + T^*$ on V such that

$$< (S + T)\alpha , \beta > = < \alpha , (S^* + T^*)\beta > \text{ for all } \alpha, \beta \in V$$

Therefore, the linear operator S + T has an adjoint. By the definition and by the uniqueness of adjoint, we get

$$(S + T)^* = S^* + T^*$$

(ii) Since T is a linear operator on V, therefore cT is also a linear operator on V. For every $\alpha, \beta \in V$, we have

$$< (cT)\alpha , \beta > = < cT\alpha , \beta > = c < T\alpha , \beta > = c < \alpha , T^*\beta >$$

$$= < \alpha , \bar{c}T^*\beta > = < \alpha , (\bar{c}T^*)\beta >$$

$$< (cT)\alpha , \beta > = < \alpha , (cT)^*\beta >$$

Thus for the linear operator cT on V , ∃ a linear operator $(cT)^*$ or $\bar{c}T^*$ on V such that

$$< (cT)\alpha , \beta > = < \alpha , (cT)^*\beta > \forall \alpha, \beta \in V.$$

Hence the linear operator cT possesses an adjoint. By the definition and by the uniqueness of adjoint, we get

$$(cT)^* = \bar{c}T^*$$

(iii) We observe that ST is a linear operator on V

Now $\forall \alpha, \beta \in V$, we have

$$< (ST)\alpha , \beta > = < ST\alpha , \beta >$$

$$= < T\alpha , S^*\beta > \qquad \text{by definition of adjoint}$$

$$= < \alpha , T^*S^*\beta >$$

$$= < \alpha , (T^*S^*)\beta >$$

Thus for the linear operator ST on V ∃ a linear operator T^*S^* on V such that

$$< (ST)\alpha , \beta > = < \alpha , (T^*S^*)\beta > \forall \alpha, \beta \in V$$

Therefore, the linear operator ST has an adjoint. By the definition and by the uniqueness of adjoint, we get $(S\,T)^* = T^*\,S^*$

(iv) The adjoint of T i.e. T^* is a linear operator on V. For every $\alpha, \beta \in V$, we have

$$< T^* \alpha , \beta > = \overline{< \beta,\ T^*\alpha >}$$

$$= \overline{< T\,\beta,\ \alpha >}$$

$$= < \alpha ,T\,\beta >$$

Thus for the linear operator T^* on V, there **exists** a linear operator T on V such that

$$< T^* \alpha , \beta > = < \alpha ,T\,\beta > \text{ for all } \alpha, \beta \in V$$

Therefore, the linear operator T^* has an adjoint. By the definition and by the uniqueness of adjoint, we have $(T^*)^* = T$

Note: (1) If V is a finite-dimensional inner product space, then the result is true for arbitrary linear operators S and T. In a finite-dimensional inner product space, each linear operator possesses and adjoint.

(2) The operation of adjoint behaves like the operation of conjugation on complex numbers.

13.4 *SELF-ADJOINT OPERATORS*

Self-adjoint transformation: A linear operator T on an inner product space V is said to be self-adjoint if $T^* = T$

A self-adjoint linear operator on a real inner product space is called **symmetric** while a self-adjoint linear operator on a complex inner product space is called **Hermitian**.

e.g. the zero operator $\hat{0}$ and the identity operator I on *any* inner product space V are self-adjoint. For every $\alpha, \beta \in V$, we have

$$< \hat{0}\,\alpha , \beta > = < 0 , \beta > = 0 = < \alpha , 0 > = < \alpha , \hat{0}\ \beta >$$

So $\hat{0}^* = \hat{0}$

Similarly, $< I\,\alpha , \beta > = < \alpha , \beta > = < \alpha , I\,\beta >$

So $I^* = I$

13.5 *SKEW-SYMMETRIC/ SKEW-HERMITION OPERATORS*

Skew-symmetric / skew-Hermitian operator: If a linear operator T on an inner product space V is such that $T^* = -T$

then T is called *skew-symmetric* or *skew-Hermitian* according as the vector space V is real or complex.

Theorem 3: Every linear operator T on a finite dimensional complex inner product space V can be **uniquely** expressed as

$T = T_1 + iT_2$, where T_1 & T_2 are self-adjoint linear operators on V.

Proof: Let $T = \frac{1}{2}(T + T^*) + i\left(\frac{T - T^*}{2i}\right)$

Suppose $T_1 = \frac{T + T^*}{2}$ and $T_2 = \frac{T - T^*}{2i}$

So, $T = T_1 + iT_2$ …..(1)

Now $T_1^* = \left(\frac{T + T^*}{2}\right)^* = \frac{1}{2}(T^* + (T^*)^*) = \frac{1}{2}(T^* + T) = T_1$

So T_1 is self-adjoint

Again $T_2^* = \left[\frac{1}{2i}(T - T^*)\right]^* = \left(\overline{\frac{1}{2i}}\right)(T - T^*)^* = \frac{1}{(-2i)}(T^* - T)$

$$T_2^* = \frac{1}{2i}(T - T^*)$$

So T_2 is also self-adjoint. Thus T can be expressed as a sum of two self-adjoint operators.

Uniqueness: Let $T = U_1 + iU_2$ where U_1 and U_2 are both self-adjoint linear operators.

So, $T^* = (U_1 + iU_2)^* = U_1^* + \bar{i}\, U_2^* = U_1^* - i\, U_2^* = U_1 - i\, U_2$

So $T + T^* = 2U_1$ or $U_1 = \frac{1}{2}(T + T^*) = T_1$

Similarly, $T - T^* = 2i\, U_2$ or $U_2 = \frac{1}{2i}(T - T^*) = T_2$

So $T = T_1 + iT_2 = U_1 + iU_2$ i.e. representation is unique.

Note: If T is linear operator on a complex inner product space V which is **Not** finite dimensional, then the above result will be still true *provided*, it is given that T possesses adjoint.

Theorem 4: Every linear operator T on a finite-dimensional inner product space V can be uniquely expressed as $T = T_1 + T_2$, where T_1 is self-adjoint and T_2 is skew.

Proof: Let $T = \frac{1}{2}(T + T^*) + \frac{1}{2}(T - T^*)$

where $T_1 = \frac{1}{2}(T + T^*)$ and $T_2 = \frac{1}{2}(T - T^*)$

then $T = T_1 + T_2$ …..(1)

Now $T_1^* = [\frac{1}{2}(T + T^*)]^* = \frac{1}{2}(T + T^*)^* = \frac{1}{2}(T^* + T) = T_1$

So T_1 is self-adjoint.

Similarly $T_2^* = [\frac{1}{2}(T - T^*)]^* = \frac{1}{2}(T - T^*)^* = \frac{1}{2}(T^* - T)$

$$T_2^* = -\frac{1}{2}(T - T^*) = -T_2$$

So T_2 is skew.

Hence T can be expressed as a sum of two linear operators where T_1 in self-adjoint and T_2 in skew.

Uniqueness: Let $T = U_1 + U_2$, where U_1 is self-adjoint and U_2 in skew.

Then $T^* = (U_1 + U_2)^* = U_1^* + U_2^* = U_1 - U_2$

So $T + T^* = 2U_1$ or $U_1 = \frac{1}{2}(T + T^*) = T_1$

and $T - T^* = 2U_2$ or $U_2 = \frac{1}{2}(T - T^*) = T_2$

Hence $T = T_1 + T_2 = U_1 + U_2$

$\Rightarrow$ The expression (1) for T is unique.

Note: If T is a linear operator on an inner product space V which is **NOT** finite-dimensional, then the above result will be still true *provided* T possesses adjoint.

Theorem 5: A necessary and sufficient condition that a linear transformation T on an inner product space V be $\hat{0}$ is that $< T\alpha, \beta > = 0, \forall \alpha, \beta \in V$

Proof: Necessary condition: Let T = 0, then $\forall \alpha, \beta \in V$, we have

$$< T\alpha, \beta > = < \hat{0}\alpha, \beta > = < 0, \beta > = 0$$

So the condition is necessary.

Sufficient condition: Let T be a linear operator such that

$$< T\alpha, \beta > = 0, \forall \alpha, \beta \in V$$

Taking $\beta = T\alpha$, we get

$$< T\alpha, T\alpha > = 0 \quad \forall \alpha \in V$$

$$\Rightarrow \quad T\alpha = 0 \quad \forall \alpha \in V$$

$$\Rightarrow \quad T = \hat{0}$$

Hence the condition is sufficient.

Theorem 6: A necessary and sufficient condition that a linear transformation T on a unitary space be $\hat{0}$ is that $< T\alpha, \alpha > = 0 \quad \forall \alpha \in V$

Proof: Necessary condition: Let $T = \hat{0}$, then $\forall \alpha \in V$

$$< T\alpha, \alpha > = < \hat{0}\alpha, \alpha > = < 0, \alpha > = 0$$

Hence the condition is necessary.

Sufficient condition: Let T be a linear operator satisfying

$$< T\alpha, \alpha > = 0 \quad \forall \alpha \in V \qquad \text{.....(1)}$$

Replacing α by $\alpha + \beta$, we get

$$< T(\alpha + \beta), \alpha + \beta > = 0$$

$\Rightarrow \quad < T\alpha + T\beta, \alpha + \beta > = 0$

$\Rightarrow \quad < T\alpha , \alpha > + < T\alpha , \beta > + < T\beta , \alpha > + < T\beta, \beta > = 0$

$\Rightarrow \quad < T\alpha , \beta > + < T\beta , \alpha > = 0 ,$ using (1)

So $\forall\ \alpha, \beta \in V$, we have

$< T\alpha , \beta > + < T\beta , \alpha > = 0$ …..(1)

Since above result is true $\forall\ \beta \in V$, so by replacing β and $i\beta$, we get

$< T\alpha , i\beta > + < T\, i\beta , \alpha > = 0$

$\bar{\imath} < T\alpha , \beta > + i < T\beta , \alpha > = 0$

$-i < T\alpha , \beta > + i < T\beta , \alpha > = 0$

$\Rightarrow \quad - < T\alpha , \beta > + < T\beta , \alpha > = 0$ ….(2)

Adding equation (1) and (2), we get

$2 < T\beta , \alpha > = 0$

$< T\beta , \alpha > = 0 \ \forall\ \alpha, \beta \in V$

Let $\alpha = T\beta$, then

$< T\beta, T\beta > = 0 \quad \forall\ \beta \in V$

$\Rightarrow \quad T\beta = 0 \quad \forall\ \beta \in V$

$\Rightarrow \quad T = \hat{0}$

Hence the condition in sufficient.

Note: (**1**) Above result **may fail** for Eulidean space, e.g., let us consider $V_2(R)$ with standard inner product space. Let T be a linear operator on $V_2(R)$ defined as

$T(a, b) = (b, -a) \qquad \forall\ (a, b) \in V_2(R)$

Then obviously $T \neq \hat{0}$. But

$$< T(a, b), (a, b) > = < (b, -a), (a, b) >$$

$$= ba - ab = 0$$

So $< T\alpha, \alpha > = 0 \quad \forall \alpha \in V_2(R)$, through $T \neq \hat{0}$.

(**2**) However if T is self-adjoint then the above theorem is true for Euclidean spaces also. Finally, we have the following theorem –

Theorem 7: A necessary and sufficient condition that a self-adjoint linear transformation T on an inner product space V be $\hat{0}$ is that

$$< T\alpha, \alpha > = 0 \text{ , for all } \alpha \in V$$

Proof: Necessary part is same as in previous theorem.

Sufficient condition: Let $< T\alpha, \alpha > = 0 \quad \forall \alpha \in V$

So $< T(\alpha + \beta), \alpha + \beta > = 0 \quad \forall \alpha, \beta \in V$

$$\Rightarrow \quad < T\alpha + T\beta, \alpha + \beta > = 0$$

$$< T\alpha, \alpha > + < T\alpha, \beta > + < T\beta, \alpha > + < T\beta, \beta > = 0$$

$$\Rightarrow \quad < T\alpha, \beta > + < T\beta, \alpha > = 0$$

$$\Rightarrow \quad < T\alpha, \beta > + < \beta, T^* \alpha > = 0$$

$$\Rightarrow \quad < T\alpha, \beta > + < \beta, T\alpha > = 0 \text{ , as given } T = T^* \qquad \text{.....(1)}$$

Now two cases may arise –

Case I: If V is a complex inner product space. Then do as in previous theorem.

Case II: If V is a real inner product space.

Then $< \beta, T\alpha > = < T\alpha, \beta >$ as $< \alpha, \beta > = < \overline{\beta, \alpha} > = < \beta, \alpha >$

So from equation (1), we have

$$2 < T\alpha , \beta > = 0 \text{ or } < T\alpha , \beta > = 0 \ \forall \ \alpha, \beta \in V$$

Let us put $\beta = T\alpha$

$$\Rightarrow < T\alpha , T\alpha > = 0 \ \forall \ \alpha \in V$$

$$\Rightarrow \quad T\alpha = 0 \ \forall \ \alpha \in V$$

$$\Rightarrow \quad T = \hat{0}$$

Theorem 8: A necessary and sufficient condition that a linear transformation T on a unitary space (of any dimension) be self-adjoint (Hermitian) is that,

$$< T\alpha , \alpha > \text{ be real } \ \forall \ \alpha \in V$$

Proof: Necessary condition: Let T be self-adjoint operator on a unitary space V i.e. $T^* = T$.

Then for every $\alpha \in V$, we have

$$< T\alpha , \alpha > = < \alpha , T^* \alpha > = < \alpha , T\alpha > = < \overline{T\alpha , \alpha} >$$

$$\Rightarrow \quad < T\alpha , \alpha > \text{ is real } \ \forall \ \alpha \in V$$

Sufficient condition: Let $< T\alpha , \alpha >$ be real $\forall \ \alpha \in V$. We have to prove that $T^* = T$. For every $\alpha, \beta \in V$, we have

$$< T (\alpha + \beta), \alpha + \beta > = < T\alpha + T\beta, \alpha + \beta >$$

$$< T (\alpha + \beta), \alpha + \beta > = < T\alpha , \alpha > + < T\alpha , \beta > + < T\beta , \alpha > + < T\beta, \beta > \quad ....(1)$$

Since $< T (\alpha + \beta), \alpha + \beta >$, $< T\alpha , \alpha >$ and $< T\beta , \beta >$ are real.

$$\Rightarrow \quad < T\alpha , \beta > + < T\beta , \alpha > \text{ must be real}$$

So $\quad < T\alpha , \beta > + < T\beta , \alpha > = \overline{< T\alpha , \beta > + < T\beta , \alpha >}$

$$= < \overline{T\alpha , \beta} > + < \overline{T\beta , \alpha} >$$

$$= < \beta , T\alpha > + < \alpha , T\beta >$$

So $\forall\ \alpha,\ \beta\ \in V$, we have

$$< T\alpha , \beta > + < T\beta , \alpha > \ = < \beta , T\alpha > + \ < \alpha , T\beta > \qquad \text{....(2)}$$

Replacing β by $i\beta$ in equation (2), we get

$$< T\alpha ,\ i\beta\ > + < T\, i\beta\ , \alpha > \ = <\ i\beta\ , T\alpha > + \ < \alpha , T\, i\beta\ >$$

$$\bar{\imath} < T\alpha , \beta > + i < T\beta , \alpha > \ = \ i < \beta , T\alpha > + \bar{\imath} < \alpha , T\beta >$$

$$-i < T\alpha , \beta > + i < T\beta , \alpha > \ = \ i < \beta , T\alpha > \ - i < \alpha , T\beta >$$

$$- < T\alpha , \beta > + \ < T\beta , \alpha > \ = \ \ < \beta , T\alpha > \ - \ < \alpha , T\beta > \qquad \text{...(3)}$$

on equation(2) – equation(3), we get

$$< T\alpha , \beta > = < \alpha , T\beta >$$

$$< T\alpha , \beta > = < \alpha , T\beta >$$

$$< T\alpha , \beta > = < T^{*}\ \alpha , \beta > \ \forall\ \alpha,\ \beta\ \in V$$

$$\Rightarrow \qquad T = T^{*}$$

Note: If V is finite-dimensional, then we can take *advantage* of the fact that T must possess adjoint. So in this case, the **converse** part of the theorem can be easily proved as:

Since $< T\alpha , \alpha >$ is real $\forall\ \alpha \in V$

So, $\quad < T\alpha , \alpha > \ = < \overline{T\alpha , \alpha} > = < \overline{\alpha , T^{*}\alpha} > = < T^{*}\alpha , \alpha >$

$$\Rightarrow \qquad < T\alpha - T^{*}\alpha,\ \alpha > \ = 0\ \ \forall\ \alpha$$

$$\Rightarrow \qquad < (T - T^{*})\alpha , \alpha > \ = 0\ \ \forall\ \ \alpha \in V \qquad \text{(by previous theorem)}$$

$$\Rightarrow \qquad T - T^{*} = \hat{0}\ \text{ or } \ T = T^{*}$$

Example 1: Let V = $V_2(C)$ with standard inner product. Let T be the linear operator defined by

T(1, 0) = (1, -2) and T(0, 1) = (i, -1)

If $\alpha = (a, b) \in V_2(C)$, then find $T^*\alpha$

Solution: Obviously B = {(1,0), (0, 1)} is an orthonormal basis of V. Let us find $[T]_B$ i.e.

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

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

$$\therefore \quad [T]_B = \begin{bmatrix} 1 & i \\ -2 & -1 \end{bmatrix} \Rightarrow [T^*]_B = \begin{bmatrix} 1 & -2 \\ -i & -1 \end{bmatrix}$$

Now, (a, b) = a(1, 0) + b(0, 1). So coordinate matrix of T^* (a, b) in B is

$$= \begin{bmatrix} 1 & -2 \\ -i & -1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} a & -2b \\ -ia & -b \end{bmatrix},$$

$$T^*(a, b) = (a - 2b)(1, 0) + (-ia - b)(0, 1) = (a - 2b, -ia - b)$$

Example 2: A linear operator on R^2 is defined by

$$T(x, y) = (x + 2y, x - y)$$

Find the adjoint T^*, if the inner product is standard one.

Solution: Let B = {(1, 0) , (0, 1)} be an orthonormal basis of V, We find $[T]_B$. By given rule.

$$T(1, 0) = (1, 1) \text{ and } t(0, 1) = (2, -1).$$

So $[T]_B = \begin{bmatrix} 1 & 2 \\ 1 & -1 \end{bmatrix}$

The matrix of T^* in the ordered basis B is the transpose of the matrix $[T]_B$.

So $[T^*]_B = \begin{bmatrix} 1 & 1 \\ 2 & -1 \end{bmatrix}$

The coordinate matrix of $T^*(x, y)$ in the basis B

$$= \begin{bmatrix} 1 & 1 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x + y \\ 2y - y \end{bmatrix}$$

So $T^*(x, y) = (x + y, 2x - y)$

Example 3: Let T be a linear operator on $V_2(C)$ defined by

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

Using the standard inner product –

(i) Find the matrix of T^* in the standard ordered basis

(ii) Does T commute with T^* ?

Solution: (i) $T(1, 0) = (1 + i, 2) = (1 + i) (1, 0) + 2(0, 1)$

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

So $[T]_B = \begin{bmatrix} 1+i & i \\ 2 & i \end{bmatrix}$

Then $[T^*]_B = \begin{bmatrix} 1-i & 2 \\ -i & -i \end{bmatrix}$

(ii) $[T]_B [T^*]_B = \begin{bmatrix} 1+i & i \\ 2 & i \end{bmatrix} \begin{bmatrix} 1-i & 2 \\ -i & -i \end{bmatrix} = \begin{bmatrix} 3 & 3+2i \\ 3-2i & 5 \end{bmatrix}$

$[T^*]_B\ [T]_B = \begin{bmatrix} 1-i & 2 \\ -i & -i \end{bmatrix} \begin{bmatrix} 1+i & i \\ 2 & i \end{bmatrix} = \begin{bmatrix} 6 & 3i+1 \\ -3i+1 & 2 \end{bmatrix}$

Since $[T]_B [T^*]_B \neq [T^*]_B\ [T]_B$

$\Rightarrow \quad [T\ T^*]_B \neq [T^*\ T]_B$

So $\quad T\ T^* \neq T^*\ T$

Example 4: Prove that the product of two self-adjoint operators on an inner product space is self-adjoint iff the two operators commute.

Solution: Let T and S be two self-adjoint operators s.t. $\quad T^* = T$ and $S^* = S$

IF PART: Let T and S commute i.e. TS = ST

Now, $\quad (TS)^* = S^*\ T^*$

$= S\ T$

$= T\ S$

So TS is also self-adjoint.

ONLY IF PART: Let ST be self-adjoint

$(ST)^* = ST$

$\Rightarrow \quad T^* S^* = ST$

$\Rightarrow \quad T S = S T$

i.e. S and T commute

Example 5: Let $\forall\ \alpha,\ \beta \in V$ and T is a linear transformation on V. Also if $f(\alpha) = <\overline{\beta, T\alpha}>, \forall\ \alpha \in V$, then prove that f is a linear functional. Also find a vector β' such that $f(\alpha) = <\alpha, \beta'> \ \forall\ \alpha \in V$

Solution: (**i**) Given that $f(\alpha) = <\overline{\beta, T\alpha}>, \forall\ \alpha \in V$

So f is a function from V into F. Let a, b $\in$ V and α_1 , $\alpha_2 \in V$. Then

$$f(a\alpha_1 + b\,\alpha_2) = <\overline{\beta, T(a\alpha_1 + b\,\alpha_2)}> = <T(a\alpha_1 + b\,\alpha_2), \beta>$$

$$= a <T\alpha_1, \beta> + b <T\alpha_2, \beta>$$

$$= a <\overline{\beta, T\alpha_1}> + <\overline{\beta, T\alpha_2}> = a\, f(\alpha_1) + b\, f(\alpha_2)$$

So f is a linear functional on V.

(**ii**) If V is finite dimensional , then there exists a unique vector β' such that

$$f(\alpha) = <\alpha, \beta'> \quad \forall\ \alpha \in V$$

We have $f(\alpha) = <\overline{\beta, T\alpha}> = <T\alpha, \beta> = <\alpha, T^*\beta> \ \forall\ \alpha$

$\therefore \quad$ if $f(\alpha) = <\alpha, \beta'> \ \forall\ \alpha$, then

$<\alpha, T^*\beta> = <\alpha, \beta'> \ \forall\ \alpha$

Hence $\beta = T^*\beta$

Example 6: Let V be a finite-dimensional inner product space and T be a linear operator on V. If T is invertible, then prove that T^* is invertible and $(T^*)^{-1} = (T^{-1})^*$.

Solution: Suppose T is invertible. Then

$T\,T^{-1} = I$

$\Rightarrow \quad (T\,T^{-1})^* = I^*$

$\Rightarrow \quad (T^{-1})^*\, T^* = I \qquad$ as $I^* = I$

$\Rightarrow$ T^* is also invertible and $(T^*)^{-1} = (T^{-1})^*$.

Example 7: Let T be a linear operator on a finite-dimensional inner product space V. Then T is self-adjoint iff its matrix in every orthonormal basis is a self-adjoint matrix.

Solution: Let B be any orthonormal basis for T. Then

$$[T^*]_B = [T]_B^* \qquad \ldots..(1)$$

IF PART: Let T be self-adjoint i.e. $T = T^*$. Then from (1), $[T]_B = [T]_B^*$ i.e. $[T]_B$ is a self-adjoint matrix.

ONLY IF PART: Let $[T]_B$ be a self-adjoint matrix. Then $[T]_B = [T]_B^*$

$= [T^*]_B$; using eq^n (1)

$\therefore \quad T = T^*$

Example 8: If T is a self-adjoint linear operator on a finite dimensional inner product Space V, then det(T) is real.

Solution: Let B be any orthonormal basis for V. Then

$$[T^*]_B = [T]_B^*$$

But $\quad T^* = T \;\Rightarrow\; [T]_B = [T]_B^* \qquad \ldots..(1)$

Let $\quad [T]_B = A \;\Rightarrow\; A = A^*$

det A = det $(A^*) = \overline{\det(A)}$ = det (A) is real.

Example 9: If T is self-adjoint, then S^*TS is self-adjoint $\forall$ S. Conversely if S is invertible and S^*TS is self-adjoint, then T is self-adjoint. Prove both results.

Solution: Given that T is self-adjoint, so $T^* = T$. Now $(S^*TS)^* = S^*T^*(S^*)^* = S^*TS$

So S^*TS is self-adjoint. Now, conversely, let S be invertible, then S^* in also invertible. If S^*TS is self-adjoint , then

$$(S^*TS)^* = S^*TS$$

$$\Rightarrow \quad S^*T^*S = S^*TS$$

So $\quad (S^*)^{-1}(S^*T^*S)S^{-1} = (S^*)^{-1}(S^*TS)S^{-1}$

$$\Rightarrow \quad ((S^*)^{-1}S^*)\ T^*(SS^{-1}) = ((S^*)^{-1}S^*)\ T(SS^{-1})$$

$$\Rightarrow \quad I\,T^*\,I = I\,T\,I$$

$$\Rightarrow \quad T^* = T$$

or T is self-adjoint.

Example 10: Let V be a finite-dimensional inner product space, and T be any linear operator on V. Suppose W is a subspace of V which is invariant under T. Then prove that the orthogonal complement of W is invariant under T^*.

Solution: Given that W is invariant under T.

Claim: $W^\perp$ is invariant under T^*.Let $\beta \in W^\perp$ be arbitrary. Then we shall prove that T^*B is in $W^\perp$ i.e. T^*B is orthogonal to every vector in W. Let $\alpha \in W$. Then

$$< \alpha , T^*\beta > \; = \; < T\alpha , \beta >$$

$= 0$, since $\alpha \in W \Rightarrow T\alpha \in W$ and β is orthogonal to every vector in W.

So $T^*\beta$ is orthogonal to every vector $\alpha \in W$

So $T^*\beta$ is in $W^\perp$.

$\Rightarrow$ $W^\perp$ is invariant under T^*.

13.6 *POSITIVE OPERATOR*

Positive operator: A linear operator T on an inner product space V is called positive (in symbols, T > 0), if –

(i) T is self adjoint i.e. $T^* = T$, and
(ii) $< T\alpha , \alpha > \; > 0 \; \forall \; \alpha \neq 0$

If $\alpha = 0$, then $< T\alpha , \alpha > = 0$. Hence if T is positive, then $< T\alpha , \alpha > \; \geq \; 0 \;\; \forall \; \alpha \in V$ and $< T\alpha , \alpha > = 0 \Rightarrow \alpha = 0$.

13.7 *NON-NEGATIVE OPERATOR*

Non-Negative operator: A linear operator T on an inner product space V is called non-negative, if –

(i) It is self-adjoint , and
(ii) $< T\alpha , \alpha > \; \geq \; 0 \;\; \forall \; \alpha \in V$

Note: (1) Every positive operator is also a non-negative operator.

(2) If T is a non-negative operator, then $< T\alpha , \alpha > = 0$, is possible even if $\alpha \neq 0$. So a non-negative operator **may not** be a positive operator

(3) If S and T are two linear operators on an inner product space V, then we define

$$S > T \text{ if } S - T > 0$$

(4) Some authors say a positive operator as '**positive definite'**.

Theorem 9: Let V be an inner product space and T be a linear operator on V. Let 'p' be the function defined on ordered pairs of $\alpha, \beta \in V$ by

$$p(\alpha, \beta) = < T\alpha , \beta >$$

Then the function p is an inner product on V iff T is a positive operator.

Proof: Step I: Let a, b $\in$ F and $\alpha_1 , \alpha_2 \in V$. Then

$$p\,(a\alpha_1 + b\,\alpha_2, \beta) \;=\; < T(a\alpha_1 \;+\; b\,\alpha_2) , \beta > = < Ta\alpha_1 \;+\; Tb\alpha_2) , \beta >$$

$$= \; a < T\alpha_1 , \beta > + \, b < T\alpha_2 , \beta >$$

$$= a\, p(\alpha_1 , \beta) + b\, p(\alpha_2 , \beta)$$

So the function p satisfies linearity property.

Step II: Now the function p will be an inner product on V if and only if

$$p(\alpha, \beta) = \overline{p(\beta, \alpha)} \text{ and } p(\alpha, \alpha) > 0, \alpha \neq 0$$

So we have $p(\alpha, \beta) = \langle T\alpha, \beta \rangle$

$$\Rightarrow \overline{p(\beta, \alpha)} = \langle \overline{T\beta, \alpha} \rangle = \langle \alpha, T\beta \rangle$$

Also $p(\alpha, \alpha) = \langle T\alpha, \alpha \rangle$.

Hence the function p will be an inner product on iff

(i) $\langle T\alpha, \beta \rangle = \langle \alpha, T\beta \rangle \ \forall \alpha, \beta \in V$ i.e. T is self-adjoint.

(ii) $\langle T\alpha, \alpha \rangle > 0$ if $\alpha \neq 0$

Hence the function p will be an inner product on V iff the linear operator T is positive.

Note: Now we shall show that if V is finite-dimensional, then every inner product on V is of the type as discussed in next theorem –

Theorem 10: Let V(F) be a finite-dimensional inner product space with inner product < , > . If p is any inner product on V, there is a unique positive linear operator T on V such that $p(\alpha, \beta) = \langle T\alpha, \beta \rangle \ \forall \alpha, \beta \in V$.

Proof: Let $\beta \in V$ be a fixed vector and $f : V \rightarrow F$ such that

$$f(\alpha) = p(\alpha, \beta) \ \forall \alpha \in V$$

As we have seen, p satisfies linearity property, so f is a linear functional on V. Hence by Riesz representation theorem, there exists a unique vector β' in such that

$$f(\alpha) = \langle \alpha, \beta' \rangle \ \forall \alpha \text{ in } V$$

$$\Rightarrow p(\alpha, \beta) = \langle \alpha, \beta' \rangle \ \forall \alpha \text{ in } V$$

Let us define $T : V \rightarrow V$ such that $T\beta = \beta'$.

So $p(\alpha, \beta) = \langle \alpha, T\beta \rangle \ \forall \alpha, \beta \in V$ .....(1)

We also have, $p(\alpha, \beta) = \langle \alpha, T\beta \rangle$

$$p(\alpha, \beta) = \overline{p(\beta, \alpha)} \text{ , by conjugacy property of inner product p}$$

$$= < \overline{\beta , T\alpha} > = < T\alpha , \beta >$$

Thus, we have, $p(\alpha, \beta) = < T\alpha , \beta > \forall \alpha, \beta \in V$ …..(2)

Linearity of T: Let α_1 , $\alpha_2 \in V$ and a_1 , $a_2 \in F$. Then for all $r \in V$, we have

$$< T(a_1\alpha_1 + a_2\alpha_2) , r > = p(a_1\alpha_1 + a_2\alpha_2 , r)$$

$$= a_1\, p(\alpha_1 , r) + a_2\, p(\alpha_2 , r) \text{ , by linearity of p}$$

$$= < (a_1T\alpha_1 + a_2T\alpha_2) , r >\text{, by linearity of inner product} < , >$$

So, we have, $T(a_1\alpha_1 + a_2\alpha_2) = a_1T\alpha_1 + a_2T\alpha_2$

Hence T is a linear operator. Thus, we have proved the existence of a linear operator T with $p(\alpha, \beta) = < T\alpha , \beta >$. Since p is an inner product, so by previous theorem, T is positive.

Uniqueness: Suppose there are two linear operators T and U such that

$$p(\alpha, \beta) = < T\alpha , \beta > = < U\alpha , \beta > \forall \alpha, \beta \in V$$

Then $< T\alpha - U\alpha , \beta > = 0 \ \forall \alpha, \beta \in V$ …..(3)

Let us keep α fixed. Then from equation (3), we see that the vector $T\alpha - U\alpha$ is orthogonal to every vector β in V.

Therefore $T\alpha - U\alpha = 0, \forall \alpha \in V$

$$\Rightarrow \quad T\alpha = U\alpha , \forall \alpha \in V$$

Hence T is unique.

Theorem 11: Let V be a finite-dimensional inner product space and T a linear operator on V. Then t is positive if and only if there is an invertible linear operator U on V such that $T = U^* U$.

Proof: Let $T = U^* U$, where U in an invertible linear operator on V.

Since $T^* = (U^*U)^* = U^*(U^*)^* = U^*U = T$

So T is self-adjoint. Also,

$< T\alpha , \alpha > \; = \; < U^*U\alpha , \alpha > \; = \; < U\alpha , U^{**} \alpha > \; = \; < U\alpha , U \alpha > \geq 0$

Also $< T\alpha , \alpha > \; = 0 \Rightarrow \; < U\alpha , U \alpha > = 0 \;\; \Rightarrow U \alpha = 0$

$\Rightarrow \quad \alpha = 0$, as U in invertible and V is finite-dimensional, so U is non-singular.

So if $\alpha \neq 0$, then $< T\alpha , \alpha > \; > 0$

Hence T is positive.

Conversely, suppose T is positive.Then $p(\alpha, \beta) = < T \alpha , \beta >$ is an inner product on V. Suppose $\{ \alpha_1 , \ldots, \alpha_n \}$ be a basis for V which is orthonormal with respect to the inner product < , > and let $\{ \beta_1 , \ldots, \beta_n\}$ be a basis orthonormal with respect to the inner product p. So,

$$p(\beta_i , \beta_j) = \delta_{ij} = < \alpha_i , \alpha_j >$$

Now, let U be the unique linear operator on V such that $U \beta_i = \alpha_i$; i = 1, 2, …., n. Obviously U is invertible, because it carries a basis onto a basis. We have

$$p(\beta_i , \beta_j) = < \alpha_i , \alpha_j > = < U \beta_i , U \beta_j >$$

Now let $\alpha, \beta \in V$; such that

$$\alpha = \sum_{i=1}^{n} x_i\beta_i \text{ and } \beta = \sum_{j=1}^{n} y_j\beta_j \text{ . Then}$$

$$< T \alpha , \beta > \; = \; p(\alpha, \beta)$$

$$< T \alpha , \beta > \; = p \left(\sum_{i=1}^{n} x_i\beta_i , \sum_{j=1}^{n} y_j\beta_j \right) = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i\overline{y_j} \; p(\beta_i\beta_j)$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} x_i\overline{y_j} < U\beta_i, U\beta_j > = < \sum_{i=1}^{n} x_iU \beta_i , \sum_{j=1}^{n} y_jU \beta_j >$$

$$= < U \sum_{i=1}^{n} x_i\beta_i , U \sum_{j=1}^{n} y_j\beta_j > = < U \alpha, U \beta > \; = < U^* U \alpha, \beta >$$

Thus $\forall \alpha, \beta \in V$, we have

$$< T \alpha , \beta > \; = < U^* U \alpha, \beta >$$

$\Rightarrow \quad T = U^* U$

Positive Matrix: Let A = $[a_{ij}]_{n \times n}$ be a square matrix of order n over the field of R or C, then A is said to be positive if :

(i) $A^* = A$, and

(ii) $\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}\, \overline{x_i}\, x_j > 0$, where $x_1, \ldots, x_n \in F$ and not all zero

Principal Minors of a Matrix: Let A = $[a_{ij}]_{n \times n}$ be a an arbitrary field **F**. The principal minors of A are the n scalars defined as –

der $A^{(K)} = \det \begin{pmatrix} a_{11} & \cdots & a_{1K} \\ \vdots & \ddots & \vdots \\ a_{K1} & \cdots & a_{KK} \end{pmatrix}$, where K = 1, 2, …, n.

Suppose A = $[a_{ij}]_{n \times n}$ over **R** or **C**. Then A is positive if the principal minors of A are all positive. (Its converse is also true).

Example 1: Which of the following matrices are positive –

(i) $\begin{bmatrix} 1 & 1+i \\ 1-i & 3 \end{bmatrix}$ (ii) $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ (iii) $\begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \end{bmatrix}$

Solution: (i) Here obviously $A^* = A$. So A is self-adjoint. Now principal minors of A are 1 and

$\begin{vmatrix} 1 & 1+i \\ 1-i & 3 \end{vmatrix}$ i.e. 1 and 1.

So both the principal minors of A are +ve . Hence A is a +ve matrix.

(ii) It is not self-adjoint. Hence it is not positive.

(iii) Here $A^* = A$. Also all the principal minors viz 1,

$\begin{bmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{3} \end{bmatrix}$ and $\begin{vmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \end{vmatrix}$ are positive (**verify**). Hence A is positive.

Example 2: Prove that every entry on the main diagonal of a positive matrix is positive.

Solution: Let A = $[a_{ij}]_{n \times n}$ be a positive matrix. So

$$\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}\, \overline{x_i}\, x_j > 0, \qquad \ldots.(1)$$

where $x_1, \ldots., x_n$ are any n scalars (not all zero). Now suppose that out of n scalars $x_1, \ldots., x_n$, we take $x_i = 1$ and each of the remaining (n – 1) scalars is taken as zero. Then from equation (1), we conclude that $a_{ii} > 0\ \forall$ i. Hence each entry on the main diagonal of a +ve matrix is positive.

13.8 *UNITARY OPERATOR*

Definition: In a inner product space *V*, let *T* be a linear operator. Then the operator *T* is called unitary operator if adjoint T^* of T exist and $TT^* = T^*T = I$

Note 1: In a finite dimensional inner product space T is unitary *iff* $T^*T = I$

2: A linear operator T on a finite dimensional inner product space V is unitary iff T preserve inner product.

13.9 *NORMAL OPERATOR*

In this section we will learn about the important topic in inner product space.

Definition: Let in a inner product space *V*, *T* be a linear operator. Then the operator *T* is called normal operator or normal if it commutes with its adjoint *i.e.*, $TT^* = T^*T$.

Note 1: If vector space is of finite dimensional then T^* will definitely exist.

2: If vector space is not of finite dimensional then definition will make sense only if T possesses adjoint.

Theorem 12: Every self-adjoint operator is normal.

Proof: Let we consider *T* be a self-adjoint operator then obviously, $T^* = T$.

Therefore, we can say that $TT^* = T^*T$,

Hence *T* is normal

Theorem 13: Every unitary operator is normal.

Proof: Let we consider *T* be a unitary operator then obviously, $TT^* = T^*T = I$

Therefore, we can say that $TT^* = T^*T$,

Hence *T* is normal.

Theorem 14: Let in a inner product space *V*, *T* be a normal operator. Then a necessary and sufficient condition that α be a characteristic vector of *T* is that it be a characteristic vector of T^*.

Proof: Let us consider T be a normal operator on an inner product space V. If $\alpha \in V$, then we have,

$$\|T(\alpha)\|^2 = (T\alpha, T\alpha) = (\alpha, T^*T\alpha) = (\alpha, TT^*\alpha)$$

$= (T^*\alpha, T^*\alpha) = \|T^*(\alpha)\|$

Since T is normal and if $\alpha \in V$,

$\|T\alpha\| = \|T^*\alpha\|$ ……….. (1)

If c be scalar, then (1) can be written as

$(T - cI)^* = T^* - \bar{c}I^* = T^* - \bar{c}I$

Now we have to show, $T - cI$ is normal.

We have, $(T - cI)(T - cI)^* = (T - cI)(T^* - \bar{c}I)$

$= TT^* - \bar{c}T - cT^* + c\bar{c}I$

Also $(T - cI)^*(T - cI) = (T^* - \bar{c}I)(T - cI)$

$= T^*T - \bar{c}T - cT^* + c\bar{c}I$

As we know T is normal. So,

$(T - cI)(T - cI)^* = (T - cI)^*(T - cI)$

Thus, $(T - cI)$ is normal. Now from (1),

$\|(T - cI)(\alpha)\| = \|(T - cI)^*(\alpha)\| \;\forall \alpha \in V$

$\Rightarrow \|(T - cI)\alpha\| = \|(T^* - \bar{c}I)^*(\alpha)\| \;\forall \alpha \in V$ ………. (2)

By equation (2) we can say that,

$\Rightarrow (T - cI)\alpha = 0 \text{ iff } (T^* - \bar{c}I)\alpha = 0$

i.e., $T(\alpha) = c\alpha \text{ iff } T^*\alpha = \bar{c}\alpha$

Thus, we can say that α is a eigen vector of T corresponding to the eigen value c if and only if it is a characteristic vector of T^* corresponding to the eigen value $\bar{c}$.

Remark 1: The characteristic vector for *T* belonging to distinct characteristic values is orthogonal if *T* is a normal operator on an inner product space *V*.

2: In a normal operator's characteristic spaces are pairwise orthogonal to each other.

Definition (Normal matrix): A square order complex matrix A is called normal if,

$$AA^* = A^*A.$$

If matrix is diagonal matrix D, then obviously

$$DD^* = D^*D$$

Remark 1: A unitarily equivalent to a diagonal matrix *iff* matrix is normal.

Solved example

Example 1: If in a inner product space *V*, *T* be a normal operator. Then *cT* is also a normal operator for any scalar *c*.

Proof: We have given that *T* be a normal operator i.e., $TT^* = T^*T$

Since, $(cT)^* = \bar{c}T^*$

Now, $(cT)(cT)^* = (cT)(\bar{c}T^*) = c\bar{c}(TT^*)$

Again, $(cT)^*(cT) = (\bar{c}T^*)(cT) = (\bar{c}c)(T^*T)$

Thus we can say, $(cT)(cT)^* = (cT)^*(cT)$

Hence, cT is normal.

Example 2: In a inner product space *V*, if T_1, T_2 are normal operator with the property that either commutes with the adjoint of other, then prove that $T_1 + T_2$ and T_1T_2 are also normal operator.

Solution: We have given T_1, T_2 are normal. Therefore,

$T_1T_1^* = T_1^*T_1$ and $T_2T_2^* = T_2^*T_2$

According to question it is given that,

$T_1T_2^* = T_2^*T_1$ and $T_2T_1^* = T_1^*T_2$

Now, $(T_1 + T_2)(T_1 + T_2)^* = (T_1 + T_2)(T_1^* + T_2^*)$

$$= T_1T_1^* + T_1T_2^* + T_2T_1^* + T_2T_2^*$$

$$= T_1T_1^* + T_2^*T_1 + T_1^*T_2 + T_2T_2^*$$

$$= T_1^*(T_1 + T_2) + T_2^*(T_1 + T_2) = (T_1^* + T_2^*)(T_1 + T_2)$$

$$= (T_1 + T_2)^*(T_1 + T_2)$$

Thus, $T_1 + T_2$ is normal.

Now, $(T_1T_2)(T_1T_2)^* = T_1T_2T_2^*T_1^* = T_1(T_2T_2^*)T_1^*$

$$= T_1(T_2^*T_2)T_1^*$$

$$= (T_1T_2^*)(T_2T_1^*)$$

$$= (T_2^*T_1)(T_1^*T_2)$$

$$= T_2^*(T_1T_1^*)T_2$$

$$= T_2^*(T_1^*T_1)T_2$$

$$= (T_2^*T_1^*)(T_1T_2) = (T_1T_2)^*(T_1T_2)$$

Thus, T_1T_2 is normal.

Example 3: In a finite dimensional complex inner product space let T be the linear operator. Show that T is normal if and only if its real and imaginary parts commute.

Solution: Let $T = T_1 + iT_2$. Then $T_1^* = T_1$ and $T_2^* = T_2$. Let we assume that $T_1T_2 = T_2T_1$ then we have to prove that T is normal.

We have, $T^* = (T_1 + iT_2)^* = T_1^* + \bar{i}\,T_2^* = T_1 - iT_2$

$\therefore TT^* = (T_1 + iT_2)(T_1 - iT_2) = T_1^2 - iT_1T_2 + iT_2T_1 + T_2^2 = T_1^2 + T_2^2$ $\quad [\because T_1T_2 = T_2T_1]$

Also, $T^*T = (T_1 - iT_2)(T_1 + iT_2) = T_1^2 + iT_1T_2 - iT_2T_1 + T_2^2 = T_1^2 + T_2^2$

$\therefore TT^* = T^*T$. Hence T is normal.

Conversely, we assume that T is normal then we have to prove that $TT^* = T^*T$.

$$\Rightarrow T_1^2 - iT_1T_2 + iT_2T_1 + T_2^2 = T_1^2 + iT_1T_2 - iT_2T_1 + T_2^2$$

$$\Rightarrow 2i(T_1T_2 - T_2T_1) = 0$$

$$\Rightarrow T_1T_2 - T_2T_1 = 0 \qquad [\because 2i \neq 0]$$

$\Rightarrow T_1T_2 = T_2T_1$

Check your progress

Problem 1: In a finite dimensional complex inner product space let T be the linear operator. Show that T is normal if and only if its real and imaginary parts commute.

Solution: Let $T = T_1 + iT_2$. Then $T_1^* = T_1$ and $T_2^* = T_2$. Let we assume that $T_1T_2 = T_2T_1$ then we have to prove that T is normal.

We have, $T^* = (T_1 + iT_2)^* = T_1^* + \bar{i}\,T_2^* = T_1 - iT_2$

$\therefore TT^* = (T_1 + iT_2)(T_1 - iT_2) = T_1^2 - iT_1T_2 + iT_2T_1 + T_2^2 = T_1^2 + T_2^2 \qquad [\because T_1T_2 = T_2T_1]$

Also, $T^*T = (T_1 - iT_2)(T_1 + iT_2) = T_1^2 + iT_1T_2 - iT_2T_1 + T_2^2 = T_1^2 + T_2^2$

$\therefore TT^* = T^*T$. Hence T is normal.

Conversely, we assume that T is normal then we have to prove that $TT^* = T^*T$.

$\Rightarrow T_1^2 - iT_1T_2 + iT_2T_1 + T_2^2 = T_1^2 + iT_1T_2 - iT_2T_1 + T_2^2$

$\Rightarrow 2i(T_1T_2 - T_2T_1) = 0$

$\Rightarrow T_1T_2 - T_2T_1 = 0 \qquad [\because 2i \neq 0]$

$\Rightarrow T_1T_2 = T_2T_1$

Problem 2: Let S and T be two positive linear operators on an inner product space V. Then prove that S + T is also positive operator.

Solution: Given $S^* = S$ and $T^* = T$

So $(S + T)^* = S^* + T^* = S + T$

So S+ T is self adjoint.

Also, if $\alpha \in V$, then

$$\langle (S+T)\alpha, \alpha \rangle = \langle S\alpha + T\alpha, \alpha \rangle = \langle S\alpha, \alpha \rangle + \langle T\alpha, \alpha \rangle$$

But S and T are positive. So $< S\alpha , \alpha > \geq 0$ and $< T\alpha , \alpha > \geq 0$.

$\Rightarrow \quad < (S+T)\alpha, \alpha > \geq 0.$

Hence S + T is positive.

Problem 3: Let V be a finite-dimensional inner product space and T be a self- adjoint linear operator on V. Prove that the range of T is the orthogonal complement of the null space of T i.e. $R(T) = [N(T)]^{\perp}$.

Solution: Let $\alpha \in R(T)$. Then $\exists$ a vector $\beta \in V$ such that $\alpha = T\beta$. Let r be an arbitrary vector of $[N(T)]^{\perp}$. Then Tr = 0

We have

$$< \alpha , r > = < T\beta , r > = < \beta , T^{*}r > = < \beta , Tr > \text{ as } T^{*} = T$$

$$= < \beta , 0 > = 0$$

Thus $< \alpha , r > = 0 \quad \forall \; r \in N(T)$

So, $\alpha \in [N(T)]^{\perp} \Rightarrow R(T) \subseteq [N(T)]^{\perp}$ …..(1)

Since $V = N(T) \oplus [N(T)]^{\perp}$

$\Rightarrow \quad \dim V = \dim N(T) + \dim [N(T)]^{\perp}$ …..(2)

By **Rank- nullity theorem**, we have

$$\dim V = \dim R(T) + \dim N(T) \qquad \text{…..(3)}$$

So we conclude that $\dim R(T) = \dim [N(T)]^{\perp}$ …..(4)

From equation (1) and (4) , we conclude that

$$R(T) = [N(T)]^{\perp}$$

13.10 *SUMMARY*

In this unit we have learned about the most essential tool name as operators used in inner product space like adjoint operator, self-adjoint operator, skew-symmetric operator, positive operator, unitary operator and normal operator. Mostly, the uses of these operators to solve out the matrix problems. Other important concepts introduced in this unit were:

- Every self-adjoint operator is normal.
- Every unitary operator is normal
- The operation of adjoint behaves like the operation of conjugation on complex numbers
- Every positive operator is also a non-negative operator

13.11 *GLOSSARY*

- Unitary operator
- Normal operator
- Adjoint operator
- Self-adjoint operator
- Skew-symmetric or Hermitian operator.

13.12 *REFERENCES*

- S Kumaresan; Linear Algebra-A Geometric Approach; PHI-2016.
- K Hoffman & Ray Kunze; Linear Algebra; PHI-1971.
- P Halmos; Finite-Dimensional Vector Spaces;
- G Strang; Introduction To Linear Algebra;

13.13 *SUGGESTED READING*

- NPTEL videos.
- Schaum series.
- A R Vashishtha, Krishna Prakashan; Meerut.
- Graduate Text In Mathematics, Springer.

13.14 *TERMINAL QUESTION*

Long answer type question

1: Let S and T be linear operators on an inner product space V and c $\in$ **F**. If S and T possess adjoints, then prove that the operators S + T, cT, ST, T^* will possess adjoints.

2: Prove that Every linear operator T on a finite dimensional complex inner product space V can be uniquely expressed as

$T = T_1 + iT_2$, where T_1 & T_2 are self-adjoint linear operators on V.

3: Prove that every linear operator T on a finite-dimensional inner product space V can be uniquely expressed as $T = T_1 + T_2$, where T_1 is self-adjoint and T_2 is skew.

4: Prove that the necessary and sufficient condition that a linear transformation T on a unitary space (of any dimension) be self-adjoint (Hermitian) is that,

$$< T\alpha , \alpha > \text{ be real } \forall \ \alpha \in V$$

Short answer type question

1: Let V be a finite-dimensional inner product space and let B = { α_1 , …., α_n } be an ordered orthonormal basis for V . Let T be a linear operator on V and let $A = [a_{ij}]_{m \times n}$ be the matrix of T with respect to the ordered basis B. Then prove that $a_{ij} = < T\alpha_j, \alpha_i >$.

2: In any orthonormal basis for *V* and *T* be the linear operator on *V*, then prove that the matrix of T^* is the conjugate transpose of the matrix of *T*.

3: Prove that the necessary and sufficient condition that a linear transformation T on an inner product space V be $\hat{0}$ is that $< T\alpha, \beta > = 0, \forall \ \alpha, \beta \in V$

4: Prove that the necessary and sufficient condition that a linear transformation T on a unitary space be $\hat{0}$ is that $< T\alpha, \alpha > = 0 \ \ \forall \ \alpha \in V$

5: A linear operator on R^2 is defined by

$$T(x_1, y_1) = (x_1 + 2y_1, x_1 - y_1)$$

Find the adjoint T^*, if the inner product is standard one.

6: Prove that the product of two self-adjoint operators on an inner product space is self-adjoint iff the two operators commute.

7: If T is self-adjoint, then S^*TS is self-adjoint $\forall$ S. Conversely if S is invertible and S^*TS is self-adjoint, then T is self-adjoint. Prove both results.

8: Prove that characteristic of normal operator are pair-wise orthogonal.

9: Prove that each self-adjoint and unitary operaor are normal operator

10: If in a inner product space *V*, *T* be a normal operator. Then prove that *cT* is also a normal operator for any scalar *c*.

11: If in a finite dimensional vector space V, T be a linear operator. If $\|T\alpha\| = \|T^*\alpha\| \ \forall \alpha \in V$

Fill in the blanks

1: $(S + T)^* =$ …………..

2: A linear operator T on an inner product space V is said to be self-adjoint if ………..

3: A linear T is called *skew-symmetric* or *skew-Hermitian* according as the vector space V is ……………………

4: A necessary and sufficient condition that a linear transformation T on a unitary space be $\hat{0}$ is that ………………..

13.15 *ANSWERS*

Answer of short question

5: $[T^*]_B = \begin{bmatrix} 1 & 1 \\ 2 & -1 \end{bmatrix}$

Answer of fill in the blanks

1: $S^* + T^*$ **2:** $T^* = T$ **3:** Real or Complex

4: $< T\alpha, \alpha > = 0 \quad \forall \alpha \in V$

UNIT-14: BILINEAR FORM

CONTENTS

14.1 *INTRODUCTION*

We shall talk about bilinear and quadratic forms in this unit. We can expand our understanding of linear phenomena by utilizing bilinear forms, which are essentially linear transformations that are linear in many variables. Quadratic forms, which are (classically) homogeneous quadratic polynomials in multiple variables, are closely connected to them. It may surprise you to learn that we can still study quadratic forms using many of the same resources from linear algebra, even if they are not linear. The fundamental characteristics of bilinear and quadratic forms will be covered, with an emphasis on the concepts of positive definiteness and positive semidefiniteness as well as some of their uses in geometry, calculus, and linear algebra. Singular value decomposition, which connects many of the topics we have covered, is the topic of our final discussion.

14.2 *OBJECTIVES*

After reading this unit learners will be able to

- Visualized the concept of bilinear form and matrix of bilinear form.
- Understand the concept of degenerate, non-degenerate and symmetric bilinear form.
- Visualized the concept of quadratic form
- Application and implementation of normal operator, bilinear and normal form.

14.3 *BILINEAR FORM*

Suppose *U* and *V* are two vector spaces corresponding to the same field *F*. Let
$W = U \times V$ *i.e.*, $W = \{(\alpha, \beta) : \alpha \in U, \beta \in V\}$

If (α_1, β_1), $(\alpha_2, \beta_2) \in W$, then there equality can be defined as follows:

$(\alpha_1, \beta_1) = (\alpha_2, \beta_2)$ if $\alpha_1 = \alpha_2$ and $\beta_1 = \beta_2$

And the addition is defined as follows:

$$(\alpha_1, \beta_1) + (\alpha_2, \beta_2) = (\alpha_1 + \alpha_2,\ \beta_1 + \beta_2)$$

If *c* is any element of field and (α, β) be the element of *W*, the scalar multiplication can be defined as follows:

$$c(\alpha, \beta) = (c\alpha, c\beta)$$

It is obvious that *W* is a vector space over the field *F* with respect to addition and scalar multiplication as specified above. The external direct product of vector spaces *U* and *V* is denoted by the symbol *W*, which we will express as

$$W = U \oplus V$$

We will now discuss bilinear forms, which are a particular class of scalar-valued functions on *W*.

Definition: Suppose *U* and *V* are two vector spaces corresponding to the same field *F*. A bilinear form $W = U \oplus V$ is a function *f* from *W* into *F*, which assign to each each element (α, β) in such a way that

$$f(a\alpha_1 + b\alpha_2, \beta) = af(\alpha_1, \beta) + bf(\alpha_2, \beta)$$

and $f(\alpha, a\beta_1 + b\beta_2) = af(\alpha, \beta_1) + bf(\alpha, \beta_2)$

Here $f(\alpha, \beta)$ is an element of F. It denotes the image of (α, β) under F. Consequently, a function from *W* into *F* that, while one of its inputs is fixed, is linear as a function of the other is called a bilinear form on *W*.

If $U = V$, then to say that *f* is a bilinear form on $W = V \oplus V$ we just refer f is bilinear form on *V*.

Another definition

Let *V* denote a vector space over a field *F* of characteristic other than 2.

Definition: A bilinear form *f* is a map $f : V \times V \to F$, such that *f* is bilinear if the following properties are satisfied

(i) $f(ax + bx', y) = af(x, y) + bf(x', y)$ for every $x, x', y \in V$ and $a, b \in F$.

(ii) $f(x, cy + dy'') = cf(x, y) + df(x', y')$ for every $x, y, y' \in V$ and $c, d \in F$.

In other words, *f* is bilinear if it is separate linear in each variable.

Definition: The bilinear form *f* is said to be symmetric if $f(x, y) = f(y, x)$. It is called skew-symmetric if $f(x, y) = -f(y, x)$.

Remark: Note that the characteristic of the field is 0 and hence $-1 \neq 1$.

Example 4: Let *V* denote a vector space over a field *F*. Consider L_1, L_2 be linear function on V. If *f* be a function from $V \times V$ into F and defined as

$f(\alpha, \beta) = L_1(\alpha)L_2(\beta)$ then, *f* is bilinear form on *V*.

Solution: If $(\alpha, \beta) \in V$, then $L_1(\alpha), L_2(\beta)$ are scalars.

So, we have $f(a\alpha_1 + b\alpha_2, \beta) = L_1(a\alpha_1 + b\alpha_2)L_2(\beta)$

$$= [aL_1(\alpha_1) + bL_1(\alpha_2)]L_2(\beta)$$

$$= aL_1(\alpha_1)L_2(\beta) + bL_1(\alpha_2)L_2(\beta)$$

$$= af(\alpha_1, \beta) + bf(\alpha_2, \beta)$$

Also, $f(\alpha, a\beta_1 + b\beta_2) = L_1(\alpha)L_2(a\beta_1 + b\beta_2)$

$$= L_1(\alpha)[aL_2(\beta_1) + bL_2(\beta_2)]$$

$$= aL_1(\alpha)L_2(\beta_1) + bL_1(\alpha)L_2(\beta_2)$$

$$= af(\alpha, \beta_1) + bf(\alpha, \beta_2)$$

Thus, *f* is bilinear on *V*.

Example 5: Let *V* denote a vector space over a field *F*. Suppose linear operator (*T*) on *V* and *f* is a bilinear form on *V*. Let *g* is a function from $V \times V$ into F and defined as

$g(\alpha, \beta) = f(T\alpha, T\beta)$. Then show that *g* is also a bilinear form on *V*.

Proof: Since we have $g(a\alpha_1 + b\alpha_2, \beta) = f(T(a\alpha_1 + b\alpha_2), T\beta)$

$$= f(aT(\alpha_1 + bT\alpha_2, T\beta)$$

$$= af(T\alpha_1, T\beta) + bf(T\alpha_2, T\beta)$$

$$= ag(\alpha_1, \beta) + bg(\alpha_2, \beta)$$

Also, $g(\alpha, a\beta_1 + b\beta_2) = f(T\alpha, T(a\beta_1 + b\beta_2))$

$$= f(T\alpha, aT\beta_1 + bT\beta_2)$$

$$= af(T\alpha, T\beta_1) + bf(T\alpha, T\beta_2)$$

$$= ag(\alpha, \beta_1) + bg(\alpha, \beta_2)$$

Hence *g* is a bilinear form on *V*.

Example 6: Consider the two vector space *U* and *V* over the same field *F*. Let $W = U \oplus V$. If $\hat{0}$ is the zero function *W* into *F*. Then show that $\hat{0}$ is a bilinear form on W.

Solution: We have $\hat{0}(\alpha, \beta) = 0 \forall (\alpha, \beta) \in W$.

Now, $\hat{0}(a\alpha_1 + b\alpha_2, \beta) = 0 = 0 + 0 = a0 + b0 = a\hat{0}(\alpha_1, \beta) + b\hat{0}(\alpha_2, \beta)$

Also, $\hat{0}(\alpha, a\beta_1 + b\beta_2) = 0 = 0 + 0 = a0 + b0 = a\hat{0}(\alpha, \beta_1) + b\hat{0}(\alpha, \beta_2)$

Thus $\hat{0}$ is a bilinear form on *W*.

Remarks 1: Let $V = V_n(F)$ *i.e.*, let *V* be the vector space of *n*- tuple over the field *F*. If $\alpha = (a_1, a_2, ..., a_n)$ and $\beta = (b_1, b_2, ..., b_n)$ be any two elements in *V*, let *f* be a function from $V \times V$ into *F* defined as

$f(\alpha, \beta) = a_1 b_1 + a_2 b_2 + ... + a_n b_n$. So, *f* is a bilinear form on *V*.

2: If *U* and *V* are the two vector space over the same field *F* and *f* is a bilinear form on $U \times V$ into *F* defined as,

$(-f)(\alpha, \beta) = -f(\alpha, \beta)$, is a bilinear form on $U \times V$

Example 7: Which of the following vectors $x = (x_1, x_2), y = (y_1, y_2)$ defined on R^2 are the bilinear form

(i) $f(x, y) = x_1 y_2 - x_2 y_1$

(ii) $f(x, y) = (x_1 - y_1)^2 + x_2 y_2$

Solution: Let $x = (x_1, x_2)$,

$y = (y_1, y_2)$,

And $z = (z_1, z_2)$

Be any three vector in R^2. Let $a, b \in R$. Then

$$ax + by = a(x_1, x_2) + b(y_1, y_2)$$

$$= (ax_1 + by_1, ax_2 + by_2)$$

(i) Now, by definition of *f*, we have

$f(x, z) = f((x_1, x_2), (z_1, z_2)) = x_1 z_2 - x_2 z_1$,

$f(y, z) = f((y_1, y_2), (z_1, z_2)) = y_1 z_2 - y_2 z_1$

$f(z, x) = f((z_1, z_2), (x_1, x_2)) = z_1 x_2 - z_2 x_1$ and

$f(z, y) = f((z_1, z_2), (y_1, y_2)) = z_1 y_2 - z_2 y_1$

Now,

$$f(ax + by, z) = f((ax_1 + by_1, ax_2 + by_2), (z_1, z_2))$$

$$= (ax_1 + by_1) z_2 - (ax_2 + by_2) z_1$$

$$= a(x_1 z_2 - x_2) + b(z_1 y_2 - z_2 y_1)$$

$$= af(\alpha, \gamma) + bf(\beta, \gamma)$$

Also, $f(z, ax + by) = f((z_1, z_2), (ax_1 + by_1, ax_2 + by_2))$

$$= z_1(ax_2 + by_2) - z_2(ax_1 + by_1)$$

$$= (ax_2 - z_2 x_1) + b(z_1 y_2 - z_2 y_1)$$

$$= af(\gamma, \alpha) + bf(\gamma, \beta)$$

Thus *f* is bilinear form on R^2.

(ii) Since we have,

$$f(x, z) = (x_1 - z_1)^2 + x_2 z_2$$

And $f(y, z) = (y_1 - z_1)^2 + y_2 z_2$

Now, $f(ax + by, z) = f((ax_1 + by_1, ax_2 + by_2), (z_1, z_2))$

$$= (ax_1 + by_1 - z_1)^2 + (ax_2 + by_2) z_2$$

Also, $af(x, z) + bf(y, z) = a(x_1 - z_1)^2 + ax_2 z_2 + b(y_1 - z_1)^2 + by_2 z_2$

$$= a(x_1 - z_1)^2 + b(y_1 - z_1)^2 + (ax_2 + by_2) z_2$$

Obviously, $f(ax + by, z) \neq af(x, z) + bf(y, z)$.

Hence f is a bilinear form on R^2.

Remarks 1: If U is an $n-$dimensional vector space with basis $\{x_1, x_2, ..., x_n\}$, if *V* is an *m*-dimensional vector space with basis $\{y_1, y_2, ..., y_m\}$ and if $\{a_{ij}\}$ is any set of *nm* scalars (*i*= 1, …, *n*; *j*=1, …, *m*} then there is a one and only one bilinear form *f* on $U \oplus V$ such that

$$f(\alpha_i, \beta_j) = a_{ij} \quad \forall\, i, j$$

14.4 *MATRIX OF A BILINEAR FORM*

Definition: Let *V* be a finite dimensional vector space and $B = \{x_1, x_2, ..., x_n\}$ be an ordered basis for V. If *f* is a bilinear form on *V*, the matrix of *f* in the ordered basis *B* is the $n \times n$ matrix $A = [a_{ij}]_{n \times n}$ such that

$f(\alpha_i, \alpha_j) = a_{ij}, i = 1, ..., n;\ \ j = 1, ..., n$

We will denote this matrix A by $[f]_B$.

Example 8: If f be the bilinear form on R^2 defined by

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

Then find the matrix of f in the ordered basis $B = \{(1, -1), (1, 1)\}$ of R^2.

Solution: Let $B = \{\alpha_1, \alpha_2\}$ where $\alpha_1 = (1, -1), \alpha_2 = (1, 1)$

We have $f(\alpha_1, \alpha_1) = f((1, -1), (1, -1)) = -1 - 1 = -2$

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

$f(\alpha_2, \alpha_1) = f((1, 1), (1, -1)) = 1 - 1 = 0$

$f(\alpha_2, \alpha_2) = f((1, 1), (1, 1)) = 1 + 1 = 2$

$$\therefore [f]_B = \begin{bmatrix} -2 & 0 \\ 0 & 2 \end{bmatrix}$$

Rank of bilinear form: The rank of bilinear form is defined as the rank of the matrix of the form in any ordered basis.

OR

The rank of bilinear form f is the rank of the matrix representation of the bilinear form.

Example 9: Let $\alpha = (x_1, x_2, x_3)$ and $\beta = (y_1, y_2, y_3)$ and the bilinear form of α and β is given as

$f(\alpha, \beta) = x_1 y_1 + 2 x_1 y_2 + 5 x_1 y_3 - 2 x_2 y_1 + x_2 y_3 - 6 x_3 y_2 + 6 x_3 y_3$. Find the matrix of f and rank of f.

Solution: We have $f(\alpha, \beta) = x_1 y_1 + 2 x_1 y_2 + 5 x_1 y_3 - 2 x_2 y_1 + x_2 y_3 - 6 x_3 y_2 + 6 x_3 y_3$

$$= [x_1 \ \ x_2 \ \ x_3] \begin{bmatrix} y_1 + 2y_2 + 5y_3 \\ -2y_1 + 0y_2 + y_3 \\ 0y_1 - 6y_2 + 6y_3 \end{bmatrix}$$

$$= [x_1 \;\; x_2 \;\; x_3] \begin{bmatrix} 1 & 2 & 5 \\ -2 & 0 & 1 \\ 0 & -6 & 6 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$$

Thus matrix of *f* is $\begin{bmatrix} 1 & 2 & 5 \\ -2 & 0 & 1 \\ 0 & -6 & 6 \end{bmatrix}$

Since the determinant of above matrix is non-zero, hence its rank is 3. Thus, rank of f is also 3.

14.5 *DEGENERATE AND NON-DEGENERATE BILINEAR FORMS*

In a vector space *V*, a bilinear form *f* is called degenerate if

(a) For each non-zero α in *V*, $f(\alpha, \beta) = 0 \;\forall\; \beta \in V$ and

(b) For each non-zero β in *V*, $f(\alpha, \beta) = 0 \;\forall\; \alpha \in V$

A bilinear form is called non-degenerate if it is not degenerate. In other sense we can say that a bilinear form *f* on a vector space V is called non-degenerate if,

(a) For each $0 \neq \alpha$ in *V*, there exist an element $\beta \in V$ s.t., $f(\alpha, \beta) \neq 0$ and

(b) For each $0 \neq \beta \in V$ in *V*, there exist an element $\alpha \in V$ s.t., $f(\alpha, \beta) \neq 0$

14.6 *SYMMETRIC BILINEAR FORMS*

Definition: In a vector space *V*, a bilinear form *f* is said to be symmetric if
$f(\alpha, \beta) = f(\beta, \alpha) \;\forall\; \alpha, \beta \in V$

Theorem 4: In a finite dimensional vector space *V*, a bilinear form *f* on *V* is symmetric if and only if its matrix *A* in some ordered basis is symmetric, i.e., $A' = A$

Proof: Let in a vector space *V*, *B* is an ordered basis and the vectors $\alpha, \beta \in V$. Let X, Y be the co-ordinates vector of α, β respectively in the ordered basis *B*. If *f* is a bilinear form on *V* and *A* is the matrix of *f* in the ordered basis *B*, then

$$f(\alpha, \beta) = X' AY ,$$

and $f(\beta,\alpha) = Y^{'}AY$

So, *f* will be symmetric *iff* $X^{'}AY = Y^{'}AX$

are all column matrices *X* and *Y*.

Now $X^{'}AY$ is a 1×1 matrix, therefore we have

$X^{'}AY = (X^{'}AY)^{'} = Y^{'}A^{'}(X^{'})^{'} = Y^{'}A^{'}X$

$\therefore$ f will be symmetric if and only if

$X^{'}AY = Y^{'}AX$ for all column matrices *X* and *Y*

i.e., $A^{'} = A$

It means *A* is symmetric.

14.7 *QUADRATIC FORMS*

Definition: In a vector space *V* over the field *F*, let *f* is a bilinear form. Then the quadratic form on *V* associated with the bilinear form *f* is the function *q* from *V* into *F* defined by:

$q(\alpha) = f(\alpha,\alpha)\ \forall\ \alpha \in V$

Theorem 5: In a vector space *V* over the field *F* whose characteristic is not equal to 2 i.e., $1+1\neq 0$. Then every symmetric bilinear form on *V* is uniquely determined by the corresponding quadratic form.

Proof: In a vector space *V* over the field *F*, let *f* is a symmetric bilinear form and *q* be the quadratic form on *V* associated with *f*. Then for each $\alpha,\beta \in V$ we have

$q(\alpha+\beta) = f(\alpha+\beta,\alpha+\beta)$

$= f(\alpha,\alpha+\beta) + f(\beta,\alpha+\beta)$

$= f(\alpha,\alpha) + f(\alpha,\beta) + f(\beta,\alpha) + f(\beta,\beta)$

$= q(\alpha) + f(\alpha,\beta) + f(\beta,\alpha) + q(\beta)$

$= q(\alpha) + (1+1) + q(\beta)$

$\therefore (1+1)f(\alpha,\beta) = q(\alpha+\beta) - q(\alpha) - q(\beta)$ ……… (1)

Thus, $f(\alpha,\beta)$ is uniquely determined by q with the help of the polarization identity (1) provided $(1+1)\neq 0$ i.e., F is not of characteristic 2.

Note: According to the polarization identity we can write,

$$f(\alpha,\beta)=\frac{1}{4}q(\alpha+\beta)-\frac{1}{4}q(\alpha-\beta)$$

Theorem 6: Let *V* is a finite dimensional vector space over a subfield of the complex numbers, and let *f* be a symmetric bilinear form on *V*. Then there is an ordered basis for *V* in which *f* is represented by a diagonal matrix.

Proof: To prove the theorem we should find an ordered basis $B=\{\alpha_1,\alpha_2,...,\alpha_n\}$ for *V* such that

$f(\alpha_i,\alpha_j)=0$ for $i\neq j$.

Case I: If $f=\hat{0}$ *or* $n=1$, the term obviously true. So we will suppose $f\neq\hat{0}$ *and* $n>1$.

Case II: If $f(\alpha,\alpha)=0\ \forall\alpha\in V$ then, $q(\alpha)=0$ for every , where *q* is quadratic form associated with *f*. So, by the polarization identity $f(\alpha,\beta)=\frac{1}{4}q(\alpha+\beta)-\frac{1}{4}q(\alpha-\beta)$ we see that

$f(\alpha,\beta)=0\ \forall\alpha\in V$ and thus $f=\hat{0}$ which assure about the contradiction. Therefore there must be a vector $\alpha_1\in V$ such that $f(\alpha_1,\alpha_1)=q(\alpha_1)\neq 0$.

Let W_1 be the one dimensional subspaces of V spanned by the vector α_1 and let W_2 be the collection of all vectors β in V such that $f(\alpha_1,\beta)=0$. Obviously W_2 is a subspace of V. Now we claim that $V=W_1\oplus W_2$. We shall first prove our claim.

At, first we have to prove that subspaces W_1 and W_2 are disjoint.

Let $\gamma\in W_1\cap W_2$ then $\gamma\in W_1$ and $\gamma\in W_2$

If $\gamma\in W_1\Rightarrow\gamma=c\alpha_1$ for some scalar *c*.

Also if $\gamma\in W_1\Rightarrow f(\alpha_1,\gamma)=0$

$\Rightarrow f(\alpha_1,c\alpha_1)=0$

$\Rightarrow cf(\alpha_1,\alpha_1)=0$ $[\because f(\alpha_1,\alpha_1)\neq 0]$

$\Rightarrow c=0$

$\Rightarrow \gamma = 0\alpha_1 = 0$

$\therefore W_1$ and W_2 are disjoint.

Now we have to only prove that $V = W_1 + W_2$. For it let us consider $\gamma \in V$. Since $f(\alpha_1, \alpha_1) \neq 0$, so put

$$\beta = \gamma - \frac{f(\gamma, \alpha_1)}{f(\alpha_1, \alpha_1)} \alpha_1$$

Thus $f(\alpha_1, \beta) = f\left(\alpha_1, \gamma - \frac{f(\gamma, \alpha_1)}{f(\alpha_1, \alpha_1)} \alpha_1\right)$

$$= f(\alpha_1, \gamma) - \frac{f(\gamma, \alpha_1)}{f(\alpha_1, \alpha_1)} f(\alpha_1, \alpha_1)$$

$= f(\alpha_1, \gamma) - f(\gamma, \alpha_1)$ [$\because f$ is symmetric]

$= 0$

$\therefore \beta \in W_2$ by definition of W_2. Also by definition of W_1 the vector $\frac{f(\gamma, \alpha_1)}{f(\alpha_1, \alpha_1)} \alpha_1$ as in W_1.

$$\therefore \gamma = \frac{f(\gamma, \alpha_1)}{f(\alpha_1, \alpha_1)} \alpha_1 + \beta \in W_1 + W_2.$$

Hence $V = W_1 + W_2$

$\therefore V = W_1 \oplus W_2$

So $\dim W_2 = V - \dim W_1 = n - 1$.

Now let g be the restriction of f from V to W_2. Then g is a symmetric bilinear form on W_2 is less than $\dim V$. Now we can consider by the induction that W_2 has a basis $\{\alpha_1, \alpha_2, ..., \alpha_n\}$ such that

$g(\alpha_i, \alpha_j) = 0, i \neq j \,(i \geq 2, j \geq 2)$

$\Rightarrow f(\alpha_i, \alpha_j) = 0, i \neq j \,(i \geq 2, j \geq 2)$ [$\because g$ is restriction on f]

So, by the definition of W_2, we have

$f(\alpha_1, \alpha_j) = 0 \;\; for \;\; j = 2, 3, ..., n$

Since $\{\alpha_1\}$ is a basis of for W_1 and $V = W_1 \oplus W_2$, therefore $\{\alpha_1, \alpha_2, ..., \alpha_n\}$ is a basis for V such that

$f(\alpha_i, \alpha_j) = 0 \;\; for\ i \neq j$

Example 10: Find the quadratic form of the symmetric matrix A = $\begin{bmatrix} 2 & -3 \\ -3 & 3 \end{bmatrix}$

Solution: Let α = (*x*, *y*). Then the quadratic form *q*(α) of *A* is given by,

$$q(\alpha) = f\left(\begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} x \\ y \end{pmatrix}\right) = [x\ y]\begin{bmatrix} 2 & -3 \\ -3 & 3 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = 2x^2 - 6xy + 3y^2$$

Thus, $2x^2 - 6xy + 3y^2$ is the quadratic form of the given matrix.

14.8 *SKEW-SYMMETRIC BILINEAR FORMS*

Definition: In a vector space *V*, a bilinear form *f* is said to be skew-symmetric if $f(\alpha, \beta) = f(\beta, \alpha)\ \forall\ \alpha, \beta \in V$

Remarks 1: Every bilinear form on the vector space *V* over a subfield *F* of the complex numbers can be uniquely expressed as the sum of a symmetric and skew-symmetric bilinear forms.

2: If *V* is a finite-dimensional vector space, then a bilinear form *f* on *V* is skew-symmetric if and only if its matrix A in some (or every) ordered basis is skew-symmetric, *i.e.*, $A' = -A$.

14.9 *REFLEXIVITY AND ORTHOGONALITY*

Definition: A bilinear form *B*: $V \times V \to K$ is called **reflexive** if,

B(*v*, *w*) = 0 implies *B*(*w*, *v*) = 0 for all *v*, *w* in *V*.

Definition: Let *B*: $V \times V \to K$ be a reflexive bilinear form. v, w in *V* are **orthogonal with respect to *B*** if *B*(*v*, *w*) = 0.

A bilinear form *B* is reflexive if and only if it is either symmetric or alternating. In the absence of reflexivity we have to distinguish left and right orthogonality. In a reflexive space the left and right radicals agree and are termed the *kernel* or the *radical* of the bilinear form: the subspace of all vectors orthogonal with every other vector. A vector **v**, with matrix representation *x*, is in the radical of a bilinear form with matrix representation *A*, if and only if $Ax = 0 \Leftrightarrow x^T A = 0$. The radical is always a subspace of *V*. It is trivial if and only if the matrix *A* is nonsingular, and thus if and only if the bilinear form is nondegenerate.

Suppose W is a subspace. Define the orthogonal complement.

$$W^{\perp} = \{v \mid B(v,w) = 0 \; \forall \; w \in W\}$$

For a non-degenerate form on a finite-dimensional space, the map $V/W \to W^{\perp}$ is bijective, and the dimension of $W^{\perp}$ is $\dim(V) - \dim(W)$

Check your progress

Problem 1: Which of the following function *f* on the vector space $V_2(R)$ are forms bilinear form where $\alpha = (x_1\, x_2), \beta = (y_1, y_2) \in V_2(R)$

(*i*) $f(\alpha, \beta) = (x_1 y_1 + x_2 y_2)$

(*ii*) $f(\alpha, \beta) = 1$

(*iii*) $f(\alpha, \beta) = x_1 y_1 + x_1 y_2 + x_2 y_1 + x_2 y_2$

Problem 2: If *f* is the bilinear form on the vector space $V_2(R)$ defined by $f(\alpha, \beta) = x_1 y_1 + x_2 y_2$, where $\alpha = (x_1\, x_2), \beta = (y_1, y_2) \in V_2(R)$. Then find the matrix of *f* for the bases $\{(1,0),(0,1)\}$

14.10 *SUMMARY*

In this unit we have learned about the unitary operator and normal operator which are essential tool in the inner product space. Also in this unit we have learned about the important concept which commonly solve out many matrix related problems like bilinear form, quadratic form, symmetric and skew-symmetric bilinear form and there related important theorems and applications. Other important concepts introduced in this unit were:

- The rank of bilinear form *f* is the rank of the matrix representation of the bilinear form
- The reason the symmetric ones are significant is that, at least when the field characteristic is not 2, the vector space admits an especially basic type of basis called an orthogonal basis for them.
- It is possible to uniquely represent every bilinear form on the vector space V over a subfield F of the complex numbers as the sum of symmetric and skew-symmetric bilinear forms.
- We can explain how the matrices associated with bilinear forms relate to coordinate vectors, how they change when the basis changes, and how to utilize them to translate back and forth between matrices and bilinear forms, just like we can with the matrices associated with linear transformations.

14.11 *GLOSSARY*

- Bilinear form
- Quadratic form

14.12 *REFERENCES*

- Gel′fand I.M. Lectures on linear algebra (1989), Courier Corporation.
- Kenneth Hoffman & Ray Kunze, Linear Algebra (2nd edition)(2015), Prentice-Hall.
- David C. Lay, Linear Algebra and its Application (3rd Edition) (2007) Pearson Education Asia, India Reprint.
- Seymour Lipshutz and Marc Lipson, Schaum's outlines "Linear Algebra" (3rd Edition)(2012), Mc Graw Hill Education.
- J. N. Sharma and A. R. Vasistha, Linear Algebra (29th Edition) (1999), Krishna Prakashan.

14.13 *SUGGESTED READING*

- Minking Eie & Shou-Te Chang (2020), A First Course In Linear Algebra, World Scientific.
- Axler, Sheldon (2015), Linear algebra done right. Springer.
- https://nptel.ac.in/courses/111106051
- https://archive.nptel.ac.in/courses/111/104/111104137
- https://epgp.inflibnet.ac.in/Home/ViewSubject?catid=ZLCHeZEhCZ8yCri36nSF3A==
- https://old.sggu.ac.in/wp-content/uploads/2020/07/Linear-AlgebraTY_506.pdf
- https://orb.binghamton.edu
- https://en.wikipedia.org/wiki/Bilinear_form#:~:text=In%20other%20words%2C%20a%20bilinear,%3D%20%CE%BBB(u%2C%20v)

14.14 *TERMINAL QUESTION*

Long Answer Type Question:

1. Let *V* denote a vector space over a field *F*. Suppose linear operator (*T*) on *V* and *f* is a bilinear form on *V*. Let *g* is a function from $V \times V$ into F and defined as

$g(\alpha, \beta) = f(T\alpha, T\beta)$. Then show that *g* is also a bilinear form on *V*.

2. Prove that in a finite dimensional vector space *V*, a bilinear form *f* on *V* is symmetric if and only if its matrix *A* in some ordered basis is symmetric, i.e., $A' = A$

3. Let V is a finite dimensional vector space over a subfield of the complex numbers, and let f be a symmetric bilinear form on V. Then prove that there is an ordered basis for V in which f is represented by a diagonal matrix.

4. If f is the bilinear form on the vector space $V_2(R)$ defined by $f(\alpha,\beta) = x_1y_1 + x_2y_2$, where $\alpha = (x_1\ x_2), \beta = (y_1, y_2) \in V_2(R)$. Then find the matrix of f for the following bases

(i) $\{(1,2),(3,4)\}$

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

5. Let f be bilinear form on $V_2(R)$ defined by $f((x_1,x_2),(y_1,y_2)) = (x_1 + x_2)(y_1 + y_2)$

(i) Find the matrix of f corresponding to the standard basis $B = \{(1,0),(0,1)\}$

(ii) Find the transition matrix from the matrix B to the basis $B' = \{(1,-1),(1,1)\}$

(iii) Find the matrix of f in the basis B'

6. Described explicitly about the all bilinear form f on $V_3(R)$ with the property defined by,

$$f(\alpha,\beta) = f(\beta,\alpha)\ \forall \alpha,\beta \in V_3(R)$$

7. If f is the bilinear form on the vector space $V_2(R)$ defined by

$f((x_1,x_2),(y_1,y_2)) = 2x_1y_1 - 3x_1y_2 + x_2y_2$. Then find matrix of f corresponding to the basis $\{(1,0),(1,1)\}$

Short Answer Type Question:

1: Consider the two vector space U and V over the same field F. Let $W = U \oplus V$. If $\hat{0}$ is the zero function W into F. Then show that $\hat{0}$ is a bilinear form on W.

2: Which of the following vectors $x = (x_1, x_2), y = (y_1, y_2)$ defined on R^2 are the bilinear form

(i) $f(x,y) = x_1y_2 - x_2y_1$

(ii) $f(x,y) = (x_1 - y_1)^2 + x_2y_2$

3: If f be the bilinear form on R^2 defined by

$f((x_1,y_1),(x_2,y_2) = x_1y_1 + x_2y_2$. Then find the matrix of f in the ordered basis $B = \{(1,-1),(1,1)\}$ of R^2.

4: Let $\alpha = (x_1, x_2, x_3)$ and $\beta = (y_1, y_2, y_3)$ and the bilinear form of α and β is given as

$f(\alpha, \beta) = x_1 y_1 + 2 x_1 y_2 + 5 x_1 y_3 - 2 x_2 y_1 + x_2 y_3 - 6 x_3 y_2 + 6 x_3 y_3$. Find the matrix of f and rank of f.

5. Find the quadratic form of the symmetric matrix A = $\begin{bmatrix} 2 & -3 \\ -3 & 3 \end{bmatrix}$

6. Find all the bilinear forms on the vector space F^2, where F is field.

Fill in the blanks

1. Each self-adjoint operator is ……………

2. Each Unitary operator is ………….

3. In a vector space V, a bilinear form f is said to be symmetric if ……….

4. The operator T is called unitary operator if ………..

5. The operator T is called unitary operator if ………..

14.15 *ANSWERS*

Answer of check your progress

1: **(i)** Bilinear form **(ii)** Not a bilinear form **(iii)** Bilinear form

2: $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$

Answer of long answer type question

4: **(i)** $\begin{bmatrix} 4 & 14 \\ 14 & 24 \end{bmatrix}$ **(ii)** $\begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix}$

6: **(i)** $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ **(ii)** $\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$ **(iii)** $\begin{bmatrix} 0 & 0 \\ 0 & 4 \end{bmatrix}$

7: $\begin{bmatrix} 2 & -1 \\ 2 & 0 \end{bmatrix}$

Answer of short answer type question

2: **(i)** f is bilinear form on R^2. **(ii)** f is bilinear form on R^2.

3: $[f]_B = \begin{bmatrix} -2 & 0 \\ 0 & 2 \end{bmatrix}$

4: $\begin{bmatrix} 1 & 2 & 5 \\ -2 & 0 & 1 \\ 0 & -6 & 6 \end{bmatrix}$

5: $2x^2 - 6xy + 3y^2$

6: Let A be any 2×2 matrix over F and B be any ordered basis of F^2. Then the bilinear forms on F^2 are precisely those obtained by $f(\alpha,\beta) = X^{'}AY$, where X, Y are the co-ordinates matrices of α and β in the ordered basis *B*.

Answer of fill in the blanks

1: Normal Opertor

2: Normal operator

3: $f(\alpha,\beta) = f(\beta,\alpha)\ \forall\ \alpha,\beta \in V$

4: $TT^* = T^*T = I$

5: $TT^* = T^*T = I$

Teen Pani Bypass Road, Transport Nagar
Uttarakhand Open University,
Haldwani, Nainital-263139
Phone No. 05946-261122, 261123
Toll free No. 18001804025
Fax No. 05946-264232,

E-mail:info@uou.ac.in
Website: https://www.uou.ac.in/