Krishna's

TEXT BOOK on

LINEAR ALGEBRA

(For B.A. and B.Sc. Vth Semester students of Kumaun University)

Kumaun University Semester Syllabus *w.e.f.* 2018-19

By

A.R. Vasishtha
Retired Head, Dept. of Mathematics
Meerut College, Meerut

A.K. Vasishtha
M.Sc., Ph.D.
C.C.S. University, Meerut

Kumaun

KRISHNA Prakashan Media (P) Ltd.
KRISHNA HOUSE, 11, Shivaji Road, Meerut-250001 (U.P.), India

Jai Shri Radhey Shyam

Dedicated

to

Lord

Krishna

Authors & Publishers

Preface

This book on **Linear Algebra** has been specially written according to the latest **Syllabus** to meet the requirements of **B.A. and B.Sc. Semester-V Students** of all colleges affiliated to Kumaun University.

The subject matter has been discussed in such a simple way that the students will find no difficulty to understand it. The proofs of various theorems and examples have been given with minute details. Each chapter of this book contains complete theory and a fairly large number of solved examples. Sufficient problems have also been selected from various university examination papers. At the end of each chapter an exercise containing objective questions has been given.

We have tried our best to keep the book free from misprints. The authors shall be grateful to the readers who point out errors and omissions which, inspite of all care, might have been there.

The authors, in general, hope that the present book will be warmly received by the students and teachers. We shall indeed be very thankful to our colleagues for their recommending this book to their students.

The authors wish to express their thanks to Mr. S.K. Rastogi, *M.D.,* Mr. Sugam Rastogi, *Executive Director,* Mrs. Kanupriya Rastogi, *Director* and entire team of **KRISHNA Prakashan Media (P) Ltd., Meerut** for bringing out this book in the present nice form.

The authors will feel amply rewarded if the book serves the purpose for which it is meant. Suggestions for the improvement of the book are always welcome.

— Authors

Syllabus

Linear Algebra

B.A./B.Sc. V Semester

Kumaun University

Fifth Semester – First Paper

B.A./B.Sc. Paper-I M.M.-60

Linear Transformations: Linear transformations, rank and nullity, Linear operators, Algebra of linear transformations, Invertible linear transformations, isomorphism; Matrix of a linear transformation, Matrix of the sum and product of linear transformations, Change of basis, similarity of matrices.

Linear Functionals: Linear functional, Dual space and dual basis, Double dual space, Annihilators, hyperspace; Transpose of a linear transformation.

Eigen vectors and Eigen values: Eigen vectors and Eigen values of a matrix, product of characteristic roots of a matrix and basic results on characteristic roots, nature of the characteristic roots of Hermitian, skew-Hermitian, unitary and orthogonal matrices, characteristic equation of a matrix, Cayley-Hamilton theorem and its use in finding inverse of a matrix.

Bilinear forms: Bilinear forms, symmetric and skew-symmetric bilinear forms, quadratic form associated with a bilinear form.

Brief Contents

Krishna's

LINEAR ALGEBRA

Chapters

Chapter

1

Linear Transformations

1 Linear Transformation or Vector Space Homomorphism

(Garhwal 2010)

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*

$$T(a\alpha + b\beta) = aT(\alpha) + bT(\beta) \qquad \ldots(1)$$

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

The condition (1) is also called **linearity property.** It can be easily seen that the condition (1) is equivalent to the condition

$$T(a\alpha + \beta) = aT(\alpha) + T(\beta)$$

for all α, β in U and all scalars a in F.

Linear Operator: Definition: *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 α, β in V and for all a, b in F. **(Kumaun 2014)**

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

Illustration 1: *The function*

$$T : V_3 (\mathbf{R}) \to V_2 (\mathbf{R})$$

defined by $T (a, b, c) = (a, b) \ \forall \ a, b, c \in \mathbf{R}$ *is a linear transformation from* $V_3 (\mathbf{R})$ *into* $V_2 (\mathbf{R})$.

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

If $a, b \in \mathbf{R}$, then

$$\begin{aligned} T (a\alpha + b\beta) &= T [a (a_1, b_1, c_1) + b (a_2, b_2, c_2)] \\ &= T (aa_1 + ba_2, ab_1 + bb_2, ac_1 + bc_2) \\ &= (aa_1 + ba_2, ab_1 + bb_2) && \text{[By def. of } T \text{]} \\ &= (aa_1, ab_1) + (ba_2, bb_2) \\ &= a (a_1, b_1) + b (a_2, b_2) \\ &= aT (a_1, b_1, c_1) + bT (a_2, b_2, c_2) \\ &= aT (\alpha) + bT (\beta). \end{aligned}$$

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

Illustration 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

$$\begin{aligned} T (aA + bB) &= P (aA + bB) Q && \text{[By def. of } T \text{]} \\ &= (aPA + bPB) Q = a \ PAQ + b \ PBQ \\ &= aT (A) + bT (B). \end{aligned}$$

$\therefore$ T is a linear transformation from V into V. Thus T is a linear operator on V.

Illustration 3: *Let* $V (F)$ *be the vector space of all polynomials over the field* F. *Let* $f (x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n \in V$ *be a polynomial of degree* n *in the indeterminate* x. *Let us define*

$$D f (x) = a_1 + 2a_2 x + \ldots + na_n x^{n-1} \text{ if } n > 1$$

and $\quad D f (x) = 0$ *if* $f (x)$ *is a constant polynomial.*

Then the correspondence D *from* V *into* V *is a linear operator on* V.

If $f (x)$ is a polynomial over the field F, then $Df (x)$ as defined above is also a polynomial over the field F. Thus if $f (x) \in V$, then $Df (x) \in V$. Therefore D is a function from V into V.

Also if $\quad f (x), g (x) \in V$ and $a, b \in F$, then

$$D [a \ f (x) + bg (x)] = a \ Df (x) + b \ Dg (x).$$

$\therefore$ D is a linear transformation from V into V.

The operator D on V is called the differentiation operator. It should be noted that for polynomials the definition of differentiation can be given purely algebraically, and does not require the usual theory of limiting processes.

Illustration 4: *Let $V(\mathbf{R})$ be the vector space of all continuous functions from $\mathbf{R}$ into $\mathbf{R}$. If $f \in V$ and we define T by*

$$(Tf)(x) = \int_0^x f(t)\, dt \ \forall \ x \in \mathbf{R},$$

then T is a linear transformation from V into V.

If f is real valued continuous function, then Tf, as defined above, is also a real valued continuous function. Thus

$$f \in V \ \Rightarrow \ Tf \in V.$$

Also the operation of integration satisfies the linearity property. Therefore T is a linear transformation from V into V.

2 Some Particular Transformations

1. Zero Transformation: *Let $U(F)$ and $V(F)$ be two vector spaces. The function T, from U into V defined by*

$$T(\alpha) = \mathbf{0} \ (\text{zero vector of } V) \ \forall \alpha \in U$$

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 $T(a\alpha + b\beta) = \mathbf{0}$ [By def. of T]

$$= a\mathbf{0} + b\mathbf{0} = aT(\alpha) + bT(\beta).$$

$\therefore$ T is a linear transformation from U into V. It is called **zero transformation** and we shall in future denote it by $\hat{\mathbf{0}}$.

2. Identity Operator: *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

$$I(a\alpha + b\beta) = a\alpha + b\beta \quad \text{[By def. of } I\text{]}$$

$$= aI(\alpha) + bI(\beta).$$

$\therefore$ I is a linear transformation from V into V. The transformation I is called **identity operator** on V and we shall always denote it by I.

3. Negative of a Linear Transformation: *Let $U(F)$ and $V(F)$ be two vector spaces. Let T be a linear transformation from U into V. The correspondence $-T$ defined by*

$$(-T)(\alpha) = -[T(\alpha)] \ \forall \ \alpha \in U$$

is a linear transformation from U into V. **(Garhwal 2008)**

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

$$(-T)(a\alpha + b\beta) = -[T(a\alpha + b\beta)] \quad \text{[By def. of } -T\text{]}$$

$$= -[aT(\alpha) + bT(\beta)]$$

[$\because$ 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.

3 Properties of Linear Transformations.

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

(i) $T(\mathbf{0}) = \mathbf{0}$ *where* $\mathbf{0}$ *on the left hand side is zero vector of* U *and* $\mathbf{0}$ *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) $T(a_1\alpha_1 + a_2\alpha_2 + \ldots + a_n\alpha_n) = a_1 T(\alpha_1) + a_2 T(\alpha_2) + \ldots + a_n T(\alpha_n)$

where $\alpha_1, \alpha_2, \ldots \alpha_n \in U$ *and* $a_1, a_2, \ldots, a_n \in F$.

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

$$T(\alpha) + \mathbf{0} = T(\alpha) \qquad [\because \ \mathbf{0} \text{ is zero vector of } V \text{ and } T(\alpha) \in V]$$

$$= T(\alpha + \mathbf{0}) \qquad [\because \ \mathbf{0} \text{ is zero vector of } U]$$

$$= T(\alpha) + T(\mathbf{0}) \qquad [\because \ T \text{ is a linear transformation}]$$

Now in the vector space V, we have

$$T(\alpha) + \mathbf{0} = T(\alpha) + T(\mathbf{0})$$

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

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

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

But $\quad T[\alpha + (-\alpha)] = T(\mathbf{0}) = \mathbf{0} \in V$. $\quad$ [By (i)]

Thus in V, we have

$$T(\alpha) + T(-\alpha) = \mathbf{0} \quad \Rightarrow \quad T(-\alpha) = -T(\alpha).$$

(iii) We have $T(\alpha - \beta) = T[\alpha + (-\beta)] = T(\alpha) + T(-\beta)$ $\quad [\because \ T$ is linear$]$

$$= T(\alpha) + [-T(\beta)] \qquad \text{[By (ii)]}$$

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

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

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

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

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

Note: On account of this property sometimes we say that a linear transformation **preserves linear combinations.**

4 Range and Null Space of a Linear Transformation

Range of a Linear Transformation: **(Kumaun 2007, 14; Gorakhpur 15)**

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 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.*,

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

Theorem 1: *If $U(F)$ and $V(F)$ are two vector spaces and T is a linear transformation from U into V, then range of T is a subspace of V.* **(Kumaun 2009; Gorakhpur 14)**

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

$$a\beta_1 + b\beta_2 = a\,T(\alpha_1) + b\,T(\alpha_2)$$
$$= T(a\alpha_1 + b\alpha_2) \qquad [\because \ T \text{ is a linear transformation}]$$

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

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.

Null space of Linear Transformation: **(Kumaun 2007, 14; Gorakhpur 15)**

Definition: *Let $U(F)$ and $V(F)$ be two vector spaces and let T be a linear transformation from U into V. Then the null space of T written as $N(T)$ is the set of all vectors α in U such that $T(\alpha) = \mathbf{0}$ (zero vector of V). Thus*

$$N(T) = \{\alpha \in U : T(\alpha) = \mathbf{0} \in V\}.$$

If we regard the linear transformation T from U into V as a vector space homomorphism of U into V, then the null space of T is also called the **kernel of** T.

Theorem 2: *If $U(F)$ and $V(F)$ are two vector spaces 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.*

(Gorakhpur 2014)

Proof: Let $N(T) = \{\alpha \in U : T(\alpha) = \mathbf{0} \in V\}$.

Since $T(\mathbf{0}) = \mathbf{0} \in V$, therefore at least $\mathbf{0} \in N(T)$.

Thus $N(T)$ is a non-empty subset of U.

Let $\alpha_1, \alpha_2 \in N(T)$. Then $T(\alpha_1) = \mathbf{0}$ and $T(\alpha_2) = \mathbf{0}$.

Let $a, b \in F$. Then $a\alpha_1 + b\alpha_2 \in U$ and

$$T(a\alpha_1 + b\alpha_2) = a\,T(\alpha_1) + b\,T(\alpha_2) \quad [\because T \text{ is a linear transformation}]$$
$$= a\mathbf{0} + b\mathbf{0} = \mathbf{0} + \mathbf{0} = \mathbf{0} \in V.$$

$\therefore \qquad a\alpha_1 + b\alpha_2 \in N(T)$.

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.

5 Rank and Nullity of a Linear Transformation

(Gorakhpur 2010; Garhwal 10B)

Theorem 1: *Let T be a linear transformation from a vector space $U(F)$ into a vector space $V(F)$. If U is finite dimensional, then the range of T is a finite dimensional subspace of V.*

Proof: Since U is finite dimensional, therefore there exists a finite subset of U, say $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ which spans U.

Let $\beta \in$ range of T. Then there exists α in U such that $T(\alpha) = \beta$.

Now $\alpha \in U \Rightarrow \exists\, a_1, a_2, \ldots, a_n \in F$ such that

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

$$\Rightarrow \qquad T(\alpha) = T(a_1\alpha_1 + a_2\alpha_2 + \ldots + a_n\alpha_n)$$

$$\Rightarrow \qquad \beta = a_1 T(\alpha_1) + a_2 T(\alpha_2) + \ldots + a_n T(\alpha_n). \qquad \ldots(1)$$

Now the vectors $T(\alpha_1), T(\alpha_2), \ldots, T(\alpha_n)$ are in the range of T. If β is any vector in the range of T, then from (1), we see that β can be expressed as a linear combination of $T(\alpha_1), T(\alpha_2), \ldots, T(\alpha_n)$.

Therefore range of T is spanned by the vectors

$$T(\alpha_1), T(\alpha_2), \ldots, T(\alpha_n).$$

Hence range of T is finite dimensional.

Now we are in a position to define **rank** and **nullity** of a linear transformation.

Rank and Nullity of a Linear Transformation:

(Garhwal 2007, 10; Gorakhpur 2012)

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.,

$$\nu(T) = \dim N(T).$$

Theorem 2: *Let U and V be vector spaces over the field F and let T be a linear transformation from U into V. Suppose that U is finite dimensional. Then*

$$\text{rank}(T) + \text{nullity}(T) = \dim U.$$

(Gorakhpur 2010, 12; Kumaun 13, 14)

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 $\dim N = \text{nullity}(T) = k$ and let $\{\alpha_1, \alpha_2, \ldots, \alpha_k\}$ be a basis for N.

Since $\{\alpha_1, \alpha_2, \ldots, \alpha_k\}$ is a linearly independent subset of U, therefore we can extend it to form a basis of U. Let dim $U = n$ and let

$$\{\alpha_1, \alpha_2, \ldots, \alpha_k, \alpha_{k+1}, \ldots, \alpha_n\}$$

be a basis for U.

The vectors $T(\alpha_1), T(\alpha_2), \ldots, T(\alpha_k), T(\alpha_{k+1}), \ldots, T(\alpha_n)$ are in the range of T. We claim that $\{T(\alpha_{k+1}), T(\alpha_{k+2}), \ldots, T(\alpha_n)\}$ is a basis for the range of T.

(i) First we shall prove that the vectors

$T(\alpha_{k+1}), T(\alpha_{k+2}), \ldots, T(\alpha_n)$ span the range of T.

Let $\beta \in$ range of T. Then there exists $\alpha \in U$ such that

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

Now $\alpha \in U \Rightarrow \exists\, a_1, a_2, \ldots, a_n \in F$ such that

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

$\Rightarrow$ $T(\alpha) = T(a_1\alpha_1 + a_2\alpha_2 + \ldots + a_n\alpha_n)$

$\Rightarrow$ $\beta = a_1 T(\alpha_1) + \ldots + a_k T(\alpha_k) + a_{k+1} T(\alpha_{k+1}) + \ldots + a_n T(\alpha_n)$

$\Rightarrow$ $\beta = a_{k+1} T(\alpha_{k+1}) + a_{k+2} T(\alpha_{k+2}) + \ldots + a_n T(\alpha_n)$

$[\because \alpha_1, \alpha_2, \ldots, \alpha_k \in N \Rightarrow T(\alpha_1) = \mathbf{0}, \ldots, T(\alpha_k) = \mathbf{0}]$

$\therefore$ the vectors $T(\alpha_{k+1}), \ldots, T(\alpha_n)$ span the range of T.

(ii) Now we shall show that the vectors $T(\alpha_{k+1}), \ldots, T(\alpha_n)$ are linearly independent.

Let $c_{k+1}, \ldots, c_n \in F$ be such that

$$c_{k+1} T(\alpha_{k+1}) + \ldots + c_n T(\alpha_n) = \mathbf{0}$$

$\Rightarrow$ $T(c_{k+1}\alpha_{k+1} + \ldots + c_n\alpha_n) = \mathbf{0}$

$\Rightarrow$ $c_{k+1}\alpha_{k+1} + \ldots + c_n\alpha_n \in$ null space of T *i.e.*, N

$\Rightarrow$ $c_{k+1}\alpha_{k+1} + \ldots + c_n\alpha_n = b_1\alpha_1 + b_2\alpha_2 + \ldots + b_k\alpha_k$

for some $b_1, b_2, \ldots, b_k \in F$

[$\because$ Each vector in N can be expressed as a linear combination of the vectors $\alpha_1, \ldots, \alpha_k$ forming a basis of N]

$\Rightarrow$ $b_1\alpha_1 + \ldots + b_k\alpha_k - c_{k+1}\alpha_{k+1} - \ldots - c_n\alpha_n = \mathbf{0}$

$\Rightarrow$ $b_1 = \ldots = b_k = c_{k+1} = \ldots = c_n = 0$

[$\because$ $\alpha_1, \alpha_2, \ldots, \alpha_k, \alpha_{k+1}, \ldots, \alpha_n$ are linearly independent, being a basis for U]

$\Rightarrow$ the vectors $T(\alpha_{k+1}), \ldots, T(\alpha_n)$ are linearly independent.

$\therefore$ The vectors $T(\alpha_{k+1}), \ldots, T(\alpha_n)$ 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,

$$\rho(T) + \nu(T) = n.$$

Illustrative Examples

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

$$T(a_1, a_2, a_3) = (3a_1 - 2a_2 + a_3, a_1 - 3a_2 - 2a_3)$$ **(Kumaun 2015)**

is a linear transformation from $V_3(\mathbf{R})$ into $V_2(\mathbf{R})$.

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

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

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

Also let $a, b \in \mathbf{R}$. Then $a\alpha + b\beta \in V_3(\mathbf{R})$. We have

$$\begin{aligned} T(a\alpha + b\beta) &= T[a(a_1, a_2, a_3) + b(b_1, b_2, b_3)] \\ &= T(aa_1 + bb_1, aa_2 + bb_2, aa_3 + bb_3) \\ &= (3(aa_1 + bb_1) - 2(aa_2 + bb_2) + aa_3 + bb_3, \\ &\qquad aa_1 + bb_1 - 3(aa_2 + bb_2) - 2(aa_3 + bb_3)) \\ &= (a(3a_1 - 2a_2 + a_3) + b(3b_1 - 2b_2 + b_3), \\ &\qquad a(a_1 - 3a_2 - 2a_3) + b(b_1 - 3b_2 - 2b_3)) \\ &= a(3a_1 - 2a_2 + a_3, a_1 - 3a_2 - 2a_3) \\ &\qquad + b(3b_1 - 2b_2 + b_3, b_1 - 3b_2 - 2b_3) \\ &= aT(\alpha) + bT(\beta). \end{aligned}$$

Hence T is a linear transformation from $V_3(\mathbf{R})$ into $V_2(\mathbf{R})$.

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

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

is a linear transformation from $V_2(\mathbf{R})$ into $V_3(\mathbf{R})$. Find the range, rank, null-space and nullity of T. **(Kumaun 2009, 13; Gorakhpur 12, 14)**

Solution: Let $\alpha = (a_1, b_1), \beta = (a_2, b_2) \in V_2(\mathbf{R})$.

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 $a, b \in \mathbf{R}$.

Then $a\alpha + b\beta \in V_2(\mathbf{R})$

and
$$\begin{aligned} T(a\alpha + b\beta) &= T[a(a_1, b_1) + b(a_2, b_2)] \\ &= T(aa_1 + ba_2, ab_1 + bb_2) \\ &= (aa_1 + ba_2 + ab_1 + bb_2, aa_1 + ba_2 - ab_1 - bb_2, ab_1 + bb_2) \\ &= (a[a_1 + b_1] + b[a_2 + b_2], a[a_1 - b_1] + b[a_2 - b_2], ab_1 + bb_2) \\ &= a(a_1 + b_1, a_1 - b_1, b_1) + b(a_2 + b_2, a_2 - b_2, b_2) \\ &= aT(\alpha) + bT(\beta). \end{aligned}$$

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

Now $\{(1, 0), (0, 1)\}$ is a basis for $V_2(\mathbf{R})$.

We have $T(1, 0) = (1 + 0, 1 - 0, 0) = (1, 1, 0)$

and $T(0, 1) = (0 + 1, 0 - 1, 1) = (1, -1, 1)$.

The vectors $T(1,0), T(0,1)$ span the range of T.

Thus the range of T is the subspace of V_3 ($\mathbf{R}$) spanned by the vectors $(1,1,0), (1,-1,1)$.

Now the vectors $(1,1,0), (1,-1,1) \in V_3$ ($\mathbf{R}$) are linearly independent because if $x, y \in \mathbf{R}$, then

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

$$\Rightarrow \quad (x+y, x-y, y) = (0,0,0)$$

$$\Rightarrow \quad x+y=0, x-y=0, y=0 \quad \Rightarrow \quad x=0, y=0.$$

$\therefore$ The vectors $(1,1,0), (1,-1,1)$ form a basis for the range of T.

Hence, rank T = dim of range of $T = 2$.

Nullity of T = dim of V_2 ($\mathbf{R}$) $-$ rank $T = 2 - 2 = 0$.

$\therefore$ Null space of T must be the zero subspace of V_2 ($\mathbf{R}$).

Otherwise, $(a,b) \in$ null space of T

$$\Rightarrow \quad T(a,b) = (0,0,0) \quad \Rightarrow \quad (a+b, a-b, b) = (0,0,0)$$

$$\Rightarrow \quad a+b=0, a-b=0, b=0 \quad \Rightarrow \quad a=0, b=0.$$

$\therefore$ $(0,0)$ is the only element of V_2 ($\mathbf{R}$) which belongs to the null space of T.

$\therefore$ Null space of T is the zero subspace of V_2 ($\mathbf{R}$).

Example 3: *Let V be the vector space of all $n \times n$ matrices over the field F and let B be a fixed $n \times n$ matrix. If $T(A) = AB - BA \ \forall \ A \in V$, verify that T is a linear transformation from V into V.* **(Kumaun 2013)**

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

$$\begin{aligned} T(aA_1 + bA_2) &= (aA_1 + bA_2) B - B(aA_1 + bA_2) \\ &= aA_1B + bA_2B - aBA_1 - bBA_2 \\ &= a(A_1B - BA_1) + b(A_2B - BA_2) \\ &= aT(A_1) + bT(A_2). \end{aligned}$$

$\therefore$ 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. Give an example of such a linear transformation.*

Solution: Let N be the null space of T. Then N is also the range of T.

Now $\rho(T) + \nu(T) = \dim V$

i.e., dim of range of T + dim 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 $T : V_2(\mathbf{R}) \to V_2(\mathbf{R})$ be defined by

$$T(a,b) = (b,0) \ \forall \ a, b \in \mathbf{R}.$$

Let $\alpha = (a_1, b_1), \beta = (a_2, b_2) \in V_2$ (**R**) and let $x, y \in \mathbf{R}$.

Then
$$T(x\alpha + y\beta) = T(x(a_1, b_1) + y(a_2, b_2))$$
$$= T(xa_1 + ya_2, xb_1 + yb_2) = (xb_1 + yb_2, 0)$$
$$= (xb_1, 0) + (yb_2, 0) = x(b_1, 0) + y(b_2, 0)$$
$$= xT(a_1, b_1) + yT(a_2, b_2) = xT(\alpha) + yT(\beta).$$

$\therefore$ 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 spanning 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 the range of $T = \{a(1, 0) : a \in \mathbf{R}\} = \{(a, 0) : a \in \mathbf{R}\}$.

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 $T = \{(a, 0) : a \in \mathbf{R}\}$.

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 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., $R(T) \cap N(T) = \{\mathbf{0}\}$.*

(ii) $T[T(\alpha)] = \mathbf{0} \Rightarrow T(\alpha) = \mathbf{0}$.

Solution: First we shall show that (i) $\Rightarrow$ (ii).

We have $T[T(\alpha)] = \mathbf{0} \Rightarrow T(\alpha) \in N(T)$

$\Rightarrow$ $T(\alpha) \in R(T) \cap N(T)$ $[\because \alpha \in V \Rightarrow T(\alpha) \in R(T)]$

$\Rightarrow$ $T(\alpha) = \mathbf{0}$ because $R(T) \cap N(T) = \{\mathbf{0}\}$.

Now we shall show that (ii) $\Rightarrow$ (i).

Let $\alpha \neq \mathbf{0}$ and $\alpha \in R(T) \cap N(T)$.

Then $\alpha \in R(T)$ and $\alpha \in N(T)$.

Since $\alpha \in N(T)$, therefore $T(\alpha) = \mathbf{0}$. ...(1)

Also $\alpha \in R(T) \Rightarrow \exists \ \beta \in V$ such that $T(\beta) = \alpha$.

Now $T(\beta) = \alpha \Rightarrow T[T(\beta)] = T(\alpha) = \mathbf{0}$ [From (1)]

Thus $\exists \ \beta \in V$ such that $T[T(\beta)] = \mathbf{0}$ but $T(\beta) = \alpha \neq \mathbf{0}$.

This contradicts the given hypothesis (ii).

Therefore there exists no $\alpha \in R(T) \cap N(T)$ such that $\alpha \neq \mathbf{0}$.

Hence $R(T) \cap N(T) = \{\mathbf{0}\}$.

Example 6: *Consider the basis $S = \{\alpha_1, \alpha_2, \alpha_3\}$ of $\mathbf{R}^3$ where*

$$\alpha_1 = (1, 1, 1), \alpha_2 = (1, 1, 0), \alpha_3 = (1, 0, 0).$$

Express $(2, -3, 5)$ in terms of the basis $\alpha_1, \alpha_2, \alpha_3$.

Let $T : \mathbf{R}^3 \to \mathbf{R}^2$ *be defined as* $T(\alpha_1) = (1, 0), T(\alpha_2) = (2, -1), T(\alpha_3) = (4, 3)$.
Find $T(2, -3, 5)$. **(Meerut 2003)**

Solution: Let $(2, -3, 5) = a\alpha_1 + b\alpha_2 + c\alpha_3$

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

Then $a + b + c = 2, a + b = -3, a = 5.$

Solving these equations, we get

$$a = 5, b = -8, c = 5.$$

$\therefore$ $(2, -3, 5) = 5\alpha_1 - 8\alpha_2 + 5\alpha_3$.

Now $T(2, -3, 5) = T(5\alpha_1 - 8\alpha_2 + 5\alpha_3)$

$$= 5T(\alpha_1) - 8T(\alpha_2) + 5T(\alpha_3)$$

[$\because$ T is a linear transformation]

$$= 5(1, 0) - 8(2, -1) + 5(4, 3)$$
$$= (5, 0) - (16, -8) + (20, 15)$$
$$= (9, 23).$$

Comprehensive Exercise 1

1. Show that the mapping $T : V_3(\mathbf{R}) \to V_2(\mathbf{R})$ defined as
$$T(a_1, a_2, a_3) = (a_1 - a_2, a_1 - a_3)$$
is a linear transformation.

2. Show that the mapping $T : \mathbf{R}^3 \to \mathbf{R}^2$ defined as $T(x, y, z) = (z, x + y)$ is linear. **(Gorakhpur 2013)**

3. Show that the following functions are linear :
   (i) $T : \mathbf{R}^2 \to \mathbf{R}^2$ defined by $T(a, b) = (b, a)$
   (ii) $T : \mathbf{R}^2 \to \mathbf{R}^2$ defined by $T(a, b) = (a + b, a)$
   (iii) $T : \mathbf{R}^3 \to \mathbf{R}$ defined by $T(a, b, c) = 2a - 3b + 4c$.

4. Show that the following mappings T are not linear :
   (i) $T : \mathbf{R}^2 \to \mathbf{R}$ defined by $T(x, y) = xy$;
   (ii) $T : \mathbf{R}^2 \to \mathbf{R}^2$ defined by $T(x, y) = (1 + x, y)$; **(Kumaun 2008)**
   (iii) $T : \mathbf{R}^3 \to \mathbf{R}^2$ defined by $T(x, y, z) = (|x|, 0)$;
   (iv) $T : \mathbf{R}^2 \to \mathbf{R}$ defined by $T(x, y) = |x - y|$. **(Kumaun 2010)**

5. Show that the mapping $T : \mathbf{R}^2 \to \mathbf{R}^3$ defined as
$$T(a, b) = (a - b, b - a, -a)$$
is a linear transformation from $\mathbf{R}^2$ into $\mathbf{R}^3$. Find the range, rank, null space and nullity of T. **(Gorakhpur 2010, 11)**

6. Let F be the field of complex numbers and let T be the function from F^3 into F^3 defined by

$$T(x_1, x_2, x_3) = (x_1 - x_2 + 2x_3, 2x_1 + x_2 - x_3, -x_1 - 2x_2).$$

Verify that T is a linear transformation. Describe the null space of T.

(Meerut 2002; Garhwal 07; Kumaun 09)

7. Let F be the field of complex numbers and let T be the function from F^3 into F^3 defined by

$$T(a, b, c) = (a - b + 2c, 2a + b, -a - 2b + 2c).$$

Show that T is a linear transformation. Find also the rank and the nullity of T.

8. Let $T : \mathbf{R}^3 \to \mathbf{R}^3$ be the linear transformation defined by :

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

Find a basis and the dimension of (i) the range of T (ii) the null space of T.

(Garhwal 2011)

9. Let $T : V_4(\mathbf{R}) \to V_3(\mathbf{R})$ be a linear transformation defined by

$$T(a, b, c, d) = (a - b + c + d, a + 2c - d, a + b + 3c - 3d).$$

Then obtain the basis and dimension of the range space of T and null space of T.

10. Let V be the vector space of polynomials in x over $\mathbf{R}$. Let $D : V \to V$ be the differential operator : $D(f) = \dfrac{df}{dx} \cdot$ Find the image (*i.e.* range) and kernel of D.

11. Let V be the vector space of $n \times n$ matrices over the field F and let E be an arbitrary matrix in V. Let $T : V \to V$ be defined as

$$T(A) = AE + EA,\ A \in V.$$

Show that T is linear.

12. Let V be the vector space of 2×2 matrices over $\mathbf{R}$ and let $M = \begin{bmatrix} 1 & -1 \\ -2 & 2 \end{bmatrix} \cdot$ Let $T : V \to V$ be the linear function defined by $T(A) = MA$ for $A \in V$. Find a basis and the dimension of

(i) the kernel of T and (ii) the range of T.

13. Let V be the vector space of 2×2 matrices over $\mathbf{R}$ and let $M = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix} \cdot$ Let $T : V \to V$ be the linear transformation defined by $T(A) = AM - MA$. Find a basis and the dimension of the kernel of T.

14. Let V be the space of $n \times 1$ matrices over a field F and let W be the space of $m \times 1$ matrices over F. Let A be a fixed $m \times n$ matrix over F and let T be the linear transformation from V into W defined by $T(X) = AX$.

Prove that T is the zero transformation if and only if A is the zero matrix.

15. Let $U(F)$ and $V(F)$ be two vector spaces and let T_1, T_2 be two linear transformations from U to V. Let x, y be two given elements of F. Then the mapping T defined as

$$T(\alpha) = x\, T_1(\alpha) + y\, T_2(\alpha)\ \forall\ \alpha \in U$$

is a linear transformation from U into V.

Answers 1

5. Null space of $T = \{\mathbf{0}\}$; nullity of $T = 0$, rank $T = 2$

The set $\{(1, -1, -1), (-1, 1, 0)\}$ is a basis set for $R(T)$

6. Null space of $T = \{\mathbf{0}\}$

7. Rank $T = 2$; nullity of $T = 1$

8. (i) $\{(1, 0, 1), (2, 1, 1)\}$ is a basis of $R(T)$ and dim. $R(T) = 2$

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

9. $\{(1, 1, 1), (0, 1, 2)\}$ is a basis of $R(T)$ and dim. $R(T) = 2$

$\{(1, 2, 0, 1), (2, 1, -1, 0)\}$ is a basis of $N(T)$ and dim. $N(T) = 2$

10. The image of D is the whole space V

The kernel of D is the set of constant polynomials

12. $\left\{\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}\right\}$ is a basis of the kernel T and dim (kernel T) $= 2$

$\left\{\begin{bmatrix} 1 & 0 \\ -2 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & -2 \end{bmatrix}\right\}$ is a basis for $R(T)$ and dim $R(T) = 2$

13. $\left\{\begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right\}$ is a basis of the kernel of T and dim (kernel T) $= 2$

6 Linear Transformations as Vectors

Let $L(U, V)$ be the set of all linear transformations from a vector space $U(F)$ into a vector space $V(F)$. Sometimes we denote this set by Hom (U, V). Now we want to impose a vector space structure on the set $L(U, V)$ over the same field F. For this purpose we shall have to suitably define addition in $L(U, V)$ and scalar multiplication in $L(U, V)$ over F.

Algebra of Linear Transformations:

Theorem 1: *Let U and V be vector spaces over the field F. Let T_1 and T_2 be linear transformations from U into V. The function $T_1 + T_2$ defined by*

$$(T_1 + T_2)(\alpha) = T_1(\alpha) + T_2(\alpha) \ \forall \ \alpha \in U$$

is a linear transformation from U into V. If c is any element of F, the function (cT) defined by

$$(cT)(\alpha) = cT(\alpha) \ \forall \ \alpha \in U$$

is a linear transformation from U into V. The set $L(U, V)$ of all linear transformations from U into V, together with the addition and scalar multiplication defined above is a vector space over the field F.

Proof: Suppose T_1 and T_2 are linear transformations from U into V and we define $T_1 + T_2$ as follows :

$$(T_1 + T_2)(\alpha) = T_1(\alpha) + T_2(\alpha) \; \forall \; \alpha \in U. \qquad \ldots(1)$$

Since $T_1(\alpha) + T_2(\alpha) \in V$, therefore $T_1 + T_2$ is a function from U into V.

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

$$(T_1 + T_2)(a\alpha + b\beta) = T_1(a\alpha + b\beta) + T_2(a\alpha + b\beta) \qquad \text{[By (1)]}$$
$$= [aT_1(\alpha) + bT_1(\beta)] + [aT_2(\alpha) + bT_2(\beta)]$$
$$[\because T_1 \text{ and } T_2 \text{ are linear transformations}]$$
$$= a[T_1(\alpha) + T_2(\alpha)] + b[T_1(\beta) + T_2(\beta)] \qquad [\because V \text{ is a vector space}]$$
$$= a(T_1 + T_2)(\alpha) + b(T_1 + T_2)(\beta) \qquad \text{[By (1)]}$$

$\therefore$ $T_1 + T_2$ is a linear transformation from U into V. Thus

$$T_1, T_2 \in L(U, V) \Rightarrow T_1 + T_2 \in L(U, V).$$

Therefore $L(U, V)$ is closed with respect to addition defined in it.

Again let $T \in L(U, V)$ and $c \in F$. Let us define cT as follows :

$$(cT)(\alpha) = cT(\alpha) \; \forall \; \alpha \in U. \qquad \ldots(2)$$

Since $cT(\alpha) \in V$, therefore cT is a function from U into V.

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

$$(cT)(a\alpha + b\beta) = cT(a\alpha + b\beta) \qquad \text{[By (2)]}$$
$$= c[aT(\alpha) + bT(\beta)] \qquad [\because T \text{ is a linear transformation}]$$
$$= c[aT(\alpha)] + c[bT(\beta)] = (ca)T(\alpha) + (cb)T(\beta)$$
$$= (ac)T(\alpha) + (bc)T(\beta) = a[cT(\alpha)] + b[cT(\beta)]$$
$$= a[(cT)(\alpha)] + b[(cT)(\beta)].$$

$\therefore$ cT is a linear transformation from U into V. Thus

$$T \in L(U, V) \text{ and } c \in F \Rightarrow cT \in L(U, V).$$

Therefore $L(U, V)$ is closed with respect to scalar multiplication defined in it.

Associativity of addition in $L(U, V)$:

Let $T_1, T_2, T_3 \in L(U, V)$. If $\alpha \in U$, then

$$[T_1 + (T_2 + T_3)](\alpha) = T_1(\alpha) + (T_2 + T_3)(\alpha)$$
$$\text{[By (1) } i.e., \text{ by def. of addition in } L(U, V)]$$
$$= T_1(\alpha) + [T_2(\alpha) + T_3(\alpha)] \qquad \text{[By (1)]}$$
$$= [T_1(\alpha) + T_2(\alpha)] + T_3(\alpha)\backslash$$
$$[\because \text{Addition in } V \text{ is associative}]$$
$$= (T_1 + T_2)(\alpha) + T_3(\alpha) \qquad \text{[By (1)]}$$
$$= [(T_1 + T_2) + T_3](\alpha) \qquad \text{[By (1)]}$$

$\therefore$ $$T_1 + (T_2 + T_3) = (T_1 + T_2) + T_3$$
[By def. of equality of two functions]

Commutativity of addition in $L(U, V)$: Let $T_1, T_2 \in L(U, V)$. If α is any element of U, then

$$(T_1 + T_2)(\alpha) = T_1(\alpha) + T_2(\alpha) \qquad \text{[By (1)]}$$
$$= T_2(\alpha) + T_1(\alpha) \qquad [\because \text{Addition in } V \text{ is commutative}]$$
$$= (T_2 + T_1)(\alpha) \qquad \text{[By (1)]}$$

$\therefore$ $$T_1 + T_2 = T_2 + T_1 \qquad \text{[By def. of equality of two functions]}$$

Existence of additive identity in $L(U, V)$: Let $\hat{\mathbf{0}}$ be the zero function from U into V *i.e.*,

$$\hat{\mathbf{0}}(\alpha) = \mathbf{0} \in V \;\forall\; \alpha \in U.$$

Then $\hat{\mathbf{0}} \in L(U, V)$. If $T \in L(U, V)$ and $\alpha \in U$, we have

$$(\hat{\mathbf{0}} + T)(\alpha) = \hat{\mathbf{0}}(\alpha) + T(\alpha) \qquad \text{[By (1)]}$$
$$= \mathbf{0} + T(\alpha) \qquad \text{[By def. of } \hat{\mathbf{0}}\text{]}$$
$$= T(\alpha) \qquad [\mathbf{0} \text{ being additive identity in } V]$$

$\therefore \quad \hat{\mathbf{0}} + T = T \;\forall\; T \in L(U, V).$

$\therefore \quad \hat{\mathbf{0}}$ is the additive identity in $L(U, V)$.

Existence of additive inverse of each element in $L(U, V)$:

Let $T \in L(U, V)$. Let us define $-T$ as follows :

$$(-T)(\alpha) = -T(\alpha) \;\forall\; \alpha \in U.$$

Then $-T \in L(U, V)$.

If $\alpha \in U$, we have

$$(-T + T)(\alpha) = (-T)(\alpha) + T(\alpha) \qquad \text{[By def. of addition in } L(U, V)\text{]}$$
$$= -T(\alpha) + T(\alpha) \qquad \text{[By def. of } -T\text{]}$$
$$= \mathbf{0} \in V$$
$$= \hat{\mathbf{0}}(\alpha) \qquad \text{[By def. of } \hat{\mathbf{0}}\text{]}$$

$\therefore \quad -T + T = \hat{\mathbf{0}}$ for every $T \in L(U, V)$.

Thus each element in $L(U, V)$ possesses additive inverse.

Therefore $L(U, V)$ is an **abelian group** with respect to addition defined in it.

Further we make the following observations :

(i) Let $c \in F$ and $T_1, T_2 \in L(U, V)$. If α is any element in U, we have

$$[c(T_1 + T_2)](\alpha) = c[(T_1 + T_2)(\alpha)] \qquad \text{[By (2) } i.e., \text{ by def. of scalar multiplication in } L(U, V)\text{]}$$
$$= c[T_1(\alpha) + T_2(\alpha)] \qquad \text{[By (1)]}$$
$$= cT_1(\alpha) + cT_2(\alpha) \qquad [\because\; c \in F \text{ and } T_1(\alpha), T_2(\alpha) \in V \text{ which is a vector space}]$$
$$= (cT_1)(\alpha) + (cT_2)(\alpha) \qquad \text{[By (2)]}$$
$$= (cT_1 + cT_2)(\alpha) \qquad \text{[By (1)]}$$

$\therefore \quad c(T_1 + T_2) = cT_1 + cT_2.$

(ii) Let $a, b \in F$ and $T \in L(U, V)$. If $\alpha \in U$, we have

$$[(a + b)T](\alpha) = (a + b)T(\alpha) \qquad \text{[By (2)]}$$
$$= aT(\alpha) + bT(\alpha) \qquad [\because\; V \text{ is a vector space}]$$
$$= (aT)(\alpha) + (bT)(\alpha) \qquad \text{[By (2)]}$$
$$= (aT + bT)(\alpha) \qquad \text{[By (1)]}$$

$\therefore \quad (a + b)T = aT + bT.$

(iii) Let $a, b \in F$ and $T \in L(U, V)$. If $\alpha \in U$, we have

$$[(ab)T](\alpha) = (ab)T(\alpha) \qquad \text{[By (2)]}$$
$$= a[bT(\alpha)] \qquad [\because V \text{ is a vector space}]$$
$$= a[(bT)(\alpha)] \qquad \text{[By (2)]}$$
$$= [a(bT)](\alpha) \qquad \text{[By (2)]}$$

$\therefore \qquad (ab)T = a(bT).$

(iv) Let $1 \in F$ and $T \in L(U, V)$. If $\alpha \in U$, we have

$$(1T)(\alpha) = 1T(\alpha) \qquad \text{[By (2)]}$$
$$= T(\alpha) \qquad [\because V \text{ is a vector space}]$$

$\therefore \qquad 1T = T.$

Hence $L(U, V)$ is a vector space over the field F.

Note: If in place of the vector space V, we take U, then we observe that the set of all linear operators on U forms a vector space with respect to addition and scalar multiplication defined as above.

Dimension of $L(U, V)$: Now we shall prove that if $U(F)$ and $V(F)$ are finite dimensional, then the vector space of linear transformations from U into V is also finite dimensional. For this purpose we shall require an **important result** which we prove in the following theorem :

Theorem 2: *Let U be a finite dimensional vector space over the field F and let $B = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ be an ordered basis for U. Let V be a vector space over the same field F and let $\beta_1, \ldots, \beta_n$ be any vectors in V. Then there exists a unique linear transformation T from U into V such that*

$$T(\alpha_i) = \beta_i, \; i = 1, 2, \ldots, n.$$

Proof: Existence of T:

Let $\qquad \alpha \in U.$

Since $B = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ is a basis for U, therefore there exist unique scalars $x_1, x_2, \ldots, x_n$ such that

$$\alpha = x_1\alpha_1 + x_2\alpha_2 + \ldots + x_n\alpha_n.$$

For this vector α, let us define

$$T(\alpha) = x_1\beta_1 + x_2\beta_2 + \ldots + x_n\beta_n.$$

Obviously $T(\alpha)$ as defined above is a unique element of V. Therefore T is a well-defined rule for associating with each vector α in U a unique vector $T(\alpha)$ in V. Thus T is a function from U into V.

The unique representation of $\alpha_i \in U$ as a linear combination of the vectors belonging to the basis B is

$$\alpha_i = 0\alpha_1 + 0\alpha_2 + \ldots + 1\alpha_i + 0\alpha_{i+1} + \ldots + 0\alpha_n.$$

Therefore according to our definition of T, we have

$$T(\alpha_i) = 0\beta_1 + 0\beta_2 + \ldots + 1\beta_i + 0\beta_{i+1} + \ldots + 0\beta_n$$

i.e., $\qquad T(\alpha_i) = \beta_i, \; i = 1, 2, \ldots, n.$

Now to show that T is a linear transformation.

Let $a, b \in F$ and $\alpha, \beta \in U$. Let

$$\alpha = x_1\alpha_1 + \ldots + x_n\,\alpha_n \quad \text{and} \quad \beta = y_1\alpha_1 + \ldots + y_n\,\alpha_n.$$

Then $T\,(a\alpha + b\beta) = T\,[a\,(x_1\alpha_1 + \ldots + x_n\,\alpha_n) + b\,(y_1\alpha_1 + \ldots + y_n\,\alpha_n)]$

$$= T\,[(ax_1 + by_1)\,\alpha_1 + \ldots + (ax_n + by_n)\,\alpha_n]$$

$$= (ax_1 + by_1)\,\beta_1 + \ldots + (ax_n + by_n)\,\beta_n \qquad \text{[By def. of } T\text{]}$$

$$= a\,(x_1\,\beta_1 + \ldots + x_n\,\beta_n) + b\,(y_1\,\beta_1 + \ldots + y_n\,\beta_n)$$

$$= aT\,(\alpha) + bT\,(\beta) \qquad \text{[By def. of } T\text{]}$$

$\therefore$ T is a linear transformation from U into V. Thus there exists a linear transformation T from U into V such that

$$T\,(\alpha_i) = \beta_i\,,\ i = 1, 2, \ldots, n.$$

Uniqueness of T: Let $T\,'$ be a linear transformation from U into V such that

$$T\,'\,(\alpha_i) = \beta_i\,,\ i = 1, 2, \ldots, n.$$

For the vector $\alpha = x_1\alpha_1 + \ldots + x_n\,\alpha_n \in U$, we have

$$T\,'\,(\alpha) = T\,'\,(x_1\alpha_1 + \ldots + x_n\,\alpha_n)$$

$$= x_1 T\,'\,(\alpha_1) + \ldots + x_n\,T\,'\,(\alpha_n)$$

[$\because$ $T\,'$ is a linear transformation]

$$= x_1\,\beta_1 + \ldots + x_n\,\beta_n \qquad \text{[By def. of } T\,'\text{]}$$

$$= T\,(\alpha). \qquad \text{[By def. of } T\text{]}$$

Thus $T\,'\,(\alpha) = T\,(\alpha)\ \forall\ \alpha \in U.$

$\therefore$ $T\,' = T.$

This shows the uniqueness of T.

Note: From this theorem we conclude that if T is a linear transformation from a finite dimensional vector space $U\,(F)$ into a vector space $V\,(F)$, then T is completely defined if we mention under T the images of the elements of a basis set of U. If S and T are two linear transformations from U into V such that

$S\,(\alpha_i) = T\,(\alpha_i)\ \forall\ \alpha_i$ belonging to a basis of U, then

$S\,(\alpha) = T\,(\alpha)\ \forall\ \alpha \in U,$ *i.e.,* $S = T.$

Thus two linear transformations from U into V are equal if they agree on a basis of U.

Theorem 3: *Let U be an n-dimensional vector space over the field F, and let V be an m-dimensional vector space over F. Then the vector space L (U, V) of all linear transformations from U into V is finite dimensional and is of dimension mn.*

Proof: Let $B = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ and $B\,' = \{\beta_1, \beta_2, \ldots, \beta_m\}$ be ordered bases for U and V respectively. By theorem 2, there exists a unique linear transformation T_{11} from U into V such that

$$T_{11}\,(\alpha_1) = \beta_1, T_{11}\,(\alpha_2) = \mathbf{0}, \ldots, T_{11}\,(\alpha_n) = \mathbf{0}$$

where $\beta_1, \mathbf{0}, \ldots, \mathbf{0}$ are vectors in V.

In fact, for each pair of integers (p, q) with $1 \le p \le m$ and $1 \le q \le n$, there exists a unique linear transformation T_{pq} from U into V such that

$$T_{pq}\,(\alpha_i) = \begin{cases} \mathbf{0}, & \text{if} \quad i \ne q \\ \beta_p, & \text{if} \quad i = q \end{cases}$$

i.e., $T_{pq}(\alpha_i) = \delta_{iq}\,\beta_p$, ...(1)

where $\delta_{iq} \in F$ is Kronecker delta *i.e.*, $\delta_{iq} = 1$ if $i = q$ and $\delta_{iq} = 0$ if $i \neq q$.

Since p can be any of $1, 2, \ldots, m$ and q any of $1, 2, \ldots, n$, there are mn such T_{pq}'s. Let B_1 denote the set of these mn transformations T_{pq}'s. We shall show that B_1 is a basis for $L(U, V)$.

(i) First we shall show that $L(U, V)$ is a linear span of B_1.

Let $T \in L(U, V)$. Since $T(\alpha_1) \in V$ and any element in V is a linear combination of $\beta_1, \beta_2, \ldots, \beta_m$, therefore

$$T(\alpha_1) = a_{11}\,\beta_1 + a_{21}\,\beta_2 + \ldots + a_{m1}\,\beta_m,$$

for some $a_{11}, a_{21}, \ldots, a_{m1} \in F$. In fact for each $i, 1 \leq i \leq n$,

$$T(\alpha_i) = a_{1i}\,\beta_1 + a_{2i}\,\beta_2 + \ldots + a_{mi}\,\beta_m = \sum_{p=1}^{m} a_{pi}\,\beta_p \quad \ldots(2)$$

Now consider $S = \sum_{p=1}^{m} \sum_{q=1}^{n} a_{pq}\,T_{pq}$.

Obviously S is a linear combination of elements of B_1 which is a subset of $L(U, V)$. Since $L(U, V)$ is a vector space, therefore $S \in L(U, V)$ *i.e.*, S is also a linear transformation from U into V. We shall show that $S = T$.

Let us compute $S(\alpha_i)$ where α_i is any vector in the basis B of U. We have

$$S(\alpha_i) = \left[\sum_{p=1}^{m} \sum_{q=1}^{n} a_{pq}\,T_{pq}\right](\alpha_i) = \sum_{p=1}^{m} \sum_{q=1}^{n} a_{pq}\,T_{pq}(\alpha_i)$$

$$= \sum_{p=1}^{m} \sum_{q=1}^{n} a_{pq}\,\delta_{iq}\,\beta_p \qquad \text{[From (1)]}$$

$$= \sum_{p=1}^{m} a_{pi}\,\beta_p$$

[On summing with respect to q. Remember that $\delta_{iq} = 1$ when $q = i$ and $\delta_{iq} = 0$ when $q \neq i$]

$$= T(\alpha_i). \qquad \text{[From (2)]}$$

Thus $S(\alpha_i) = T(\alpha_i)$ $\forall$ $\alpha_i \in B$. Therefore S and T agree on a basis of U. So we must have $S = T$. Thus T is also a linear combination of the elements of B_1. Therefore $L(U, V)$ is a linear span of B_1.

(ii) Now we shall show that B_1 is linearly independent. For b_{pq}'s $\in F$, let

$$\sum_{p=1}^{m} \sum_{q=1}^{n} b_{pq}\,T_{pq} = \hat{\mathbf{0}} \text{ i.e., zero vector of } L(U, V)$$

$$\Rightarrow \left[\sum_{p=1}^{m} \sum_{q=1}^{n} b_{pq}\,T_{pq}\right](\alpha_i) = \hat{\mathbf{0}}(\alpha_i)\ \forall\ \alpha_i \in B$$

$$\Rightarrow \sum_{p=1}^{m} \sum_{q=1}^{n} b_{pq}\,T_{pq}(\alpha_i) = \mathbf{0} \in V \qquad [\because \hat{\mathbf{0}} \text{ is zero transformation}]$$

$$\Rightarrow \sum_{p=1}^{m} \sum_{q=1}^{n} b_{pq}\,\delta_{iq}\,\beta_p = \mathbf{0} \quad \Rightarrow \quad \sum_{p=1}^{m} b_{pi}\,\beta_p = \mathbf{0}$$

$$\Rightarrow b_{1i}\,\beta_1 + b_{2i}\,\beta_2 + \ldots + b_{mi}\,\beta_m = \mathbf{0}, 1 \leq i \leq n$$

$$\Rightarrow b_{1i} = 0, b_{2i} = 0, \ldots, b_{mi} = 0, 1 \leq i \leq n$$

[$\because$ $\beta_1, \beta_2, \ldots, \beta_m$ are linearly independent]

$\Rightarrow$ $b_{pq} = 0$ where $1 \le p \le m$ and $1 \le q \le n$

$\Rightarrow$ B_1 is linearly independent. Therefore B_1 is a basis of $L(U, V)$.

$\therefore$ dim $L(U, V)$ = number of elements in $B_1 = mn$.

Corollary: *The vector space $L(U, U)$ of all linear operators on an n-dimensional vector space U is of dimension n^2.*

Note: Suppose $U(F)$ is an n-dimensional vector space and $V(F)$ is an m-dimensional vector space. If $U \ne \{\mathbf{0}\}$ and $V \ne \{\mathbf{0}\}$, then $n \ge 1$ and $m \ge 1$. Therefore $L(U, V)$ does not just consist of the element $\hat{\mathbf{0}}$, because dimension of $L(U, V)$ is $mn \ge 1$.

7 Product of Linear Transformations

Theorem 1: *Let U, V and W be vector spaces over the field F. Let T be a linear transformation from U into V and S a linear transformation from V into W. Then the composite function ST (called product of linear transformations) defined by*

$$(ST)(\alpha) = S[T(\alpha)] \ \forall \ \alpha \in U$$

is a linear transformation from U into W. **(Kumaun 2007)**

Proof: T is a function from U into V and S is a function from V into W.

So $\alpha \in U \Rightarrow T(\alpha) \in V$.

Further $T(\alpha) \in V \Rightarrow S[T(\alpha)] \in W$.

Thus $(ST)(\alpha) \in W$.

Therefore ST is a function from U into W. Now to show that ST is a linear transformation from U into W.

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

$$(ST)(a\alpha + b\beta) = S[T(a\alpha + b\beta)] \quad \text{[By def. of product of two functions]}$$
$$= S[aT(\alpha) + bT(\beta)] \quad [\because T \text{ is a linear transformation}]$$
$$= aS[T(\alpha)] + bS[T(\beta)] \quad [\because S \text{ is a linear transformation}]$$
$$= a(ST)(\alpha) + b(ST)(\beta).$$

Hence ST is a linear transformation from U into W.

Note: If T and S are linear operators on a vector space $V(F)$, then both the products ST as well as TS exist and each is a linear operator on V. However, in general $TS \ne ST$ as is obvious from the following examples.

Theorem 2: *Let $V(F)$ be a vector space and A, B, C be linear transformations on V. Then*

(i) $A\hat{\mathbf{0}} = \hat{\mathbf{0}} = \hat{\mathbf{0}}A$ *(ii)* $AI = A = IA$

(iii) $A(BC) = (AB)C$ *(iv)* $A(B + C) = AB + AC$

(v) $(A + B)C = AC + BC$ **(Kumaun 2011)**

(vi) $c(AB) = (cA)B = A(cB)$ *where c is any element of F.* **(Kumaun 2008, 11)**

Proof: Just for the sake of convenience we first mention here our definitions of addition, scalar multiplication and product of linear transformations :

$$(A+B)(\alpha) = A(\alpha) + B(\alpha) \quad \ldots(1)$$
$$(cA)(\alpha) = cA(\alpha) \quad \ldots(2)$$
$$(AB)(\alpha) = A[B(\alpha)] \quad \ldots(3)$$

$\forall \ \alpha \in V$ and $\forall \ c \in F$.

Now we shall prove the above results.

(i) We have $\forall \ \alpha \in V$, $(A\hat{\mathbf{0}})(\alpha) = A[\hat{\mathbf{0}}(\alpha)]$ [By (3)]

$= A(\mathbf{0})$ [$\because$ $\hat{\mathbf{0}}$ is zero transformation]

$= \mathbf{0} = \hat{\mathbf{0}}(\alpha)$.

$\therefore$ $A\hat{\mathbf{0}} = \hat{\mathbf{0}}$. [By def. of equality of two functions]

Similarly we can show that $\hat{\mathbf{0}} \ A = \hat{\mathbf{0}}$.

(ii) We have $\forall \ \alpha \in V$, $(AI)(\alpha) = A[I(\alpha)]$

$= A(\alpha)$ [$\because$ I is identity transformation]

$\therefore$ $AI = A$.

Similarly we can show that $IA = A$.

(iii) We have $\forall \ \alpha \in V$, $[A(BC)](\alpha) = A[(BC)(\alpha)]$ [By (3)]

$= A[B(C(\alpha))]$ [By (3)]

$= (AB)[C(\alpha)]$ [By (3)]

$= [(AB)C](\alpha)$. [By (3)]

$\therefore$ $A(BC) = (AB)C$.

(iv) We have $\forall \ \alpha \in V$, $[A(B+C)](\alpha) = A[(B+C)(\alpha)]$ [By (3)]

$= A[B(\alpha) + C(\alpha)]$ [By (1)]

$= A[B(\alpha)] + A[C(\alpha)]$ [$\because$ A is a linear transformation and $B(\alpha), C(\alpha) \in V$]

$= (AB)(\alpha) + (AC)(\alpha)$ [By (3)]

$= (AB + AC)(\alpha)$ [By (1)]

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

(v) We have $\forall \ \alpha \in V$, $[(A+B)C](\alpha) = (A+B)[C(\alpha)]$ [By (3)]

$= A[C(\alpha)] + B[C(\alpha)]$ [By (1) since $C(\alpha) \in V$]

$= (AC)(\alpha) + (BC)(\alpha)$ [By (3)]

$= (AC + BC)(\alpha)$ [By (1)]

$\therefore$ $(A+B)C = AC + BC$.

(vi) We have $\forall \ \alpha \in V$, $[c(AB)](\alpha) = c[(AB)(\alpha)]$ [By (2)]

$= c[A(B(\alpha))]$ [By (3)]

$= (cA)[B(\alpha)]$ [By (2) since $B(\alpha) \in V$]

$= [(cA)B](\alpha)$ [By (3)]

$\therefore$ $c(AB) = (cA)B$.

Again $[c(AB)](\alpha) = c[(AB)(\alpha)]$ [By (2)]

$= c\,[A\,(B\,(\alpha))]$ [By (3)]

$= A\,[c\,B\,(\alpha)]$

[$\because$ A is a linear transformation and $B\,(\alpha) \in V$]

$= A\,[(c\,B\,)\,(\alpha)]$ [By (2)]

$= [A\,(c\,B\,)]\,(\alpha).$ [By (3)]

$\therefore \quad c\,(AB\,) = A\,(c\,B\,).$

Illustrative Examples

Example 7: *Let T_1 and T_2 be linear operators on $\mathbf{R}^2$ defined as follows :*

$$T_1\,(x_1, x_2) = (x_2\,, x_1) \quad \text{and} \quad T_2\,(x_1, x_2) = (x_1, 0).$$

Show that $T_1 T_2 \neq T_2 T_1$. **(Kumaun 2013)**

Solution: We have

$(T_1 T_2)\,(x_1, x_2) = T_1\,[T_2\,(x_1, x_2)]$

$= T_1\,(x_1, 0),$ [By def. of T_2]

$= (0, x_1).$ [By def. of T_1]

Also $(T_2\,T_1)\,(x_1, x_2) = T_2\,[T_1\,(x_1, x_2)],$ [By def. of $T_2 T_1$]

$= T_2\,(x_2\,, x_1),$ [By def. of T_1]

$= (x_2\,, 0).$ [By def. of T_2]

Thus we see that $(T_1 T_2)\,(x_1, x_2) \neq (T_2 T_1)\,(x_1, x_2)\ \forall\ (x_1, x_2) \in \mathbf{R}^2$.

Hence by the definition of equality of two mappings, we have $T_1 T_2 \neq T_2 T_1$.

Example 8: *Let $S\,(\mathbf{R})$ be the vector space of all polynomial functions in x with coefficients as elements of the field $\mathbf{R}$ of real numbers. Let D and T be two linear operators on V defined by*

$$D\,(\,f\,(x)) = \frac{d}{dx}\,f\,(x) \qquad \ldots(1)$$

and

$$T\,(\,f\,(x)) = \int_0^x f\,(x)\,dx \qquad \ldots(2)$$

for every $f\,(x) \in V$.

Then show that $DT = I$ (identity operator) and $TD \neq I$.

Solution: Let $f\,(x) = a_0 + a_1 x + a_2 x^2 + \ldots \in V$.

We have $(DT\,)\,(\,f\,(x)) = D\,[T\,(\,f\,(x))]$

$$= D\left[\int_0^x f\,(x)\,dx\right] = D\left[\int_0^x (a_0 + a_1 x + a_2 x^2 + \ldots)\,dx\right]$$

$$= D\left[a_0\,x + \frac{a_1}{2}\,x^2 + \frac{a_2}{3}\,x^3 + \ldots\right]_0^x = \frac{d}{dx}\left[a_0\,x + \frac{a_1}{2}\,x^2 + \ldots\right]$$

$$= a_0 + a_1 x + a_2 x^2 + \ldots = f\,(x) = I\,[\,f\,(x)].$$

Thus we have $(DT\,)\,[\,f\,(x)] = I\,[\,f\,(x)]\ \ \forall\ f\,(x) \in V$. Therefore $DT = I$.

Now $(TD\,)\,f\,(x) = T\,[D\,f\,(x)]$

$$= T\left[\frac{d}{dx}(a_0 + a_1 x + a_2 x^2 + \dots)\right] = T(a_1 + 2a_2 x + \dots)$$

$$= \int_0^x (a_1 + 2a_2 x + \dots)\, dx = [a_1 x + a_2 x^2 + \dots]_0^x$$

$$= a_1 x + a_2 x^2 + \dots$$

$$\neq f(x) \text{ unless } a_0 = 0.$$

Thus $\exists\, f(x) \in V$ such that $(TD)[f(x)] \neq I[f(x)]$.

$\therefore \quad TD \neq I.$

Hence $\quad TD \neq DT,$

showing that product of linear operators is not in general commutative.

Example 9: *Let V (**R**) be the vector space of all polynomials in x with coefficients in the field **R**. Let D and T be two linear transformations on V defined as*

$$D[f(x)] = \frac{d}{dx} f(x) \;\forall\; f(x) \in V \quad \textit{and} \quad T[f(x)] = x\, f(x) \;\forall\; f(x) \in V.$$

Then show that $DT \neq TD$.

Solution: We have

$$(DT)[f(x)] = D[T(f(x))]$$

$$= D[x\, f(x)] = \frac{d}{dx}[x\, f(x)]$$

$$= f(x) + x \frac{d}{dx} f(x). \qquad \dots(1)$$

Also

$$(TD)[f(x)] = T[D(f(x))] = T\left[\frac{d}{dx}(f(x))\right]$$

$$= x \frac{d}{dx} f(x). \qquad \dots(2)$$

From (1) and (2), we see that $\exists\, f(x) \in V$ such that

$$(DT)(f(x)) \neq (TD)(f(x)) \quad \Rightarrow \quad DT \neq TD.$$

Also we see that

$$(DT - TD)(f(x)) = (DT)(f(x)) - (TD)(f(x))$$

$$= f(x) = I(f(x)).$$

$\therefore \quad DT - TD = I.$

8 Ring of Linear Operators on a Vector Space

Ring: Definition: *A non-empty set R with two binary operations, to be denoted additively and multiplicatively, is called a ring if the following postulates are satisfied :*

R_1. *R is closed with respect to addition i.e.,*
$a + b \in R \;\forall\; a, b \in R.$

R_2. $(a + b) + c = a + (b + c) \;\forall\; a, b, c \in R.$

R_3. $a + b = b + a \;\forall\; a, b \in R.$

R_4. *$\exists$ an element 0 (called zero element) in R such that*
$0 + a = a \;\forall\; a \in R.$

R_5. $a \in R \ \Rightarrow \ \exists - a \in R$ *such that*
$(- a) + a = 0$.
R_6. *R is closed with respect to multiplication i.e.,*
$ab \in R, \forall \ a, b \in R$
R_7. $(ab)\, c = a\,(bc) \ \forall \ a, b, c \in R$.
R_8. *Multiplication is distributive with respect to addition, i.e.,*
$a\,(b + c) = ab + ac$ *and* $(a + b)\, c = ac + bc \ \forall \ a, b, c \in R$.

Ring with unity element: Definition:

If in a ring R there exists an element $1 \in R$ *such that*

$$1a = a = a1 \ \forall \ a \in R,$$

then R is called a ring with unity element. The element 1 is called the unity element of the ring.

Theorem: *The set* $L\,(V, V)$ *of all linear transformations from a vector space* $V\,(F)$ *into itself is a ring with unity element with respect to addition and multiplication of linear transformations defined as below :*

$$(S + T)\,(\alpha) = S\,(\alpha) + T\,(\alpha)$$

and $$(ST)\,(\alpha) = S\,[T\,(\alpha)] \ \forall \ S, T \in L\,(V, V) \text{ and } \forall \, \alpha \in V.$$

Proof: The students should themselves write the complete proof of this theorem. We have proved all the steps here and there. They should show here that all the ring postulates are satisfied in the set $L\,(V, V)$. The transformation $\hat{\mathbf{0}}$ will act as the zero element and the identity transformation I will act as the unity element of this ring.

9 Algebra or Linear Algebra

(Garhwal 2006, 12)

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

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

If there is an element 1 in V such that

$$1\alpha = \alpha = \alpha 1 \ \forall \ \alpha \in V,$$

then we call V a **linear algebra with identity over** F. Also 1 is then called the identity of V. The algebra V is **Commutative** if

$$\alpha\beta = \beta\alpha \ \forall \ \alpha, \beta \in V.$$

Theorem: *Let* $V\,(F)$ *be a vector space. The vector space* $L\,(V, V)$ *over F of all linear transformations from V into V is a linear algebra with identity with respect to the product of linear transformations as the multiplication composition in* $L\,(V, V)$.

Proof: The students should write the complete proof here. All the necessary steps have been proved here and there.

10 Polynomials

Let T be a linear transformation on a vector space $V(F)$. Then TT is also a linear transformation on V. We shall write $T^1 = T$ and $T^2 = TT$. Since the product of linear transformations is an associative operation, therefore if m is a positive integer, we shall define

$$T^m = TTT \ldots \text{upto } m \text{ times.}$$

Obviously T^m is a linear transformation on V.

Also we define $T^0 = I$ (identity transformation).

If m and n are non-negative integers, it can be easily seen that

$$T^m T^n = T^{m+n} \text{ and } (T^m)^n = T^{mn}.$$

The set $L(V, V)$ of all linear transformations on V is a vector space over the field F. If $a_0, a_1, \ldots, a_n \in F$, then

$$p(T) = a_0 I + a_1 T + a_2 T^2 + \ldots + a_n T^n \in L(V, V)$$

i.e., $p(T)$ is also a linear transformation on V because it is a linear combination over F of elements of $L(V, V)$. We call $p(T)$ as a polynomial in linear transformation T. The polynomials in a linear transformation behave like ordinary polynomials.

11 Invertible Linear Transformations

(Garhwal 2006)

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

If T is a function from U into V, then T is said to be 1-1 if

$$\alpha_1, \alpha_2 \in U \text{ and } \alpha_1 \neq \alpha_2 \;\Rightarrow\; T(\alpha_1) \neq T(\alpha_2).$$

In other words T is said to be 1-1 if

$$\alpha_1, \alpha_2 \in U \text{ and } T(\alpha_1) = T(\alpha_2) \;\Rightarrow\; \alpha_1 = \alpha_2.$$

Further T is said to be onto if

$$\beta \in V \;\Rightarrow\; \exists\, \alpha \in U \text{ such that } T(\alpha) = \beta.$$

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

Let β be any vector in V. Since T is onto, therefore

$$\beta \in V \;\Rightarrow\; \exists\, \alpha \in U \text{ such that } T(\alpha) = \beta.$$

Also α determined in this way is a unique element of U because T is one-one and therefore

$$\alpha_0, \alpha \in U \text{ and } \alpha_0 \neq \alpha \;\Rightarrow\; \beta = T(\alpha) \neq T(\alpha_0).$$

We define $T^{-1}(\beta)$ to be α. Thus

$$T^{-1} : V \Rightarrow U \text{ such that}$$

$$T^{-1}(\beta) = \alpha \;\Leftrightarrow\; T(\alpha) = \beta.$$

The function T^{-1} is itself one-one and onto. In the following theorem, we shall prove that T^{-1} is a linear transformation from V into U.

Theorem 1: *Let U and V be vector spaces over the field F and let T be a linear transformation from U into V. If T is one-one and onto, then the inverse function T^{-1} is a linear transformation from V into U.*

Proof: Let $\beta_1, \beta_2 \in V$ and $a, b \in F$.

Since T is one-one and onto, therefore there exist unique vectors $\alpha_1, \alpha_2 \in U$ such that $T(\alpha_1) = \beta_1, T(\alpha_2) = \beta_2$.

By definition of T^{-1}, we have

$$T^{-1}(\beta_1) = \alpha_1, T^{-1}(\beta_2) = \alpha_2.$$

Now $a\alpha_1 + b\alpha_2 \in U$ and we have by linearity of T,

$$T(a\alpha_1 + b\alpha_2) = aT(\alpha_1) + bT(\alpha_2)$$
$$= a\beta_1 + b\beta_2 \in V.$$

$\therefore$ by def. of T^{-1}, we have

$$T^{-1}(a\beta_1 + b\beta_2) = a\alpha_1 + b\alpha_2$$
$$= aT^{-1}(\beta_1) + bT^{-1}(\beta_2).$$

$\therefore$ T^{-1} is a linear transformation from V into U.

Theorem 2: *Let T be an invertible linear transformation on a vector space $V(F)$. Then*

$$T^{-1}T = I = TT^{-1}.$$

(Kumaun 2013)

Proof: Let α be any element of V and let $T(\alpha) = \beta$. Then

$$T^{-1}(\beta) = \alpha.$$

We have $T(\alpha) = \beta \quad \Rightarrow \quad T^{-1}[T(\alpha)] = T^{-1}(\beta)$

$\Rightarrow \quad (T^{-1}T)(\alpha) = \alpha \quad \Rightarrow \quad (T^{-1}T)(\alpha) = I(\alpha)$

$\Rightarrow \quad T^{-1}T = I.$

Let β be any element of V. Since T is onto, therefore $\beta \in V \Rightarrow \exists\, \alpha \in V$ such that $T(\alpha) = \beta$. Then $T^{-1}(\beta) = \alpha$.

Now $T^{-1}(\beta) = \alpha \quad \Rightarrow \quad T[T^{-1}(\beta)] = T(\alpha)$

$\Rightarrow \quad (TT^{-1})(\beta) = \beta \quad \Rightarrow \quad (TT^{-1})(\beta) = \beta = I(\beta)$

$\Rightarrow \quad TT^{-1} = I.$

Theorem 3: *If A, B and C are linear transformations on a vector space $V(F)$ such that*

$$AB = CA = I,$$

then A is invertible and $A^{-1} = B = C$.

(Kumaun 2011)

Proof: In order to show that A is invertible, we are to show that A is one-one and onto.

(i) A is one-one:

Let $\alpha_1, \alpha_2 \in V$. Then

$$A(\alpha_1) = A(\alpha_2)$$

$\Rightarrow \quad C[A(\alpha_1)] = C[A(\alpha_2)] \quad \Rightarrow \quad (CA)(\alpha_1) = (CA)(\alpha_2)$

$\Rightarrow \quad I(\alpha_1) = I(\alpha_2) \quad \Rightarrow \quad \alpha_1 = \alpha_2.$

$\therefore$ A is **one-one.**

(ii) A is onto:

Let β be any element of V. Since B is a linear transformation on V, therefore $B(\beta) \in V$. Let $B(\beta) = \alpha$. Then

$B(\beta) = \alpha \quad \Rightarrow \quad A[B(\beta)] = A(\alpha)$

$\Rightarrow \quad (AB)(\beta) = A(\alpha) \quad \Rightarrow \quad I(\beta) = A(\alpha) [\because AB = I]$

$\Rightarrow \quad \beta = A(\alpha).$

Thus $\beta \in V \Rightarrow \exists\, \alpha \in V$ such that $A(\alpha) = \beta$.

$\therefore$ A is onto.

Since A is one-one and onto therefore A is invertible *i.e.*, A^{-1} exists.

(iii) Now we shall show that $A^{-1} = B = C$.

We have $\quad AB = I \quad \Rightarrow \quad A^{-1}(AB) = A^{-1} I$

$\Rightarrow \quad (A^{-1} A) B = A^{-1} \quad \Rightarrow \quad IB = A^{-1}$

$\Rightarrow \quad B = A^{-1}.$

Again $\quad CA = I \quad \Rightarrow \quad (CA) A^{-1} = IA^{-1}$

$\Rightarrow \quad C(AA^{-1}) = A^{-1} \quad \Rightarrow \quad CI = A^{-1}$

$\Rightarrow \quad C = A^{-1}.$

Hence the theorem.

Theorem 4: *The necessary and sufficient condition for a linear transformation A on a vector space $V(F)$ to be invertible is that there exists a linear transformation B on V such that*

$$AB = I = BA.$$

Proof: The condition is necessary. For proof see theorem 2.

The condition is sufficient: For proof see theorem 3. Take B in place of C.

Also we note that $B = A^{-1}$ and $A = B^{-1}$.

Theorem 5: Uniqueness of inverse: *Let A be an invertible linear transformation on a vector space $V(F)$. Then A possesses unique inverse.*

Proof: Let B and C be two inverses of A. Then

$AB = I = BA$

and $\quad AC = I = CA.$

We have $\quad C(AB) = CI = C. \quad \ldots(1)$

Also $\quad (CA) B = IB = B. \quad \ldots(2)$

Since product of linear transformations is associative, therefore from (1) and (2), we get

$C(AB) = (CA) B$

$\Rightarrow \quad C = B.$

Hence the inverse of A is unique.

Theorem 6: *Let $V(F)$ be a vector space and let A, B be linear transformations on V. Then show that*

(i) If A and B are invertible, then AB is invertible and

$$(AB)^{-1} = B^{-1} A^{-1}.$$

(Kumaun 2007)

(ii) If A is invertible and $a \neq 0 \in F$, then aA is invertible and

$$(aA)^{-1} = \frac{1}{a} A^{-1}.$$

(iii) If A is invertible, then A^{-1} is invertible and $(A^{-1})^{-1} = A$.

Proof: (i) We have

$$(B^{-1} A^{-1})(AB) = B^{-1}[A^{-1}(AB)] = B^{-1}[(A^{-1} A) B]$$
$$= B^{-1}(IB) = B^{-1} B = I.$$

Also $$(AB)(B^{-1} A^{-1}) = A[B(B^{-1} A^{-1})] = A[(BB^{-1}) A^{-1}]$$
$$= A(IA^{-1}) = AA^{-1} = I.$$

Thus $$(AB)(B^{-1} A^{-1}) = I = (B^{-1} A^{-1})(AB).$$

$\therefore$ By theorem 3, AB is invertible and $(AB)^{-1} = B^{-1} A^{-1}$.

(ii) We have $$(aA)\left(\frac{1}{a} A^{-1}\right) = a\left[A\left(\frac{1}{a} A^{-1}\right)\right] = a\left[\frac{1}{a}(AA^{-1})\right]$$
$$= \left(a \frac{1}{a}\right)(AA^{-1}) = 1I = I.$$

Also $$\left(\frac{1}{a} A^{-1}\right)(aA) = \frac{1}{a}[A^{-1}(aA)] = \frac{1}{a}[a(A^{-1} A)]$$
$$= \left(\frac{1}{a} a\right)(A^{-1} A) = 1I = I.$$

Thus $$(aA)\left(\frac{1}{a} A^{-1}\right) = I = \left(\frac{1}{a} A^{-1}\right)(aA).$$

$\therefore$ by theorem 3, aA is invertible and $(aA)^{-1} = \frac{1}{a} A^{-1}$.

(iii) Since A is invertible, therefore

$$AA^{-1} = I = A^{-1} A.$$

$\therefore$ By theorem 3, A^{-1} is invertible and $A = (A^{-1})^{-1}$.

12 Singular and Non-singular Transformations

(Garhwal 2006, 07; Kumaun 07, 09)

Definition: *Let T be a linear transformation from a vector space $U(F)$ into a vector space $V(F)$. Then T is said to be* **non-singular** *if the null space of T (i.e., $\ker T$) consists of the zero vector alone i.e., if*

$$\alpha \in U \text{ and } T(\alpha) = \mathbf{0} \Rightarrow \alpha = \mathbf{0}.$$

If there exists a vector $\mathbf{0} \neq \alpha \in U$ such that $T(\alpha) = \mathbf{0}$, then T is said to be singular.

Theorem 1: *Let T be a linear transformation from a vector space $U(F)$ into a vector space $V(F)$. Then T is non-singular if and only if T is one-one.* **(Kumaun 2007, 09)**

Proof: Given that T is non-singular. Then to prove that T is one-one.

Let $\alpha_1, \alpha_2 \in U$. Then

$$T(\alpha_1) = T(\alpha_2)$$

$$\Rightarrow \quad T(\alpha_1) - T(\alpha_2) = \mathbf{0}$$

$$\Rightarrow \quad T(\alpha_1 - \alpha_2) = \mathbf{0}$$

$$\Rightarrow \quad \alpha_1 - \alpha_2 = \mathbf{0} \qquad [\because \ T \text{ is non-singular}]$$

$$\Rightarrow \quad \alpha_1 = \alpha_2.$$

$\therefore$ T is one-one.

Conversely let T be one-one. We know that $T(\mathbf{0}) = \mathbf{0}$. Since T is one-one, therefore

$$\alpha \in U \text{ and } T(\alpha) = \mathbf{0} = T(\mathbf{0}) \Rightarrow \alpha = \mathbf{0}.$$

Thus the null space of T consists of zero vector alone. Therefore T is non-singular.

Theorem 2: *Let T be a linear transformation from U into V. Then T is non-singular if and only if T carries each linearly independent subset of U onto a linearly independent subset of V.*

Proof: First suppose that T is non-singular.

Let $\quad B = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$

be a linearly independent subset of U. Then image of B under T is the subset B' of V given by

$$B' = \{T(\alpha_1), T(\alpha_2), \ldots, T(\alpha_n)\}.$$

To prove that B' is linearly independent.

Let $\quad a_1, a_2, \ldots, a_n \in F$

and let $\quad a_1 T(\alpha_1) + \ldots + a_n T(\alpha_n) = \mathbf{0}$

$$\Rightarrow \quad T(a_1 \alpha_1 + \ldots + a_n \alpha_n) = \mathbf{0} \qquad [\because \ T \text{ is linear}]$$

$$\Rightarrow \quad a_1 \alpha_1 + \ldots + a_n \alpha_n = \mathbf{0} \qquad [\because \ T \text{ is non-singular}]$$

$$\Rightarrow \quad a_i = 0, i = 1, 2, \ldots, n \qquad [\because \ \alpha_1, \ldots, \alpha_n \text{ are linearly independent}]$$

Thus the image of B under T is linearly independent.

Conversely suppose that T carries independent subsets onto independent subsets. Then to prove that T is non-singular.

Let $\alpha \neq \mathbf{0} \in U$. Then the set $S = \{\alpha\}$ consisting of the one non-zero vector α is linearly independent. The image of S under T is the set

$$S' = \{T(\alpha)\}.$$

It is given that S' is also linearly independent. Therefore $T(\alpha) \neq \mathbf{0}$ because the set consisting of zero vector alone is linearly dependent. Thus

$$\mathbf{0} \neq \alpha \in U \Rightarrow T(\alpha) \neq \mathbf{0}.$$

This shows that the null space of T consists of the zero vector alone.

Therefore T is non-singular.

Theorem 3: *Let U and V be finite dimensional vector spaces over the field F such that dim U = dim V. If T is a linear transformation from U into V, the following are equivalent.*

(i) *T is invertible.*

(ii) *T is non-singular.*

(iii) *The range of T is V.*

(iv) *If $\{\alpha_1, \ldots, \alpha_n\}$ is any basis for U, then $\{T(\alpha_1), \ldots, T(\alpha_n)\}$ is a basis for V.*

(v) *There is some basis $\{\alpha_1, \ldots, \alpha_n\}$ for U such that $\{T(\alpha_1), \ldots, T(\alpha_n)\}$ is a basis for V.*

Proof: (i) $\Rightarrow$ (ii).

If T is invertible, then T is one-one. Therefore T is non-singular.

(ii) $\Rightarrow$ (iii).

Let T be non-singular. Let $\{\alpha_1, \ldots, \alpha_n\}$ be a basis for U. Then $\{\alpha_1, \ldots, \alpha_n\}$ is a linearly independent subset of U. Since T is non-singular therefore $\{T(\alpha_1), \ldots, T(\alpha_n)\}$ is a linearly independent subset of V and it contains n vectors. Since dim V is also n, therefore this set of vectors is a basis for V. Now let β be any vector in V. Then there exist scalars $a_1, \ldots, a_n \in F$ such that

$$\beta = a_1 T(\alpha_1) + \ldots + a_n T(\alpha_n) = T(a_1 \alpha_1 + \ldots + a_n \alpha_n)$$

which shows that β is in the range of T because

$$a_1 \alpha_1 + \ldots + a_n \alpha_n \in U.$$

Thus every vector in V is in the range of T.

Hence range of T is V.

(iii) $\Rightarrow$ (iv).

Now suppose that range of T is V *i.e.,* T is onto. If $\{\alpha_1, \ldots, \alpha_n\}$ is any basis for U, then the vectors $T(\alpha_1), \ldots, T(\alpha_n)$ span the range of T which is equal to V. Thus the vectors $T(\alpha_1), \ldots, T(\alpha_n)$ which are n in number span V whose dimension is also n.

Therefore $\{T(\alpha_1), \ldots, T(\alpha_n)\}$ must be a basis set for V.

(iv) $\Rightarrow$ (v).

Since U is finite dimensional, therefore there exists a basis for U. Let $\{\alpha_1, \ldots, \alpha_n\}$ be a basis for U. Then $\{T(\alpha_1), \ldots, T(\alpha_n)\}$ is a basis for V as it is given in (iv).

(v) $\Rightarrow$ (i).

Suppose there is some basis $\{\alpha_1, \ldots, \alpha_n\}$ for U such that $\{T(\alpha_1), \ldots, T(\alpha_n)\}$ is a basis for V. The vectors $\{T(\alpha_1), \ldots, T(\alpha_n)\}$ span the range of T. Also they span V. Therefore the range of T must be all of V *i.e.,* T is onto.

If $\alpha = c_1 \alpha_1 + \ldots + c_n \alpha_n$ is in the null space of T, then

$$T(c_1 \alpha_1 + \ldots + c_n \alpha_n) = \mathbf{0}$$

$\Rightarrow$ $c_1 T(\alpha_1) + \ldots + c_n T(\alpha_n) = \mathbf{0}$

$\Rightarrow$ $c_i = 0, 1 \le i \le n$ because $T(\alpha_1), \ldots, T(\alpha_n)$ are linearly independent

$\Rightarrow$ $\alpha = \mathbf{0}$.

$\therefore$ T is non-singular and consequently T is one-one. Hence T is invertible.

Illustrative Examples

Example 10: *Describe explicitly the linear transformation $T : \mathbf{R}^2 \to \mathbf{R}^2$ such that*

$$T(2,3) = (4,5) \text{ and } T(1,0) = (0,0).$$

Solution: First we shall show that the set $\{(2,3),(1,0)\}$ is a basis of $\mathbf{R}^2$. For linear independence of this set let

$$a(2,3) + b(1,0) = (0,0), \text{ where } a, b \in \mathbf{R}.$$

Then $(2a + b, 3a) = (0,0)$

$\Rightarrow$ $2a + b = 0, 3a = 0$

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

Hence the set $\{(2,3),(1,0)\}$ is linearly independent.

Now we shall show that the set $\{(2,3),(1,0)\}$ spans $\mathbf{R}^2$. Let $(x_1, x_2) \in \mathbf{R}^2$ and let $(x_1, x_2) = a(2,3) + b(1,0) = (2a + b, 3a)$.

Then $2a + b = x_1, 3a = x_2$. Therefore

$$a = \frac{x_2}{3}, b = \frac{3x_1 - 2x_2}{3}.$$

$$\therefore \quad (x_1, x_2) = \frac{x_2}{3}(2,3) + \frac{3x_1 - 2x_2}{3}(1,0). \qquad \ldots(1)$$

From the relation (1) we see that the set $\{(2,3),(1,0)\}$ spans $\mathbf{R}^2$. Hence this set is a basis for $\mathbf{R}^2$.

Now let (x_1, x_2) be any member of $\mathbf{R}^2$. Then we are to find a formula for $T(x_1, x_2)$ under the conditions that $T(2,3) = (4,5), T(1,0) = (0,0)$.

We have $T(x_1, x_2) = T\left[\frac{x_2}{3}(2,3) + \frac{3x_1 - 2x_2}{3}(1,0)\right]$, by (1)

$$= \frac{x_2}{3} T(2,3) + \frac{3x_1 - 2x_2}{3} T(1,0), \text{ by linearity of } T$$

$$= \frac{x_2}{3}(4,5) + \frac{3x_1 - 2x_2}{3}(0,0)$$

$$= \left(\frac{4x_2}{3}, \frac{5x_2}{3}\right).$$

Example 11: *Describe explicitly a linear transformation from $V_3(\mathbf{R})$ into $V_3(\mathbf{R})$ which has its range the subspace spanned by $(1,0,-1)$ and $(1,2,2)$.*

Solution: The set $B = \{(1,0,0),(0,1,0),(0,0,1)\}$ is a basis for $V_3(\mathbf{R})$.

Also $\{(1,0,-1),(1,2,2),(0,0,0)\}$ is a subset of $V_3(\mathbf{R})$. It should be noted that in this subset the number of vectors has been taken the same as is the number of vectors in the set B.

There exists a unique linear transformation T from $V_3(\mathbf{R})$ into $V_3(\mathbf{R})$ such that

$$\text{and} \qquad \left.\begin{aligned} T(1,0,0) &= (1,0,-1), \\ T(0,1,0) &= (1,2,2), \\ T(0,0,1) &= (0,0,0). \end{aligned}\right\} \qquad \ldots(1)$$

Now the vectors $T(1,0,0), T(0,1,0), T(0,0,1)$ span the range of T. In other words the vectors

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

span the range of T. Thus the range of T is the subspace of $V_3(\mathbf{R})$ spanned by the set $\{(1,0,-1), (1,2,2)\}$ because the zero vector $(0,0,0)$ can be omitted from the spanning set. Therefore T defined in (1) is the required transformation.

Now let us find an explicit expression for T. Let (a, b, c) be any element of $V_3(\mathbf{R})$. Then we can write

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

$$\begin{aligned} \therefore \qquad T(a,b,c) &= aT(1,0,0) + bT(0,1,0) + cT(0,0,1) \\ &= a(1,0,-1) + b(1,2,2) + c(0,0,0) \qquad \text{[From (1)]} \\ &= (a+b, 2b, 2b-a). \end{aligned}$$

Example 12: *Let T be a linear operator on $V_3(\mathbf{R})$ defined by*

$$T(a,b,c) = (3a, a-b, 2a+b+c) \ \forall \ (a,b,c) \in V_3(\mathbf{R}).$$

Is T invertible ? If so, find a rule for T^{-1} like the one which defines T.

Solution: Let us see that T is one-one or not.

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

$$\begin{aligned} \text{Then} \qquad & T(\alpha) = T(\beta) \\ \Rightarrow \qquad & T(a_1, b_1, c_1) = T(a_2, b_2, c_2) \\ \Rightarrow \qquad & (3a_1, a_1 - b_1, 2a_1 + b_1 + c_1) = (3a_2, a_2 - b_2, 2a_2 + b_2 + c_2) \\ \Rightarrow \qquad & 3a_1 = 3a_2, a_1 - b_1 = a_2 - b_2, 2a_1 + b_1 + c_1 = 2a_2 + b_2 + c_2 \\ \Rightarrow \qquad & a_1 = a_2, b_1 = b_2, c_1 = c_2. \end{aligned}$$

$\therefore$ T is one-one.

Now T is a linear transformation on a finite dimensional vector space $V_3(\mathbf{R})$ whose dimension is 3. Since T is one-one, therefore T must be onto also and thus T is invertible.

If $T(a,b,c) = (p,q,r)$, then $T^{-1}(p,q,r) = (a,b,c)$.

$$\begin{aligned} \text{Now} \qquad & T(a,b,c) = (p,q,r) \\ \Rightarrow \qquad & (3a, a-b, 2a+b+c) = (p,q,r) \\ \Rightarrow \qquad & p = 3a, q = a-b, r = 2a+b+c \\ \Rightarrow \qquad & a = \frac{p}{3}, b = \frac{p}{3} - q, c = r - 2a - b = r - \frac{2p}{3} - \frac{p}{3} + q = r - p + q. \end{aligned}$$

$$\therefore \qquad T^{-1}(p,q,r) = \left(\frac{p}{3}, \frac{p}{3} - q, r - p + q\right) \ \forall \ (p,q,r) \in V_3(\mathbf{R})$$

is the rule which defines T^{-1}.

Example 13: *A linear transformation T is defined on* V_2 (**C**) *by*

$$T(a, b) = (\alpha a + \beta b, \gamma a + \delta b),$$

where $\alpha, \beta, \gamma, \delta$ *are fixed elements of* **C**. *Prove that T is invertible if and only if* $\alpha\delta - \beta\gamma \neq 0$.

Solution: The vector space V_2 (**C**) is of dimension 2. Therefore T is a linear transformation on a finite-dimensional vector space. T will be invertible if and only if the null space of T consists of zero vector alone. The zero vector of the space V_2 (**C**) is the ordered pair $(0, 0)$. Thus T is invertible

iff $T(x, y) = (0, 0) \Rightarrow x = 0, y = 0$

i.e., iff $(\alpha x + \beta y, \gamma x + \delta y) = (0, 0) \Rightarrow x = 0, y = 0$

i.e., iff $\alpha x + \beta y = 0, \gamma x + \delta y = 0 \Rightarrow x = 0, y = 0$.

Now the necessary and sufficient condition for the equations $\alpha x + \beta y = 0, \gamma x + \delta y = 0$ to have the only solution $x = 0, y = 0$ is that

$$\begin{vmatrix} \alpha & \beta \\ \gamma & \delta \end{vmatrix} \neq 0 .$$

Hence T is invertible iff $\alpha\delta - \beta\gamma \neq 0$.

Example 14: *Find two linear operators T and S on* V_2 (**R**) *such that*

$$TS = \hat{\mathbf{0}} \text{ but } ST \neq \hat{\mathbf{0}}.$$

(Kumaun 2009, 14)

Solution: Consider the linear transformations T and S on V_2 (**R**) defined by

$T(a, b) = (a, 0) \ \forall \ (a, b) \in V_2(\mathbf{R})$

and $S(a, b) = (0, a) \ \forall \ (a, b) \in V_2(\mathbf{R})$.

We have

$$(TS)(a, b) = T[S(a, b)]$$
$$= T(0, a) = (0, 0)$$
$$= \hat{\mathbf{0}}(a, b) \ \forall \ (a, b) \in V_2(\mathbf{R}).$$

$\therefore$ $TS = \hat{\mathbf{0}}$.

Again

$$(ST)(a, b) = S[T(a, b)]$$
$$= S(a, 0) = (0, a)$$
$$\neq \hat{\mathbf{0}}(a, b) \ \forall \ (a, b) \in V_2(\mathbf{R}).$$

Thus $ST \neq \hat{\mathbf{0}}$.

Example 15: *Let V be a vector space over the field F and T a linear operator on V. If* $T^2 = \hat{\mathbf{0}}$, *what can you say about the relation of the range of T to the null space of T? Give an example of a linear operator T on* V_2 (**R**) *such that* $T^2 = \hat{\mathbf{0}}$ *but* $T \neq \hat{\mathbf{0}}$.

Solution: We have $T^2 = \hat{\mathbf{0}}$

$\Rightarrow$ $T^2(\alpha) = \hat{\mathbf{0}}(\alpha) \ \forall \ \alpha \in V$

$\Rightarrow$ $T[T(\alpha)] = \mathbf{0} \ \forall \ \alpha \in V$

$\Rightarrow \qquad T(\alpha) \in$ null space of $T \ \forall\ \alpha \in V$.

But $T(\alpha) \in$ range of $T \ \forall\ \alpha \in V$.

$\therefore \qquad T^2 = \hat{\mathbf{0}} \Rightarrow$ range of $T \subseteq$ null space of T.

For the second part of the question, consider the linear transformation T on $V_2(\mathbf{R})$ defined by

$$T(a, b) = (0, a) \ \forall\ (a, b) \in V_2(\mathbf{R}).$$

Then obviously $T \neq \hat{\mathbf{0}}$.

We have $\qquad T^2(a, b) = T[T(a, b)] = T(0, a) = (0, 0)$

$$= \hat{\mathbf{0}}(a, b) \ \forall\ (a, b) \in V_2(\mathbf{R}).$$

$\therefore \qquad T^2 = \hat{\mathbf{0}}.$

Example 16: *If $T : U \to V$ is a linear transformation and U is finite dimensional, show that U and range of T have the same dimension iff T is non-singular. Determine all non-singular linear transformations*

$$T : V_4(\mathbf{R}) \to V_3(\mathbf{R}).$$

Solution: We know that

$$\dim U = \text{rank}(T) + \text{nullity}(T)$$
$$= \text{dim of range of } T + \text{dim of null space of } T.$$

$\therefore \qquad \dim U =$ dim of range of T

iff dim of null space of T is zero

i.e., iff null space of T consists of zero vector alone

i.e., iff T is non-singular.

Let T be a linear transformation from $V_4(\mathbf{R})$ into $V_3(\mathbf{R})$. Then T will be non-singular iff

$$\text{dim of } V_4(\mathbf{R}) = \text{dim of range of } T.$$

Now $\dim V_4(\mathbf{R}) = 4$ and dim of range of $T \leq 3$ because range of $T \subseteq V_3(\mathbf{R})$.

$\therefore$ $\dim V_4(\mathbf{R})$ cannot be equal to dim of range of T.

Hence T cannot be non-singular. Thus there can be no non-singular linear transformation from $V_4(\mathbf{R})$ into $V_3(\mathbf{R})$.

Example 17: *If A and B are linear transformations (on the same vector space), then a necessary and sufficient condition that both A and B be invertible is that both AB and BA be invertible.*

Solution: Let A and B be two invertible linear transformations on a vector space V.

We have $\qquad (AB)(B^{-1}A^{-1}) = I = (B^{-1}A^{-1})(AB).$

$\therefore$ AB is invertible.

Also we have

$$(BA)(A^{-1}B^{-1}) = I = (A^{-1}B^{-1})(BA).$$

$\therefore$ BA is also invertible.

Thus the condition is necessary.

Conversely, let AB and BA be both invertible. Then AB and BA are both one-one and onto.

First we shall show that A is invertible.

A is one-one:S

Let $\alpha_1, \alpha_2 \in V$. Then

$$A(\alpha_1) = A(\alpha_2)$$

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

$$\Rightarrow \quad (BA)(\alpha_1) = (BA)(\alpha_2)$$

$$\Rightarrow \quad \alpha_1 = \alpha_2 . \qquad [\because \; BA \text{ is one-one}]$$

$\therefore$ A is one-one.

A is onto:

Let $\beta \in V$. Since AB is onto, therefore there exists $\alpha \in V$ such that

$$(AB)(\alpha) = \beta$$

$$\Rightarrow \quad A[B(\alpha)] = \beta.$$

Thus $\beta \in V \Rightarrow \exists\, B(\alpha) \in V$ such that $A[B(\alpha)] = \beta$.

$\therefore$ A is onto.

$\therefore$ A is invertible.

Interchanging the roles played by AB and BA in the above proof, we can prove that B is invertible.

Example 18: *If A is a linear transformation on a vector space V such that*

$$A^2 - A + I = \hat{\mathbf{0}},$$

then A is invertible.

Solution: If $A^2 - A + I = \hat{\mathbf{0}}$, then

$$A^2 - A = -I.$$

First we shall prove that A is one-one Let $\alpha_1, \alpha_2 \in V$.

Then $\quad A(\alpha_1) = A(\alpha_2) \qquad \ldots(1)$

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

$$\Rightarrow \quad A^2(\alpha_1) = A^2(\alpha_2) \qquad \ldots(2)$$

$$\Rightarrow \quad A^2(\alpha_1) - A(\alpha_1) = A^2(\alpha_2) - A(\alpha_2) \qquad [\text{From (2) and (1)}]$$

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

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

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

$$\Rightarrow \quad -\alpha_1 = -\alpha_2 \;\Rightarrow\; \alpha_1 = \alpha_2.$$

$\therefore$ A is one-one.

Now to prove that A is onto.

Let $\alpha \in V$. Then

$$\alpha - A(\alpha) \in V.$$

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

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

$$= I(\alpha) \qquad [\because A^2 - A = -I \Rightarrow A - A^2 = I]$$

$$= \alpha .$$

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

$$A[\alpha - A(\alpha)] = \alpha.$$

$\therefore$ A is onto.

Hence A is invertible.

Example 19: *Let V be a finite dimensional vector space and T be a linear operator on V. Suppose that rank (T^2) = rank (T). Prove that the range and null space of T are disjoint i.e., have only the zero vector in common.*

Solution: We have

$$\dim V = \text{rank}(T) + \text{nullity}(T)$$

and $\dim V = \text{rank}(T^2) + \text{nullity}(T^2)$.

Since rank (T) = rank (T^2), therefore we get

$$\text{nullity}(T) = \text{nullity}(T^2)$$

i.e., dim of null space of T = dim of null space of T^2.

Now $T(\alpha) = \mathbf{0}$

$\Rightarrow$ $T[T(\alpha)] = T(\mathbf{0})$

$\Rightarrow$ $T^2(\alpha) = \mathbf{0}$.

$\therefore$ $\alpha \in$ null space of T $\Rightarrow$ $\alpha \in$ null space of T^2.

$\therefore$ null space of $T \subseteq$ null space of T^2.

But null space of T and null space of T^2 are both subspaces of V and have the same dimension.

$\therefore$ null space of T = null space of T^2.

$\therefore$ null space of $T^2 \subseteq$ null space of T

i.e., $T^2(\alpha) = \mathbf{0} \Rightarrow T(\alpha) = \mathbf{0}$.

$\therefore$ range and null space of T are disjoint. [See Example 1 after article 5]

Example 20: *Let V be a finite dimensional vector space over the field F. Let $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ and $\{\beta_1, \beta_2, \ldots, \beta_n\}$ be two ordered bases for V. Show that there exists a unique invertible linear transformation T on V such that*

$$T(\alpha_i) = \beta_i, 1 \le i \le n.$$

(Kumaun 2015)

Solution: We have proved in one of the previous theorems that there exists a unique linear transformation T on V such that $T(\alpha_i) = \beta_i, 1 \le i \le n$.

Here we are to show that T is invertible. Since V is finite dimensional therefore in order to prove that T is invertible, it is sufficient to prove that T is non-singular.

Let $\alpha \in V$

and $T(\alpha) = \mathbf{0}$.

Let $\alpha = a_1\alpha_1 + \ldots + a_n\alpha_n$ where $a_1, \ldots, a_n \in F$.

We have $T(\alpha) = \mathbf{0}$

$\Rightarrow$ $T(a_1\alpha_1 + \ldots + a_n\alpha_n) = \mathbf{0}$

$\Rightarrow$ $a_1 T(\alpha_1) + \ldots + a_n T(\alpha_n) = \mathbf{0}$

$\Rightarrow$ $a_1\beta_1 + \ldots + a_n\beta_n = \mathbf{0}$

$\Rightarrow$ $a_i = 0$ for each $1 \le i \le n$ [$\because$ $\beta_1, \ldots, \beta_n$ are linearly independent]

$\Rightarrow$ $\alpha = \mathbf{0}$.

$\therefore$ T is non-singular because null space of T consists of zero vector alone.
Hence T is invertible.

Comprehensive Exercise 2

1. Describe explicitly the linear transformation T from F^2 to F^2 such that $T(e_1) = (a, b), T(e_2) = (c, d)$ where $e_1 = (1, 0), e_2 = (0, 1)$. **(Kumaun 2008)**
2. Find a linear transformation $T : \mathbf{R}^2 \to \mathbf{R}^2$ such that $T(1, 0) = (1, 1)$ and $T(0, 1) = (-1, 2)$. Prove that T maps the square with vertices $(0, 0), (1, 0), (1, 1)$ and $(0, 1)$ into a parallelogram.
3. Let $T : \mathbf{R}^2 \to \mathbf{R}$ be the linear transformation for which $T(1, 1) = 3$ and $T(0, 1) = -2$. Find $T(a, b)$.
4. Describe explicitly a linear transformation from V_3 **(R)** into V_4 **(R)** which has its range the subspace spanned by the vectors $(1, 2, 0, -4), (2, 0, -1, -3)$.
5. Find a linear mapping $T : \mathbf{R}^3 \to \mathbf{R}^4$ whose image is generated by $(1, -1, 2, 3)$ and $(2, 3, -1, 0)$.
6. Let F be any field and let T be a linear operator on F^2 defined by $T(a, b) = (a + b, a)$. Show that T is invertible and find a rule for T^{-1} like the one which defines T.
7. Show that the operator T on $\mathbf{R}^3$ defined by $T(x, y, z) = (x + z, x - z, y)$ is invertible and find similar rule defining T^{-1}.
8. Let T be a linear operator on V_3 (**R**) defined by

$$T(a, b, c) = (3a, a - b, 2a + b + c) \ \forall \ (a, b, c) \in V_3(\mathbf{R}).$$

Prove that $(T^2 - I)(T - 3I) = \hat{\mathbf{0}}$. **(Kumaun 2011)**

9. Let T and U be the linear operators on $\mathbf{R}^2$ defined by $T(a,b)=(b,a)$ and $U(a,b)=(a,0)$. Give rules like the one defining T and U for each of the linear transformation $(U+T), UT, TU, T^2, U^2$.

10. Let T be the (unique) linear operator on $\mathbf{C}^3$ for which

$$T(1,0,0)=(1,0,i),\ T(0,1,0)=(0,1,1), T(0,0,1)=(i,1,0).$$

Show that T is not invertible.

11. Show that if two linear transformations of a finite dimensional vector space coincide on a basis of that vector space, then they are identical.

12. If T is a linear transformation on a finite dimensional vector space V such that range (T) is a proper subset of V, show that there exists a non-zero element α in V with $T(\alpha)=\mathbf{0}$.

13. Let $T:\mathbf{R}^3 \to \mathbf{R}^3$ be defined as $T(a,b,c)=(0,a,b)$. Show that

$$T \neq \hat{\mathbf{0}},\ T^2 \neq \hat{\mathbf{0}} \text{ but } T^3 = \hat{\mathbf{0}}.$$

(Kumaun 2008)

14. Let T be a linear transformation from a vector space U into a vector space V with Ker $T \neq \mathbf{0}$. Show that there exist vectors α_1 and α_2 in U such that $\alpha_1 \neq \alpha_2$ and $T\alpha_1 = T\alpha_2$. **(Kumaun 2008)**

15. Let T be a linear transformation from V_3 **(R)** into V_2 **(R)**, and let S be a linear transformation from V_2 **(R)** into V_3 **(R)**. Prove that the transformation ST is not invertible.

16. Let A and B be linear transformations on a finite dimensional vector space V and let $AB=I$. Then A and B are both invertible and $A^{-1}=B$. Give an example to show that this is false when V is not finite dimensional.

17. If A and B are linear transformations (on the same vector space) and if $AB=I$, then A is called a left inverse of B and B is called a right inverse of A. Prove that if A has exactly one right inverse, say B, then A is invertible.

18. Prove that the set of invertible linear operators on a vector space V with the operation of composition forms a group. Check if this group is commutative.

19. Let V and W be vector spaces over the field F and let U be an isomorphism of V onto W. Prove that $T \to UTU^{-1}$ is an isomorphism of $L(V,V)$ onto $L(W,W)$.

20. If $\{\alpha_1, \ldots, \alpha_k\}$ and $\{\beta_1, \ldots \beta_k\}$ are linearly independent sets of vectors in a finite dimensional vector space V, then there exists an invertible linear transformation T on V such that

$$T(\alpha_i)=\beta_i, i=1, \ldots, k.$$

Answers 2

1. $T(x_1, x_2) = (x_1 a + x_2 c, x_1 b + x_2 d)$
2. $T(x_1, x_2) = (x_1 - x_2, x_1 + 2x_2)$
3. $T(a, b) = 5a - 2b$
4. $T(a, b, c) = (a + 2b, 2a - b, -4a - 3b)$
5. $T(a, b, c) = (a + 2b, -a + 3b, 2a - b, 3a)$
6. $T^{-1}(p, q) = (q, p - q)$
7. $T^{-1}(x, y, z) = \left(\frac{1}{2}x + \frac{1}{2}y, z, \frac{1}{2}x - \frac{1}{2}y\right)$
9. $(U + T)(a, b) = (a + b, a)$; $(UT)(a, b) = (b, 0)$; $(TU)(a, b) = (0, a)$
   $T^2(a, b) = (a, b)$; $U^2(a, b) = (a, 0)$
18. Not commutative

13 Isomorphism

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

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 a homomorphism of $U(F)$ into $V(F)$, then f will become 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.

14 Theorems on Isomorphism

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

$$\{\alpha_1, \alpha_2, \ldots, \alpha_n\} \quad \text{and} \quad \{\beta_1, \beta_2, \ldots, \beta_n\}$$

be the bases of U and V respectively.

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

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

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

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

Since in the expression of α as a linear combination of $\alpha_1, \alpha_2, \ldots, \alpha_n$ the scalars $a_1, a_2, \ldots, a_n$ 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

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

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

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

$\Rightarrow$ $a_1 - b_1 = 0, a_2 - b_2 = 0, \ldots, a_n - b_n = 0$, because $\beta_1, \beta_2, \ldots, \beta_n$ are linearly independent

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

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

$\therefore$ f is one-one.

***f* is onto *V*:** If $a_1 \beta_1 + a_2 \beta_2 + \ldots + a_n \beta_n$ is any element of V, then $\exists$ an element $a_1\alpha_1 + a_2 \alpha_2 + \ldots + a_n \alpha_n \in U$ such that

$$f(a_1\alpha_1 + a_2 \alpha_2 + \ldots + a_n \alpha_n) = a_1 \beta_1 + a_2 \beta_2 + \ldots + a_n \beta_n .$$

$\therefore$ f is onto V.

***f* is a linear transformation:** We have

$$\begin{aligned} f[a(a_1\alpha_1 + a_2 \alpha_2 + \ldots + a_n \alpha_n) &+ b(b_1\alpha_1 + b_2 \alpha_2 + \ldots + b_n \alpha_n)] \\ &= f[(aa_1 + bb_1)\alpha_1 + (aa_2 + bb_2)\alpha_2 + \ldots + (aa_n + bb_n)\alpha_n] \\ &= (aa_1 + bb_1)\beta_1 + (aa_2 + bb_2)\beta_2 + \ldots + (aa_n + bb_n)\beta_n \\ &= a(a_1 \beta_1 + a_2 \beta_2 + \ldots + a_n \beta_n) + b(b_1 \beta_1 + b_2 \beta_2 + \ldots + b_n \beta_n) \\ &= a f(a_1\alpha_1 + a_2 \alpha_2 + \ldots + a_n \alpha_n) + b f(b_1\alpha_1 + b_2 \alpha_2 + \ldots + b_n \alpha_n). \end{aligned}$$

$\therefore$ f is a linear transformation.

Hence f is an isomorphism of U onto V.

$\therefore$ $U \cong V$.

Conversely, let $U(F)$ and $V(F)$ be two isomorphic finite dimensional vector spaces. Then to prove that dim U = dim V.

Let dim $U = n$. Let $S = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ be a basis of U. If f is an isomorphism of U onto V, we shall show that

$$S' = \{f(\alpha_1), f(\alpha_2), \ldots, f(\alpha_n)\}$$

is a basis of V. Then V will also be of dimension n.

First we shall show that S' is linearly independent.

Let $\quad a_1 f(\alpha_1) + a_2 f(\alpha_2) + \ldots + a_n f(\alpha_n) = \mathbf{0}'$ [zero vector of V]

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

$\Rightarrow \quad a_1\alpha_1 + a_2\alpha_2 + \ldots + a_n\alpha_n = \mathbf{0}$

[$\because f$ is one-one and $f(\mathbf{0}) = \mathbf{0}'$ where $\mathbf{0}$ is zero vector of U]

$\Rightarrow \quad a_1 = 0, a_2 = 0, \ldots, a_n = 0,$

since $\alpha_1, \alpha_2, \ldots, \alpha_n$ are linearly independent.

$\therefore \quad S'$ is linearly independent.

Now to prove that $L(S') = V$. For this we shall prove that 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 $\quad \alpha = c_1\alpha_1 + c_2\alpha_2 + \ldots + c_n\alpha_n$.

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

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

Thus β is a linear combination of the vectors of S'.

Hence $\quad V = L(S')$.

$\therefore \quad S'$ 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 $V_n(F)$.*

(Gorakhpur 2014)

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

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

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

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

Since in the expression of α as a linear combination of $\alpha_1, \alpha_2, \ldots, \alpha_n$ the scalars $a_1, a_2, \ldots, a_n$ 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 $\alpha = a_1\alpha_1 + a_2\alpha_2 + \ldots + a_n\alpha_n$

and $\quad \beta = b_1\alpha_1 + b_2\alpha_2 + \ldots + b_n\alpha_n$

be any two elements of V. We have

$\quad f(\alpha) = f(\beta)$

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

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

$\Rightarrow \quad a_1 = b_1, a_2 = b_2, \ldots, a_n = b_n$

$\Rightarrow \quad \alpha = \beta$.

$\therefore \quad f$ is one-one.

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

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

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

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

$$\begin{aligned}
& f(a\alpha + b\beta) \\
&= f[a(a_1\alpha_1 + a_2\alpha_2 + \ldots + a_n\alpha_n) + b(b_1\alpha_1 + b_2\alpha_2 + \ldots + b_n\alpha_n)] \\
&= f[(aa_1 + bb_1)\alpha_1 + (aa_2 + bb_2)\alpha_2 + \ldots + (aa_n + bb_n)\alpha_n] \\
&= (aa_1 + bb_1, aa_2 + bb_2, \ldots, aa_n + bb_n) \\
&= (aa_1, aa_2, \ldots, aa_n) + (bb_1, bb_2, \ldots, bb_n) \\
&= a(a_1, a_2, \ldots, a_n) + b(b_1, b_2, \ldots, b_n) \\
&= af(a_1\alpha_1 + a_2\alpha_2 + \ldots + a_n\alpha_n) + bf(b_1\alpha_1 + b_2\alpha_2 + \ldots + b_n\alpha_n) \\
&= af(\alpha) + bf(\beta).
\end{aligned}$$

$\therefore$ f is a linear transformation.

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

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

Illustrative Examples

Example 21: *Let $V(\mathbf{R})$ be the vector space of all complex numbers $a + ib$ over the field of reals $\mathbf{R}$ and let T be a mapping from $V(\mathbf{R})$ to $V_2(\mathbf{R})$ defined as*

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

Show that T is an isomorphism. **(Meerut 2003)**

Solution: **T is one-one:** Let

$$\alpha = a + ib, \beta = c + id$$

be any two members of $V(\mathbf{R})$. Then $a, b, c, d \in \mathbf{R}$.

We have $T(\alpha) = T(\beta)$

$\Rightarrow$ $(a, b) = (c, d)$ $[\because T(\alpha) = (a, b)]$

$\Rightarrow$ $a = c, b = d$

$\Rightarrow$ $a + ib = c + id$

$\Rightarrow$ $\alpha = \beta$.

$\therefore$ T is one-one.

***T* is onto:** Let (a, b) be an arbitrary member of $V_2(\mathbf{R})$. Then $\exists$ a vector $a + ib \in V(\mathbf{R})$ such that $T(a + ib) = (a, b)$. Hence T is onto.

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

$$\begin{aligned}
k_1\alpha + k_2\beta &= k_1(a + ib) + k_2(c + id) \\
&= (k_1a + k_2c) + i(k_1b + k_2d).
\end{aligned}$$

We have

$$T(k_1\alpha + k_2\beta) = (k_1 a + k_2 c, k_1 b + k_2 d), \text{ by definition of } T$$
$$= (k_1 a, k_1 b) + (k_2 c, k_2 d)$$
$$= k_1 (a, b) + k_2 (c, d)$$
$$= k_1 T(a + ib) + k_2 T(c + id) \qquad [\text{by definition of } T]$$
$$= k_1 T(\alpha) + k_2 T(\beta).$$

Hence T is a linear transformation.

Hence T is an isomorphism.

Example 22: *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.*

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

Let $S = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ be a basis of V. We shall first prove that

$$S' = \{f(\alpha_1), f(\alpha_2), \ldots, f(\alpha_n)\}$$

is also a basis of V. We claim that S' is linearly independent. The proof is as follows:

Let $a_1 f(\alpha_1) + a_2 f(\alpha_2) + \ldots + a_n f(\alpha_n) = \mathbf{0}$ [zero vector of V]

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

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

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

since $\alpha_1, \alpha_2, \ldots, \alpha_n$ are linearly independent.

$\therefore$ S' is linearly independent.

Now V is of dimension n and S' is a 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 $c_1, c_2, \ldots, c_n$ such that

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

Now $c_1\alpha_1 + c_2\alpha_2 + \ldots + c_n\alpha_n \in V$ and the f-image of this element is α. Therefore f is onto V. Hence f is an isomorphism of V onto V.

Example 23: *If V is 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 $S = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ be a basis of V.

We shall first prove that

$$S' = \{f(\alpha_1), f(\alpha_2), \ldots, f(\alpha_n)\}$$

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 $f(\alpha_1), f(\alpha_2), \ldots, f(\alpha_n)$. Since f is onto V, therefore $\alpha \in V$ implies that there exists $\beta \in V$ such that $f(\beta) = \alpha$. Now β can be expressed as a linear combination of $\alpha_1, \alpha_2, \ldots, \alpha_n$. Let

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

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

Thus α has been expressed as a linear combination of

$$f(\alpha_1), f(\alpha_2), \ldots, f(\alpha_n).$$

Therefore $L(S') = V$.

Since V is of 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 the 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

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

We have

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

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

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

$\Rightarrow$ $(c_1 - d_1) f(\alpha_1) + (c_2 - d_2) f(\alpha_2) + \ldots + (c_n - d_n) f(\alpha_n) = \mathbf{0}$

$\Rightarrow$ $c_1 - d_1 = 0, c_2 - d_2 = 0, \ldots, c_n - d_n = 0$, since $f(\alpha_1), f(\alpha_2), \ldots, f(\alpha_n)$ are linearly independent

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

$\Rightarrow$ $\gamma = \delta$.

$\therefore$ f is one-one.

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

Example 24: *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 independent if and only if the set $\{\alpha_1, \alpha_2, \ldots, \alpha_r\}$ 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

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

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

$$f(\alpha_1), f(\alpha_2), \ldots, f(\alpha_r)$$

are also linearly independent.

We have

$$a_1 f(\alpha_1) + a_2 f(\alpha_2) + \ldots + a_r f(\alpha_r) = \mathbf{0}', \text{ where } a_1, a_2, \ldots, a_r \in F$$

$\Rightarrow \quad f(a_1\alpha_1 + a_2\alpha_2 + \ldots + a_r\alpha_r) = \mathbf{0}' \quad$ [$\because f$ is a linear transformation]

$\Rightarrow \quad f(a_1\alpha_1 + a_2\alpha_2 + \ldots + a_r\alpha_r) = f(\mathbf{0}) \quad$ [$\because f(\mathbf{0}) = \mathbf{0}'$]

$\Rightarrow \quad a_1\alpha_1 + a_2\alpha_2 + \ldots + a_r\alpha_r = \mathbf{0} \quad$ [$\because f$ is 1-1]

$\Rightarrow \quad a_1 = 0, a_2 = 0, \ldots, a_r = 0$

since the vectors $\alpha_1, \alpha_2, \ldots, \alpha_r$ are linearly independent.

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

Conversely suppose that the vectors

$$f(\alpha_1), f(\alpha_2), \ldots, f(\alpha_r)$$

are linearly independent. Then to show that the vectors $\alpha_1, \alpha_2, \ldots, \alpha_r$ are also linearly independent.

We have

$$a_1 \alpha_1 + a_2 \alpha_2 + \ldots + a_r \alpha_r = \mathbf{0}, \text{ where } a_1, a_2, \ldots, a_r \in F$$

$\Rightarrow \quad f(a_1 \alpha_1 + a_2 \alpha_2 + \ldots + a_r \alpha_r) = f(\mathbf{0})$

$\Rightarrow \quad a_1 f(\alpha_1) + a_2 f(\alpha_2) + \ldots + a_r f(\alpha_r) = \mathbf{0}'$

[$\because f$ is a linear transformation]

$\Rightarrow \quad a_1 = 0, a_2 = 0, \ldots, a_r = 0$

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

Hence the vectors $\alpha_1, \alpha_2, \ldots, \alpha_r$ are also linearly independent.

Comprehensive Exercise 3

1. Let $T : V_2(\mathbf{R}) \to V_2(\mathbf{R})$ be defined as $T(a_1, b_1) = (b_1, a_1)$.
   Show that T is an isomorphism.
2. $V(F)$ and $W(F)$ are two finite dimensional vector spaces such that dim $V = \dim W$. If f is an isomorphism of V into W, prove that f must map V into W. **(Meerut 2002, 03; Garhwal 10B; Kumaun 14)**
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. Give an example of a one-one linear transformation of an infinite dimensional vector space which is not an isomorphism.
5. 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.
6. Prove that a finite dimensional vector space $V(\mathbf{R})$ with dimension $V = n$ is isomorphic to $\mathbf{R}^n$.

Objective Type Questions

Multiple Choice Questions

Indicate the correct answer for each question by writing the corresponding letter from (a), (b), (c) *and* (d).

1. If $U(F)$ and $V(F)$ are two vector spaces and T is a linear transformation from U into V, then range of T is a subspace of
   (a) U (b) V
   (c) $U \cup V$ (d) none of these

2. 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 is the dimension of the
   (a) range of T (b) null space of T
   (c) vector space U (d) vector space V

3. Let U be an n-dimensional vector space over the field F, and let V be an m-dimensional vector space over F. Then the vector space $L(U, V)$ of all linear transformations from U into V is finite dimensional and is of dimension
   (a) m (b) n
   (c) mn (d) none of these

4. If $T : V_2(\mathbf{R}) \to V_3(\mathbf{R})$ defined as $T(a, b) = (a + b, a - b, b)$ is a linear transformation, then nullity of T is
   (a) 0 (b) 1
   (c) 2 (d) none of these
   (Kumaun 2015)

5. If T is a linear transformation from a vector space V into a vector space W, then the condition for T^{-1} to be a linear transformation from W to V is
   (a) T should be one-one (b) T should be onto
   (c) T should be one-one and onto (d) none of these

6. Let F be any field and let T be a linear operator on F^2 defined by $T(a, b) = (a + b, a)$. Then $T^{-1}(a, b) =$
   (a) $(b, a - b)$ (b) $(a - b, b)$
   (c) $(a, a + b)$ (d) none of these
   (Kumaun 2014)

7. If T is a linear transformation $T(a, b) = (\alpha a + \beta b, \gamma a + \gamma b)$, the T is invertible if
   (a) $\alpha\beta - \gamma\delta = 0$ (b) $\alpha\beta - \gamma\delta \neq 0$
   (c) $\alpha\delta - \beta\gamma = 0$ (d) $\alpha\delta - \beta\gamma \neq 0$ **(Kumaun 2015)**

8. Let $V(F)$ be a vector space and let T_1, T_2 be linear Transformations on V. If T_1 and T_2 are invertible then
   (a) $(T_1T_2)^{-1} = T_2\, T_1^{-1}$
   (b) $(T_1T_2)^{-1} = T_1\, T_2^{-1}$
   (c) $(T_1T_2)^{-1} = T_2^{-1}T_1^{-1}$
   (d) $(T_1T_2)^{-1} = T_1^{-1}T_2^{-1}$.

9. If the mapping $f: U \to V$ is a linear transformation from the vector space $U(V)$ to the vector space $V(U)$, then:
   (a) $f(-\alpha) = f(\alpha)$
   (b) $f(-\alpha) = -f(\alpha)$
   (c) $f(-\alpha) = \pm f(\alpha)$
   (d) none of these

(Kumaun 2007)

10. Which of the following function $T: \mathbf{R}^2 \to \mathbf{R}^2$ is not linear ?
   (a) $T(a, b) = (b, a)$
   (b) $T(a, b) = (a + b, a)$
   (c) $T(a, b) = (a, a + b)$
   (d) $T(a, b) = (1 + a, b)$

(Kumaun 2010)

11. Let $T: U \to V$ be a linear transformation then:
   (a) $T(\mathbf{0}) = \mathbf{0}$
   (b) $T(-\alpha) = -T(\alpha),\ \alpha \in U$
   (c) $T(\alpha - \beta) = T(\alpha) - T(\beta)\ \forall\ \alpha, \beta \in U$
   (d) all above

(Kumaun 2014)

12. Which of the following function is linear ?
   (a) $T: \mathbf{R}^2 \to \mathbf{R}^2$ such that $T(a, b) = (b, a)$
   (b) $T: \mathbf{R}^2 \to \mathbf{R}^2$ such that $T(x, y) = xy$
   (c) $T: \mathbf{R}^2 \to \mathbf{R}^2$ such that $T(x, y) = (1 + x, y)$
   (d) $T: \mathbf{R}^2 \to \mathbf{R}^2$ such that $T(x, y) = |x + y|$

13. If $T: V_2(\mathbf{R}) \to V_3(\mathbf{R})$ is a linear transformation defined by $T(x, y) = (x + y, x - y, y)$, then nullity of T is
   (a) 0
   (b) 1
   (c) 2
   (d) 3

14. If T is a linear transformation $T: V_F \to W_F$ and $\{v_1, v_2, \ldots, v_n\}$ is a basis for V_F then $\{T(v_1), T(v_2), \ldots, T(v_n)\}$ will be a basis for W_F if following is true
   (a) T is one-one
   (b) T is onto
   (c) T is one-one and onto
   (d) none of above

15. If A, B and C are linear transformations on a vector space $V(F)$ such that $AB = CA = I$, then
   (a) A is invertible and $A^{-1} = B = C$
   (b) B is invertible and $B^{-1} = C$
   (c) C is invertible and $C^{-1} = B$
   (d) A, B and C are invertible and $A^{-1} = B^{-1} = C^{-1}$

16. Let T be a linear transformation from a vector space $U(F)$ into a vector space $V(F)$. Then T is non-singular if and only if :

(a) T is one-one onto
(b) T is one-one
(c) T is onto
(d) none of these

17. T is a linear transformation on an n-dimensional vector space $V(F)$ such that range and null spaces of T are identical. Then

(a) $n = 0$
(b) $n > 0$
(c) n is prime
(d) n is even

18. If linear transformation $T : V \to W$ is invertible then :

(a) $T(\alpha) = T(\beta) \Leftrightarrow \alpha = \beta$
(b) $T(\alpha) = T(\beta) \Leftrightarrow \alpha \neq \beta$
(c) $T(\alpha) = T(\beta) \Leftrightarrow TT^{-1}(\alpha) \neq \beta$
(d) none of these

19. Let T_1 and T_2 be linear operators on $\mathbf{R}^2$ defined as follows:

$T_1(x_1, x_2) = (x_2, x_1)$, $T_2(x_1, x_2) = (x_2, 0)$. Then $(T_1T_2)(x_1, x_2) =$

(a) $(x_1, 0)$
(b) $(x_2, 0)$
(c) $(0, x_2)$
(d) $(0, x_1)$

20. T is non singular iff

(a) nullity $(T) \neq \mathbf{0}$
(b) nullity $(T) = \mathbf{0}$
(c) rank $(T) = \mathbf{0}$
(d) nullity $(T) = \text{rank}\,(T)$

Fill in the Blank(s)

Fill in the blanks "......" so that the following statements are complete and correct.

1. Let $V(F)$ be a vector space. A linear operator on V is a function T from V into V such that $T(a\alpha + b\beta) =$ ...... for all α, β in V and for all a, b in F.

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

3. Let U and V be vector spaces over the field F and let T be a linear transformation from U into V. Suppose that U is finite dimensional. Then

$$\text{rank}\,(T) + \text{nullity}\,(T) = \ldots\ldots$$

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. Then n is ....

5. The vector space of all linear operators on an n-dimensional vector space U is of dimension ....

6. Let $V(F)$ be a vector space and let A, B be linear transformations on V. If A and B are invertible then AB is invertible and $(AB)^{-1} =$ ......

7. A linear operator T on $\mathbf{R}^2$ defined by $T(x, y) = (ax + by, cx + dy)$ will be invertible iff ....

True or False

Write 'T' for true and 'F' for false statement.

1. Let $U(F)$ and $V(F)$ be two vector spaces and let T be a linear transformation from U into V. Then the null space of T is the set of all vectors α in U such that $T(\alpha) = \alpha$.
2. The function $T : \mathbf{R}^2 \to \mathbf{R}^2$ defined by $T(a, b) = (1 + a, b)$ is a linear transformation.
3. Two linear transformations from U into V are equal if they agree on a basis of U.
4. For two linear operators T and S on $\mathbf{R}^2$, $TS = \hat{\mathbf{0}} \to ST = \hat{\mathbf{0}}$.
5. If S and T are linear operators on a vector space U, then
$$(S + T)^2 = S^2 + 2ST + T^2.$$
6. The identity operator on a vector space is always invertible.

Answers

Multiple Choice Questions

1.	(b)	2.	(a)	3.	(c)	4.	(a)	5.	(c)
6.	(a)	7.	(d)	8.	(c)	9.	(b)	10.	(d)
11.	(d)	12.	(a)	13.	(a)	14.	(c)	15.	(a)
16.	(c)	17.	(d)	18.	(a)	19.	(d)	20.	(b)

Fill in the Blank(s)

1.	$aT(\alpha) + bT(\beta)$	2.	U	3.	$\dim U$
4.	even	5.	n^2	6.	$B^{-1} A^{-1}$
7.	$ad - bc \neq 0$				

True or False

1.	F	2.	F	3.	T	4.	F	5.	F
6.	T								

□

Chapter

2

Matrices and Linear Transformations

1 Matrix

Definition: *Let F be any field. A set of mn elements of F arranged in the form of a rectangular array having m rows and n columns is called an m $\times$ n matrix over the field F.*

An $m \times n$ matrix is usually written as

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}.$$

In a compact form the above matrix is represented by $A = [a_{ij}]_{m \times n}$. The element a_{ij} is called the $(i, j)^{\text{th}}$ element of the matrix A. In this element the first suffix i will always denote the number of row in which this element occurs.

If in a matrix A the number of rows is equal to the number of columns and is equal to n, then A is called a **square matrix** of order n and the elements a_{ij} for which $i = j$ constitute its **principal diagonal.**

Unit matrix: *A square matrix each of whose diagonal elements is equal to* 1 *and each of whose non-diagonal elements is equal to zero is called a unit matrix or an identity matrix.* We shall denote it by I. Thus if I is unit matrix of order n, then $I = [\delta_{ij}]_{n \times n}$ where δ_{ij} is Kronecker delta.

Diagonal matrix: *A square matrix is said to be a diagonal matrix if all the elements lying above and below the principal diagonal are equal to* 0. For example,

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 2+i & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 5 \end{bmatrix}$$

is a diagonal matrix of order 4 over the field of complex numbers.

Null matrix: The $m \times n$ matrix whose elements are all zero is called the null matrix or (zero matrix) of the type $m \times n$.

Equality of two matrices: Definition:

Let $A = [a_{ij}]_{m \times n}$ and $B = [b_{jk}]_{m \times n}$. Then

$A = B$ if $a_{ij} = b_{ij}$ for each pair of subscripts i and j.

Addition of two matrices: Definition:

Let $A = [a_{ij}]_{m \times n}$, $B = [b_{ij}]_{m \times n}$. Then we define

$$A + B = [a_{ij} + b_{ij}]_{m \times n}.$$

Multiplication of a matrix by a scalar: Definition:

Let $A = [a_{ij}]_{m \times n}$ and $a \in F$ *i.e.,* a be a scalar. Then we define

$$aA = [aa_{ij}]_{m \times n}.$$

Multiplication of two matrices: Definition:

Let $A = [a_{ij}]_{m \times n}$, $B = [b_{jk}]_{n \times p}$ *i.e.*, the number of columns in the matrix A is equal to the number of rows in the matrix B. Then we define

$$AB = \left[\sum_{j=1}^{n} a_{ij} \, b_{jk} \right]_{m \times p}$$

i.e., AB is an $m \times p$ matrix whose $(i, k)^{\text{th}}$ element is equal to $\sum_{j=1}^{n} a_{ij} \, b_{jk}$.

If A and B are both square matrices of order n, then both the products AB and BA exist but in general $AB \neq BA$.

Transpose of a matrix: Definition:

Let $A = [a_{ij}]_{m \times n}$. The $n \times m$ matrix A^T obtained by interchanging the rows and columns of A is called the transpose of A. Thus $A^T = [b_{ij}]_{n \times m}$, where $b_{ij} = a_{ji}$, *i.e.*, the $(i, j)^{\text{th}}$ element of A^T is the $(j, i)^{\text{th}}$ element of A. If A is an $m \times n$ matrix and B is an $n \times p$ matrix, it can be shown that $(AB)^T = B^T A^T$. The transpose of a matrix A is also denoted by A^t or by A'.

Determinant of a square matrix: Let P_n denote the group of all permutations of degree n on the set $\{1, 2, \ldots, n\}$. If $\theta \in P_n$, then $\theta(i)$ will denote the image of i under θ. The symbol $(-1)^{\theta}$ for $\theta \in P_n$ will mean $+1$ if θ is an *even* permutation and -1 if θ is an *odd* permutation.

Definition: Let $A = [a_{ij}]_{n \times n}$. Then the *determinant* of A, written as det A or $|A|$ or $|a_{ij}|_{n \times n}$ is the element

$$\sum_{\theta \in P_n} (-1)^{\theta}\, a_{1\theta(1)}\, a_{2\theta(2)} \ldots a_{n\theta(n)} \text{ in } F.$$

The number of terms in this summation is $n!$ because there are $n!$ permutations in the set P_n.

We shall often use the notation

$$\begin{vmatrix} a_{11} & \ldots & a_{1n} \\ a_{21} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots \\ a_{n1} & \ldots & a_{nn} \end{vmatrix}$$

for the determinant of the matrix $[a_{ij}]_{n \times n}$.

The following properties of determinants are worth to be noted :

(i) The determinant of a unit matrix is always equal to 1.

(ii) The determinant of a null matrix is always equal to 0.

(iii) If $A = [a_{ij}]_{n \times n}$, $B = [b_{ij}]_{n \times n}$, then $\det(AB) = (\det A)(\det B)$.

Cofactors: Definition: Let $A = [a_{ij}]_{n \times n}$. We define

A_{ij} = cofactor of a_{ij} in A

$= (-1)^{i+j}$. [Determinant of the matrix of order $n-1$ obtained by deleting the row and column of A passing through a_{ij}].

It should be noted that

$$\sum_{i=1}^{n} a_{ik} A_{ij} = 0 \text{ if } k \neq j$$

or $\qquad = \det A$ if $k = j$.

Adjoint of a square matrix: Definition: Let $A = [a_{ij}]_{n \times n}$.

The $n \times n$ matrix which is the transpose of the matrix of cofactors of A is called the adjoint of A and is denoted by adj A.

It should be remembered that

$$A\,(\text{adj } A) = (\text{adj } A)\, A = (\det A)\, I$$

where I is the unit matrix of order n.

Inverse of a square matrix: Definition: Let A be a square matrix of order n. If there exists a square matrix B of order n such that

$$AB = I = BA$$

then A is said to be invertible and B is called the inverse of A.

Also we write $B = A^{-1}$.

The following results should be remembered :

(i) The necessary and sufficient condition for a square matrix A to be invertible is that det $A \neq 0$.

(ii) If A is invertible, then A^{-1} is unique and

$$A^{-1} = \frac{1}{\det A} (\text{adj. } A).$$

(iii) If A and B are invertible square matrices of order n, then AB is also invertible and $(AB)^{-1} = B^{-1} A^{-1}$.

(iv) If A is invertible, so is A^{-1} and $(A^{-1})^{-1} = A$.

Elementary row operations on a matrix:

Definition: Let A be an $m \times n$ matrix over the field F. The following three operations are called elementary row operations :

(1) multiplication of any row of A by a non-zero element c of F.

(2) addition to the elements of any row of A the corresponding elements of any other row of A multiplied by any element a in F.

(3) interchange of two rows of A.

Row equivalent matrices: Definition: If A and B are $m \times n$ matrices over the field F, then B is said to be row equivalent to A if B can be obtained from A by a finite sequence of elementary row operations. It can be easily seen that the relation of being row equivalent is an equivalence relation in the set of all $m \times n$ matrices over F.

Row reduced Echelon matrix:

Definition:

An $m \times n$ matrix R is called a row reduced echelon matrix if :

(1) Every row of R which has all its entries 0 occurs below every row which has a non-zero entry.

(2) The first non-zero entry in each non-zero row is equal to 1.

(3) If the first non-zero entry in row i appears in column k_i, then all other entries in column k_i are zero.

(4) If r is the number of non-zero rows, then

$$k_1 < k_2 < \ldots < k_r$$

(*i.e.*, the first non-zero entry in row i is to the left of the first non-zero entry in row $i+1$).

Row and column rank of a matrix:

Definition: Let $A = [a_{ij}]_{m \times n}$ be an $m \times n$ matrix over the field F. The row vectors of A are the vectors $\alpha_1, \ldots, \alpha_m \in V_n(F)$ defined by

$$\alpha_i = (a_{i1}, a_{i2}, \ldots, a_{in}), 1 \leq i \leq m.$$

The row space of A is the subspace of $V_n(F)$ spanned by these vectors. The **row rank** of A is the dimension of the row space of A.

The column vectors of A are the vectors $\beta_1, \ldots, \beta_n \in V_m(F)$ defined by

$$\beta_j = (a_{1j}, a_{2j}, \ldots, a_{mj}), 1 \leq j \leq n.$$

The **column space** of A is the subspace of $V_m(F)$ spanned by these vectors. The **column rank** of A is the dimension of the column space of A.

The following two results are to be remembered:

(1) Row equivalent matrices have the same row space.

(2) If R is a non-zero row reduced Echelon matrix, then the non-zero row vectors of R are linearly independent and therefore they form a basis for the row space of R.

In order to find the row rank of a matrix A, we should reduce it to row reduced Echelon matrix R by elementary row operations. The number of non-zero rows in R will give us the row rank of A.

2 Representation of Transformations by Matrices

Matrix of a linear transformation: Let U be an n-dimensional vector space over the field F and let V be an m-dimensional vector space over F. Let

$$B = \{\alpha_1, \ldots, \alpha_n\} \text{ and } B' = \{\beta_1, \ldots, \beta_m\}$$

be ordered bases for U and V respectively. Suppose T is a linear transformation from U into V. We know that T is completely determined by its action on the vectors α_j belonging to a basis for U. Each of the n vectors $T(\alpha_j)$ is uniquely expressible as a linear combination of $\beta_1, \ldots, \beta_m$ because $T(\alpha_j) \in V$ and these m vectors form a basis for V. Let for $j = 1, 2, \ldots, n$,

$$T(\alpha_j) = a_{1j}\beta_1 + a_{2j}\beta_2 + \ldots + a_{mj}\beta_m = \sum_{i=1}^{m} a_{ij}\beta_i .$$

The scalars $a_{1j}, a_{2j}, \ldots, a_{mj}$ are the coordinates of $T(\alpha_j)$ in the ordered basis B'. The $m \times n$ matrix whose j^{th} column ($j = 1, 2, \ldots, n$) consists of these coordinates is called the **matrix of the linear transformation T relative to the pair of ordered bases B and B'**. We shall denote it by the symbol $[T; B; B']$ or simply by $[T]$ if the bases are understood. Thus

$$\begin{aligned}[T] &= [T; B; B'] \\ &= \text{matrix of } T \text{ relative to the ordered bases } B \text{ and } B' \\ &= [a_{ij}]_{m \times n}\end{aligned}$$

where $$T(\alpha_j) = \sum_{i=1}^{m} a_{ij}\beta_i, \text{ for each } j = 1, 2, \ldots, n. \qquad \ldots(1)$$

The coordinates of $T(\alpha_1)$ in the ordered basis B' form the first column of this matrix, the coordinates of $T(\alpha_2)$ in the ordered basis B' form the second column of this matrix and so on.

The $m \times n$ matrix $[a_{ij}]_{m \times n}$ completely determines the linear transformation T through the formulae given in (1). Therefore the matrix $[a_{ij}]_{m \times n}$ represents the transformation T.

Note: Let T be a linear transformation from an n-dimensional vector space $V(F)$ into itself. Then in order to represent T by a matrix, it is most convenient to use the same ordered basis in each case, *i.e.*, to take $B = B'$. The representing matrix will then be called the **matrix of T relative to the ordered basis B** and will be denoted by $[T ; B]$ or sometimes also by $[T]_B$.

Thus if $B = \{\alpha_1, \ldots, \alpha_n\}$ is an ordered basis for V, then

$$[T]_B \text{ or } [T ; B] = \text{matrix of } T \text{ relative to the ordered basis } B$$
$$= [a_{ij}]_{n \times n},$$

where $\quad T(\alpha_j) = \sum_{i=1}^{n} a_{ij} \alpha_i$, for each $j = 1, 2, \ldots, n$.

Illustrative Examples

Example 1: *Let T be a linear transformation on the vector space $V_2(F)$ defined by $T(a, b) = (a, 0)$.*

Write the matrix of T relative to the standard ordered basis of $V_2(F)$.

Solution: Let $B = \{\alpha_1, \alpha_2\}$ be the standard ordered basis for $V_2(F)$. Then $\alpha_1 = (1, 0), \alpha_2 = (0, 1)$.

We have $\quad T(\alpha_1) = T(1, 0) = (1, 0)$.

Now let us express $T(\alpha_1)$ as a linear combination of vectors in B. We have

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

Thus $1, 0$ are the coordinates of $T(\alpha_1)$ with respect to the ordered basis B. These coordinates will form the first column of matrix of T relative to ordered basis B.

Again $\quad T(\alpha_2) = T(0, 1) = (0, 0) = 0(1, 0) + 0(0, 1)$.

Thus $0, 0$ are the coordinates of $T(\alpha_2)$ and will form second column of matrix of T relative to ordered basis B.

Thus matrix of T relative to ordered basis B

$$= [T]_B \quad \text{or} \quad [T ; B] = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

Example 2: *Let $V(\mathbf{R})$ be the vector space of all polynomials in x with coefficients in $\mathbf{R}$ of the form*

$$f(x) = a_0 x^0 + a_1 x + a_2 x^2 + a_3 x^3$$

i.e., the space of polynomials of degree three or less. The differentiation operator D is a linear transformation on V. The set

$$B = \{\alpha_1, \ldots, \alpha_4\} \text{ where } \alpha_1 = x^0, \alpha_2 = x^1, \alpha_3 = x^2, \alpha_4 = x^3$$

is an ordered basis for V. Write the matrix of D relative to the ordered basis B.

Solution: We have

$$D(\alpha_1) = D(x^0) = 0 = 0x^0 + 0x^1 + 0x^2 + 0x^3$$
$$= 0\alpha_1 + 0\alpha_2 + 0\alpha_3 + 0\alpha_4$$
$$D(\alpha_2) = D(x^1) = x^0 = 1x^0 + 0x^1 + 0x^2 + 0x^3$$
$$= 1\alpha_1 + 0\alpha_2 + 0\alpha_3 + 0\alpha_4$$

$$D(\alpha_3) = D(x^2) = 2x^1 = 0x^0 + 2x^1 + 0x^2 + 0x^3$$
$$= 0\alpha_1 + 2\alpha_2 + 0\alpha_3 + 0\alpha_4$$
$$D(\alpha_4) = D(x^3) = 3x^2 = 0x^0 + 0x^1 + 3x^2 + 0x^3$$
$$= 0\alpha_1 + 0\alpha_2 + 3\alpha_3 + 0\alpha_4.$$

$\therefore$ the matrix of D relative to the ordered basis B

$$= [D; B] = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}_{4\times 4}.$$

Theorem 1: *Let U be an n-dimensional vector space over the field F and let V be an m-dimensional vector space over F. Let B and B' be ordered bases for U and V respectively. Then corresponding to every matrix $[a_{ij}]_{m\times n}$ of mn scalars belonging to F there corresponds a unique linear transformation T from U into V such that*

$$[T; B; B'] = [a_{ij}]_{m\times n}.$$

Proof: Let $B = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ and $B' = \{\beta_1, \beta_2, \ldots, \beta_m\}$.

Now $\sum_{i=1}^{m} a_{i1}\beta_i, \ \sum_{i=1}^{m} a_{i2}\beta_i, \ldots, \sum_{i=1}^{m} a_{in}\beta_i$

are vectors belonging to V because each of them is a linear combination of the vectors belonging to a basis for V. It should be noted that the vector $\sum_{i=1}^{m} a_{ij}\beta_i$ has been obtained with the help of the j^{th} column of the matrix $[a_{ij}]_{m\times n}$.

Since B is a basis for U, therefore by the theorem 2 of article 6 of chapter 'Linear Transformations' there exists a unique linear transformation T from U into V such that

$$T(\alpha_j) = \sum_{i=1}^{m} a_{ij}\beta_i \text{ where } j = 1, 2, \ldots, n. \qquad \ldots(1)$$

By our definition of matrix of a linear transformation , we have from (1)

$$[T; B; B'] = [a_{ij}]_{m\times n}.$$

Note: If we take $V = U$, then in place of B', we also take B. In that case the above theorem will run as :

Let V be an n-dimensional vector space over the field F and B be an ordered basis or **co-ordinate system** *for V. Then corresponding to every matrix $[a_{ij}]_{n\times n}$ of n^2 scalars belonging to F there corresponds a unique linear transformation T from V into V such that*

$$[T; B] \text{ or } [T]_B = [a_{ij}]_{n\times n}.$$

Explicit expression for a linear transformation in terms of its matrix: Now our aim is to establish a formula which will give us the image of any vector under a linear transformation T in terms of its matrix.

Theorem 2: *Let T be a linear transformation from an n-dimensional vector space U into an m-dimensional vector space V and let B and B' be ordered bases for U and V respectively. If A is the matrix of T relative to B and B' then $\forall\, \alpha \in U$, we have*

$$[T(\alpha)]_{B'} = A[\alpha]_B$$

where $[\alpha]_B$ *is the co-ordinate matrix of* α *with respect to ordered basis* B *and* $[T(\alpha)]_{B'}$ *is co-ordinate matrix of* $T(\alpha) \in V$ *with respect to* B'.

Proof: Let $B = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$

and $B' = \{\beta_1, \beta_2, \ldots, \beta_m\}$.

Then $A = [T; B; B'] = [a_{ij}]_{m \times n}$,

where $T(\alpha_j) = \sum_{i=1}^{m} a_{ij} \beta_i, \; j = 1, 2, \ldots, n.$ ...(1)

If $\alpha = x_1\alpha_1 + \ldots + x_n \alpha_n$ is a vector in U, then

$$T(\alpha) = T\left(\sum_{j=1}^{n} x_j \alpha_j\right)$$

$$= \sum_{j=1}^{n} x_j T(\alpha_j)$$ [$\because$ T is a linear transformation]

$$= \sum_{j=1}^{n} x_j \sum_{i=1}^{m} a_{ij} \beta_i$$ [From (1)]

$$= \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} x_j\right) \beta_i.$$ ...(2)

The co-ordinate matrix of $T(\alpha)$ with respect to ordered basis B' is an $m \times 1$ matrix. From (2), we see that the i^{th} entry of this column matrix $[T(\alpha)]_{B'}$

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

i.e., the coefficient of β_i in the linear combination (2) for $T(\alpha)$.

If X is the co-ordinate matrix $[\alpha]_B$ of α with respect to ordered basis B, then X is an $n \times 1$ matrix. The product AX will be an $m \times 1$ matrix. The i^{th} entry of this column matrix AX will be

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

$\therefore$ $[T(\alpha)]_{B'} = AX = A[\alpha]_B = [T; B; B'][\alpha]_B$.

Note: If we take $U = V$, then the above result will be

$$[T(\alpha)]_B = [T]_B [\alpha]_B.$$

Matrices of Identity and Zero transformations:

Theorem 3: *Let* $V(F)$ *be an n-dimensional vector space and* B *be any ordered basis for* V. *If* I *be the identity transformation and* $\hat{0}$ *be the zero transformation on* V, *then*

(i) $[I; B] = I$ *(unit matrix of order n) and* **(Kumaun 2007)**

(ii) $[\hat{0}; B] =$ *null matrix of the type* $n \times n$. **(Kumaun 2007)**

Proof: Let $B = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$.

(i) We have $I(\alpha_j) = \alpha_j, \; j = 1, 2, \ldots, n$

$$= 0\alpha_1 + \ldots + 1\alpha_j + 0\alpha_{j+1} + \ldots + 0\alpha_n$$

$$= \sum_{i=1}^{n} \delta_{ij}\, \alpha_i \text{, where } \delta_{ij} \text{ is Kronecker delta.}$$

$\therefore$ By def. of matrix of a linear transformation , we have

$$[I \,;\, B] = [\delta_{ij}]_{n \times n} = I \quad i.e., \text{ unit matrix of order } n.$$

(ii) We have $\hat{\mathbf{0}}\,(\alpha_j) = \mathbf{0},\ j = 1, 2, \ldots, n$

$$= 0\alpha_1 + 0\alpha_2 + \ldots + 0\alpha_n$$

$$= \sum_{i=1}^{n} o_{ij}\, \alpha_i \text{, where each } o_{ij} = 0.$$

$\therefore$ By def. of matrix of a linear transformation, we have

$$[\,\hat{\mathbf{0}} \,;\, B\,] = [o_{ij}]_{n \times n} = \text{null matrix of the type } n \times n.$$

Matrix of the Sum of Linear Transformation :

Theorem 4: *Let T and S be linear transformations from an n-dimensional vector space U into an m-dimensional vector space V and let B and B' be ordered bases for U and V respectively. Then*

(i) $[T + S \,;\, B \,;\, B'] = [T \,;\, B \,;\, B'] + [S \,;\, B \,;\, B']$

(ii) $[cT \,;\, B \,;\, B'] = c\,[T \,;\, B \,;\, B']$ *where c is any scalar.*

Proof: Let $B = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ and $B' = \{\beta_1, \beta_2, \ldots, \beta_m\}$.

Let $[a_{ij}]_{m \times n}$ be the matrix of T relative to B, B'. Then

$$T\,(\alpha_j) = \sum_{i=1}^{m} a_{ij}\, \beta_i \,,\ j = 1, 2, \ldots, n.$$

Also let $[b_{ij}]_{m \times n}$ be the matrix of S relative to B, B'. Then

$$S\,(\alpha_j) = \sum_{i=1}^{m} b_{ij}\, \beta_i \,,\ j = 1, 2, \ldots, n.$$

(i) We have

$$(T + S)\,(\alpha_j) = T\,(\alpha_j) + S\,(\alpha_j),\ \ j = 1, 2, \ldots, n$$

$$= \sum_{i=1}^{m} a_{ij}\, \beta_i + \sum_{i=1}^{m} b_{ij}\, \beta_i = \sum_{i=1}^{m} (a_{ij} + b_{ij})\, \beta_i \,.$$

$\therefore$ matrix of $T + S$ relative to $B, B' = [a_{ij} + b_{ij}]_{m \times n}$

$$= [a_{ij}]_{m \times n} + [b_{ij}]_{m \times n} \,,$$

$\therefore$ $$[T + S \,;\, B \,;\, B'] = [T \,;\, B \,;\, B'] + [S \,;\, B \,;\, B'].$$

(ii) We have $(cT)\,(\alpha_j) = cT\,(\alpha_j),\ j = 1, 2, \ldots, n$

$$= c \sum_{i=1}^{m} a_{ij}\, \beta_i = \sum_{i=1}^{m} (ca_{ij})\, \beta_i \,.$$

$\therefore$ $$[cT \,;\, B \,;\, B'] = \text{matrix of } cT \text{ relative to } B, B'$$

$$= [ca_{ij}]_{m \times n} = c\,[a_{ij}]_{m \times n} = c\,[T \,;\, B \,;\, B'].$$

Matrix of the Product of Linear Transformation :

Theorem 5: *Let U, V and W be finite dimensional vector spaces over the field F; let T be a linear transformation from U into V and S a linear transformation from V into W. Further let B, B' and B'' be ordered bases for spaces U, V and W respectively. If A is the matrix of T relative to the pair B, B' and D is the matrix of S relative to the pair B', B'' then the matrix of the composite transformation ST relative to the pair B, B'' is the product matrix $C = DA$.*

Proof: Let dim $U = n$, dim $V = m$ and dim $W = p$. Further let

$$B = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}, B' = \{\beta_1, \beta_2, \ldots, \beta_m\} \text{ and } B'' = \{\gamma_1, \gamma_2, \ldots, \gamma_p\}.$$

Let $A = [a_{ij}]_{m \times n}$, $D = [d_{ki}]_{p \times m}$ and $C = [c_{kj}]_{p \times n}$. Then

$$T(\alpha_j) = \sum_{i=1}^{m} a_{ij} \beta_i, \ j = 1, 2, \ldots, n, \quad \ldots(1)$$

$$S(\beta_i) = \sum_{k=1}^{p} d_{ki} \gamma_k, \ i = 1, 2, \ldots, m. \quad \ldots(2)$$

and $$(ST)(\alpha_j) = \sum_{k=1}^{p} c_{kj} \gamma_k, j = 1, 2, \ldots, n. \quad \ldots(3)$$

We have $$(ST)(\alpha_j) = S[T(\alpha_j)], \ j = 1, 2, \ldots, n$$

$$= S\left(\sum_{i=1}^{m} a_{ij} \beta_i\right) \quad \text{[From (1)]}$$

$$= \sum_{i=1}^{m} a_{ij} S(\beta_i) \quad [\because S \text{ is linear}]$$

$$= \sum_{i=1}^{m} a_{ij} \sum_{k=1}^{p} d_{ki} \gamma_k \quad \text{[From (2)]}$$

$$= \sum_{k=1}^{p} \left(\sum_{i=1}^{m} d_{ki} a_{ij}\right) \gamma_k. \quad \ldots(4)$$

Therefore from (3) and (4), we have

$$c_{kj} = \sum_{i=1}^{m} d_{ki} a_{ij}, j = 1, 2, \ldots, n; k = 1, 2, \ldots, p.$$

$$\therefore \quad [c_{kj}]_{p \times n} = \left[\sum_{i=1}^{m} d_{ki} a_{ij}\right]_{p \times n}$$

$$= [d_{ki}]_{p \times m} [a_{ij}]_{m \times n}, \text{ by def. of product of two matrices.}$$

Thus $C = DA$.

Note: If $U = V = W$, then the statement and proof of the above theorem will be as follows :

Let V be an n-dimensional vector space over the field F; let T and S be linear transformations of V. Further let B be an ordered basis for V. If A is the matrix of T relative to B, and D is the matrix of S relative to B, then the matrix of the composite transformation ST relative to B is the product matrix

$$C = DA \ \text{i.e.,} \ [ST]_B = [S]_B [T]_B.$$

Proof: Let $B = \{\alpha_1, \alpha_2, \dots, \alpha_n\}$.

Let $A = [a_{ij}]_{n \times n}$, $D = [d_{ki}]_{n \times n}$ and $C = [c_{kj}]_{n \times n}$.

Then $$T(\alpha_j) = \sum_{i=1}^{n} a_{ij}\, \alpha_i,\ j = 1, 2, \dots, n, \quad \dots(1)$$

$$S(\alpha_i) = \sum_{k=1}^{n} d_{ki}\, \alpha_k,\ i = 1, 2, \dots, n, \quad \dots(2)$$

and $$(ST)(\alpha_j) = \sum_{k=1}^{n} c_{kj}\, \alpha_k,\ j = 1, 2, \dots, n. \quad \dots(3)$$

We have $$(ST)(\alpha_j) = S[T(\alpha_j)]$$
$$= S\left(\sum_{i=1}^{n} a_{ij}\, \alpha_i\right) = \sum_{i=1}^{n} a_{ij}\, S(\alpha_i) = \sum_{i=1}^{n} a_{ij} \sum_{k=1}^{n} d_{ki}\, \alpha_k$$
$$= \sum_{k=1}^{n}\left(\sum_{i=1}^{n} d_{ki}\, a_{ij}\right)\alpha_k. \quad \dots(4)$$

$\therefore$ from (3) and (4), we have
$$c_{kj} = \sum_{i=1}^{n} d_{ki}\, a_{ij}.$$

$\therefore$ $$[c_{kj}]_{n \times n} = \left[\sum_{i=1}^{n} d_{ki}\, a_{ij}\right]_{n \times n} = [d_{k\,i}]_{n \times n}\,[a_{i\,j}]_{n \times n}.$$

$\therefore$ $$C = DA.$$

Theorem 6: *Let U be an n-dimensional vector space over the field F and let V be an m-dimensional vector space over F. For each pair of ordered bases B, B' for U and V respectively, the function which assigns to a linear transformation T its matrix relative to B, B' is an isomorphism between the space $L(U, V)$ and the space of all $m \times n$ matrices over the field F.*

Proof: Let $B = \{\alpha_1, \dots, \alpha_n\}$ and $B' = \{\beta_1, \dots, \beta_m\}$.

Let M be the vector space of all $m \times n$ matrices over the field F. Let
$$\psi : L(U, V) \to M \text{ such that}$$
$$\psi(T) = [T ; B ; B'] \ \forall\, T \in L(U, V).$$

Let $T_1, T_2 \in L(U, V)$; and let
$$[T_1 ; B ; B'] = [a_{ij}]_{m \times n} \text{ and } [T_2 ; B ; B'] = [b_{ij}]_{m \times n}.$$

Then $$T_1(\alpha_j) = \sum_{i=1}^{m} a_{ij}\, \beta_i,\ j = 1, 2, \dots, n$$

and $$T_2(\alpha_j) = \sum_{i=1}^{m} b_{ij}\, \beta_i,\ j = 1, 2, \dots, n.$$

To prove that ψ **is one-one:**

We have $\psi(T_1) = \psi(T_2)$

$\Rightarrow$ $[T_1 ; B ; B'] = [T_2 ; B ; B']$ [By def. of ψ]

$\Rightarrow$ $[a_{ij}]_{m \times n} = [b_{ij}]_{m \times n}$

$\Rightarrow$ $a_{ij} = b_{ij}$ for $i = 1, \dots, m$ and $j = 1, \dots, n$

$\Rightarrow \quad \sum_{i=1}^{m} a_{ij} \beta_i = \sum_{i=1}^{m} b_{ij} \beta_i$ for $j = 1, \ldots, n$

$\Rightarrow \quad T_1(\alpha_j) = T_2(\alpha_j)$ for $j = 1, \ldots, n$

$\Rightarrow \quad T_1 = T_2$ $\quad$ [$\because$ T_1 and T_2 agree on a basis for U]

$\therefore$ ψ is one-one.

ψ is onto:

Let $[c_{ij}]_{m \times n} \in M$. Then there exists a linear transformation T from U into V such that

$$T(\alpha_j) = \sum_{i=1}^{m} c_{ij} \beta_i, \; j = 1, 2, \ldots, n.$$

We have $\quad [T \,;\, B \,;\, B'] = [c_{ij}]_{m \times n} \quad \Rightarrow \quad \boldsymbol{\psi(T) = [c_{ij}]_{m \times n}}$.

$\therefore$ ψ is onto.

ψ is a linear transformation:

If $a, b \in F$, then

$$\psi(aT_1 + bT_2) = [aT_1 + bT_2 \,;\, B \,;\, B'] \quad \text{[By def. of } \psi]$$
$$= [aT_1 \,;\, B \,;\, B'] + [bT_2 \,;\, B \,;\, B'] \quad \text{[By theorem 4]}$$
$$= a[T_1 \,;\, B \,;\, B'] + b[T_2 \,;\, B \,;\, B'] \quad \text{[By theorem 4]}$$
$$= a\psi(T_1) + b\psi(T_2), \text{ by def. of } \psi.$$

$\therefore$ ψ is a linear transformation.

Hence ψ is an isomorphism from $L(U, V)$ onto M.

Note: It should be noted that in the above theorem if $U \to V$, then ψ also preserves products and I *i.e.*,

$$\psi(T_1 T_2) = \psi(T_1)\, \psi(T_2)$$

and $\quad \psi(I) = I$ *i.e.*, unit matrix.

Theorem 7: *Let T be a linear operator on an n-dimensional vector space V and let B be an ordered basis for V. Prove that T is invertible iff $[T]_B$ is an invertible matrix. Also if T is invertible, then*

$$[T^{-1}]_B = [T]_B^{-1},$$

i.e., the matrix of T^{-1} relative to B is the inverse of the matrix of T relative to B.

Proof: Let T be invertible. Then T^{-1} exists and we have

$$T^{-1} T = I = T T^{-1}$$

$\Rightarrow \quad [T^{-1} T]_B = [I]_B = [T T^{-1}]_B$

$\Rightarrow \quad [T^{-1}]_B [T]_B = I = [T]_B [T^{-1}]_B$

$\Rightarrow \quad [T]_B$ is invertible and $([T]_B)^{-1} = [T^{-1}]_B$.

Conversely, let $[T]_B$ be an invertible matrix. Let $[T]_B = A$. Let $C = A^{-1}$ and let S be the linear transformation of V such that

$$[S]_B = C.$$

We have $\quad C A = I = AC$

$\Rightarrow \quad [S]_B [T]_B = I = [T]_B [S]_B$

$\Rightarrow \quad [ST]_B = [I]_B = [TS]_B$

$\Rightarrow \quad ST = I = TS$

$\Rightarrow \quad T$ is invertible.

3 Change of Basis

Suppose V is an n-dimensional vector space over the field F. Let B and B' be two ordered bases for V. If α is any vector in V, then we are now interested to know what is the relation between its coordinates with respect to B and its coordinates with respect to B'. **(Garhwal 2010)**

Theorem 1: *Let $V(F)$ be an n-dimensional vector space and let B and B' be two ordered bases for V. Then there is a unique necessarily invertible, $n \times n$ matrix A with entries in F such that*

(1) $[\alpha]_B = A [\alpha]_{B'}$ **(Garhwal 2008)**

(2) $[\alpha]_{B'} = A^{-1} [\alpha]_B$ *for every vector α in V.*

Proof: Let $B = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ and $B' = \{\beta_1, \beta_2, \ldots, \beta_n\}$.

Then there exists a unique linear transformation T from V into V such that

$$T(\alpha_j) = \beta_j, \; j = 1, 2, \ldots, n. \qquad \ldots(1)$$

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 elements in F. Also this matrix will be invertible because T is invertible.

Let $\quad [T]_B = A = [a_{ij}]_{n \times n}$.

Then $\quad T(\alpha_j) = \sum_{i=1}^{n} a_{ij} \alpha_i, j = 1, 2, \ldots, n. \qquad \ldots(2)$

Let $x_1, x_2, \ldots, x_n$ be the coordinates of α with respect to B and $y_1, y_2, \ldots, y_n$ be the coordinates of α with respect to B'. Then

$$\alpha = y_1 \beta_1 + y_2 \beta_2 + \ldots + y_n \beta_n = \sum_{j=1}^{n} y_j \beta_j$$

$$= \sum_{j=1}^{n} y_j T(\alpha_j) \qquad \text{[From (1)]}$$

$$= \sum_{j=1}^{n} y_j \sum_{i=1}^{n} a_{ij} \alpha_i \qquad \text{[From (2)]}$$

$$= \sum_{i=1}^{n} \left(\sum_{j=1}^{n} a_{ij} y_j \right) \alpha_i .$$

Also $\quad \alpha = \sum_{i=1}^{n} x_i \alpha_i . \qquad \therefore \quad x_i = \sum_{j=1}^{n} a_{ij} y_j$

because the expression for α as 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 $[\alpha]_B = x_i = \sum_{j=1}^{n} a_{ij}\, y_j = i^{\text{th}}$ entry of $A\,[\alpha]_{B'}$.

$$\therefore \quad [\alpha]_B = A\,[\alpha]_{B'} \quad \Rightarrow \quad A^{-1}\,[\alpha]_B = A^{-1}\,A\,[\alpha]_{B'}$$

$$\Rightarrow \quad A^{-1}\,[\alpha]_B = I\,[\alpha]_B \quad \Rightarrow \quad A^{-1}\,[\alpha]_B = [\alpha]_{B'}.$$

Note: The matrix $A = [T\,]_B$ is called the transition matrix from B to $B\,'$. It expresses the coordinates of each vector in V relative to B in terms of its coordinates relative to $B\,'$.

How to write the transition matrix from one basis to another ?

Let $B = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ and $B' = \{\beta_1, \beta_2, \ldots, \beta_n\}$ 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 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

$$\begin{aligned}
\beta_1 &= a_{11}\,\alpha_1 + a_{21}\,\alpha_2 + \ldots + a_{n1}\,\alpha_n \\
\beta_2 &= a_{12}\,\alpha_1 + a_{22}\,\alpha_2 + \ldots + a_{n2}\,\alpha_n \\
&\ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \\
&\ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \\
\beta_n &= a_{1n}\,\alpha_1 + a_{2n}\,\alpha_2 + \ldots + a_{nn}\,\alpha_n .
\end{aligned}$$

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

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix}.$$

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

$$[\alpha]_B = A\,[\alpha]_{B'} \text{ and } [\alpha]_{B'} = A^{-1}\,[\alpha]_B .$$

Theorem 2: *Let $B = \{\alpha_1, \ldots, \alpha_n\}$ and $B' = \{\beta_1, \ldots, \beta_n\}$ be two ordered bases for an n-dimensional vector space $V\,(F\,)$. If $(x_1, \ldots, x_n)$ 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

$$T\,(\alpha_i) = \beta_i\,,\ i = 1, 2, \ldots, n.$$

Proof: We have $T\,(\alpha) = T\left(\sum_{i=1}^{n} x_i\,\alpha_i\right) = \sum_{i=1}^{n} x_i\,T\,(\alpha_i)$ [$\because$ T is linear]

$$= \sum_{i=1}^{n} x_i\,\beta_i = \beta.$$

Similarity:

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 exists an $n \times n$ invertible square matrix C with elements in F such that*

$$B = C^{-1} AC.$$

(Gorakhpur 2010)

Theorem 3: *The relation of similarity is an equivalence relation in the set of all $n \times n$ matrices over the field F.* **(Kumaun 2010, 11)**

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.

Symmetric: Let A be similar to B. Then there exists an $n \times n$ invertible matrix P over F such that

$$A = P^{-1} BP$$

$\Rightarrow$ $PAP^{-1} = P (P^{-1} BP) P^{-1}$

$\Rightarrow$ $PAP^{-1} = B \quad \Rightarrow \quad B = PAP^{-1}$

$\Rightarrow$ $B = (P^{-1})^{-1} AP^{-1}$

[$\because$ 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

$$A = P^{-1} BP \quad \text{and} \quad B = Q^{-1} CQ,$$

where P and Q are invertible $n \times n$ matrices over F.

We have $A = P^{-1} BP = P^{-1} (Q^{-1} CQ) P$

$= (P^{-1} Q^{-1}) C (QP)$

$= (QP)^{-1} C (QP)$ [$\because$ 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 4: *Similar matrices have the same determinant.*

Proof: Let B be similar to A. Then there exists an invertible matrix C such that

$$B = C^{-1} AC$$

$\Rightarrow$ $\det B = \det (C^{-1} AC) \Rightarrow \det B = (\det C^{-1}) (\det A) (\det C)$

$\Rightarrow$ $\det B = (\det C^{-1}) (\det C) (\det A)$

$\Rightarrow$ $\det B = (\det C^{-1} C) (\det A)$

$\Rightarrow$ $\det B = (\det I) (\det A) \Rightarrow \det B = 1 (\det A) \Rightarrow \det B = \det A.$

Similarity of linear transformations: Definition: *Let A and B be linear transformations on a 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*

$$B = CAC^{-1}.$$

Theorem 5: *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 transformations 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. We can write

$$A = IAI^{-1},$$

where I is identity transformation on V.

$\therefore$ A is similar to A because I is definitely invertible.

Symmetric: Let A be similar to B. Then there exists 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.

Transitive: Let A be similar to B and B be similar to C.

Then $A = PBP^{-1}$,

and $B = QCQ^{-1}$,

where P and Q are invertible linear transformations on V.

We have $A = PBP^{-1} = P(QCQ^{-1})P^{-1}$

$$= (PQ)C(Q^{-1}P^{-1}) = (PQ)C(PQ)^{-1}.$$

$\therefore$ A is similar to C.

Hence similarity is an equivalence relation on the set of all linear transformations on $V(F)$.

Theorem 6: *Let T be a linear operator on an n-dimensional vector space V (F) and let B and B′ be two ordered bases for V. Then the matrix of T relative to B′ is similar to the matrix of T relative to B.*

Proof: Let $B = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ and $B' = \{\beta_1, \ldots, \beta_n\}$.

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 $T(\alpha_j) = \sum_{i=1}^{n} a_{ij}\alpha_i, \; j = 1, 2, \ldots, n$ ...(1)

and $$T(\beta_j) = \sum_{i=1}^{n} c_{ij}\,\beta_i,\ j = 1, 2, \ldots, n. \qquad \ldots(2)$$

Let S be the linear operator on V defined by

$$S(\alpha_j) = \beta_j,\ j = 1, 2, \ldots, n. \qquad \ldots(3)$$

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

$$S(\alpha_j) = \sum_{i=1}^{n} p_{ij}\,\alpha_i,\ j = 1, 2, \ldots, n \qquad \ldots(4)$$

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]

$\Rightarrow$ $P^{-1}AP = IC \Rightarrow P^{-1}AP = C$

$\Rightarrow$ C is similar to A.

Note: Suppose B and B' are two ordered bases 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 relative to B'. If P is the transition matrix from the basis B to the basis B', then $C = P^{-1}AP$.

This result will enable us to find the matrix of T relative to the basis B' when we already knew the matrix of T relative to the basis B.

Theorem 7: *Let V be an n-dimensional vector space over the field F and T_1 and T_2 be two linear operators on V. If there exist two ordered bases B and B' for V such that $[T_1]_B = [T_2]_{B'}$, then show that T_2 is similar to T_1.*

Proof: Let $B = \{\alpha_1, \ldots, \alpha_n\}$ and $B' = \{\beta_1, \ldots, \beta_n\}$.

Let $[T_1]_B = [T_2]_{B'} = A = [a_{ij}]_{n \times n}$.

Then $T_1(\alpha_j) = \sum_{i=1}^{n} a_{ij}\,\alpha_i,\ j = 1, 2, \ldots, n,$ ...(1)

and $T_2(\beta_j) = \sum_{i=1}^{n} a_{ij}\,\beta_i,\ j = 1, 2, \ldots, n.$ ...(2)

Let S be the linear operator on V defined by

$S(\alpha_j) = \beta_j, j = 1, 2, \ldots, n.$ ...(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)]

$= (T_2 S)(\alpha_j).$ ...(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

$(T_2 S)(\alpha_j) = (ST_1)(\alpha_j), j = 1, 2, \ldots, n.$

Since $T_2 S$ and ST_1 agree on a basis for V, therefore we have

$T_2 S = ST_1$

$\Rightarrow$ $T_2 SS^{-1} = ST_1 S^{-1}$ $\Rightarrow$ $T_2 I = ST_1 S^{-1}$

$\Rightarrow$ $T_2 = ST_1 S^{-1}$ $\Rightarrow$ T_2 is similar to T_1.

Determinant of a linear transformation on a finite dimensional vector space: Let T be a linear operator on an n-dimensional vector space $V(F)$. If B and B' are two ordered bases for V, then $[T]_B$ and $[T]_{B'}$ are similar matrices. Also similar matrices have the same determinant. This enables us to make the following definition :

Definition: *Let T be a linear operator on an n-dimensional vector space $V(F)$. Then the determinant of T is the determinant of the matrix of T relative to any ordered basis for V.*

By the above discussion the determinant of T as defined by us will be a unique element of F and thus our definition is sensible.

Scalar Transformation: Definition: *Let $V(F)$ be a vector space. A linear transformation T on V is said to be a scalar transformation of V if*

$$T(\alpha) = c\alpha \ \forall \ \alpha \in V,$$

where c is a fixed scalar in F.

Also then we write $T = c$ and we say that the linear transformation T is equal to the scalar c.

Also obviously if the linear transformation T is equal to the scalar c, then we have $T = cI$, where I is the identity transformation on V.

Trace of a Matrix: Definition: *Let A be a square matrix of order n over a field F. The sum of the elements of A lying along the principal diagonal is called the trace of A. We shall write the trace of A as* **trace A.** Thus if $A = [a_{ij}]_{n \times n}$, then

$$\text{tr } A = \sum_{i=1}^{n} a_{ii} = a_{11} + a_{22} + \ldots + a_{nn}.$$

(Kumaun 2008)

In the following two theorems we have given some fundamental properties of the trace function.

Theorem 8: *Let A and B be two square matrices of order n over a field F and $\lambda \in F$. Then*

(i) $tr(\lambda A) = \lambda \, tr \, A$; **(Kumaun 2008)**

(ii) $tr(A + B) = tr \, A + tr \, B$;

(iii) $tr(AB) = tr(BA)$.

Proof: Let $A = [a_{ij}]_{n \times n}$ and $B = [b_{ij}]_{n \times n}$.

(i) We have $\lambda A = [\lambda a_{ij}]_{n \times n}$, by def. of multiplication of a matrix by a scalar.

$$\therefore \quad \text{tr}(\lambda A) = \sum_{i=1}^{n} \lambda a_{ii} = \lambda \sum_{i=1}^{n} a_{ii} = \lambda \text{ tr } A.$$

(ii) We have $A + B = [a_{ij} + b_{ij}]_{n \times n}$.

$$\therefore \quad \text{tr}(A + B) = \sum_{i=1}^{n} (a_{ii} + b_{ii}) = \sum_{i=1}^{n} a_{ii} + \sum_{i=1}^{n} b_{ii} = \text{tr } A + \text{tr } B.$$

(iii) 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
$$\text{tr}(AB) = \sum_{i=1}^{n} c_{ii} = \sum_{i=1}^{n} \left(\sum_{k=1}^{n} a_{ik} b_{ki} \right)$$

$$= \sum_{k=1}^{n} \sum_{i=1}^{n} a_{ik} b_{ki},$$

[Interchanging the order of summation in the last sum]

$$= \sum_{k=1}^{n} \left(\sum_{i=1}^{n} b_{ki} \, a_{ik} \right) = \sum_{k=1}^{n} d_{kk}$$

$$= d_{11} + d_{22} + \ldots + d_{nn} = \text{tr } (BA).$$

Theorem 9: *Similar matrices have the same trace.*

Proof: Suppose A and B are two similar matrices. Then there exists an invertible matrix C such that $B = C^{-1} \, AC$.

Let $\quad C^{-1} \, A = D.$

Then $\quad \text{tr } B = \text{tr } (DC)$

$= \text{tr } (CD)$ [By theorem 8]

$= \text{tr } (CC^{-1} \, A) = \text{tr } (IA) = \text{tr } A.$

Trace of a linear transformation on a finite dimensional vector space: Let T be a linear operator on an n-dimensional vector space $V(F)$. If B and B' are two ordered bases for V, then $[T]_B$ and $[T]_{B'}$

are similar matrices. Also similar matrices have the same trace. This enables us to make the following definition.

Definition of trace of a linear transformation. *Let T be a linear operator on an n-dimensional vector space $V(F)$. Then the trace of T is the trace of the matrix of T relative to any ordered basis for V.*

By the above discussion the trace of T as defined by us will be a unique element of F and thus our definition is sensible.

Illustrative Examples

Example 3: *Find the matrix of the linear transformation T on $V_3(\mathbf{R})$ defined as $T(a, b, c) = (2b + c, a - 4b, 3a)$, with respect to the ordered basis B and also with respect to the ordered basis B' where*

(i) $B = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$

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

Solution: (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 $\quad T(0, 0, 1) = (1, 0, 0) = 1\,(1, 0, 0) + 0\,(0, 1, 0) + 0\,(0, 0, 1).$

$\therefore$ by def. of matrix 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 ordered 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 co-ordinates of the vectors $T\ (1, 0, 0), T\ (0, 1, 0)$ and $T\ (0, 0, 1)$ respectively constitute the first, second and third columns of the matrix $[T\ ]_B$.

(ii) We have $T\ (1, 1, 1) = (3, -3, 3)$.

Now our aim is 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., $$x = c, y = b - c, z = a - b. \qquad ...(1)$$

Putting $a = 3, b = -3$, and $c = 3$ in (1), we get
$$x = 3,\ y = -6 \text{ and } z = 6.$$

$\therefore$ $$T\ (1, 1, 1) = (3, -3, 3) = 3\ (1, 1, 1) - 6\ (1, 1, 0) + 6\ (1, 0, 0).$$

Also $$T\ (1, 1, 0) = (2, -3, 3).$$

Putting $a = 2, b = -3$ and $c = 3$ in (1), we get
$$T\ (1, 1, 0) = (2, -3, 3) = 3\ (1, 1, 1) - 6\ (1, 1, 0) + 5\ (1, 0, 0).$$

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 4: *Let T be the linear operator on* $\mathbf{R}^3$ *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*
$$\alpha_1 = (1, 0, 1), \alpha_2 = (-1, 2, 1) \ \textit{and}\ \alpha_3 = (2, 1, 1)?$$

Solution: By def. 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 have
$$x = \frac{-a - 3b + 5c}{4},\ \ y = \frac{b + c - a}{4},\ \ z = \frac{b - c + a}{2}. \qquad ...(1)$$

Putting $a = 4, b = -2, c = 3$ in (1), we get

$$x = \frac{17}{4}, y = -\frac{3}{4}, z = -\frac{1}{2}.$$

$\therefore$ $$T(\alpha_1) = \frac{17}{4}\alpha_1 - \frac{3}{4}\alpha_2 - \frac{1}{2}\alpha_3.$$

Also $$T(\alpha_2) = T(-1, 2, 1) = (-2, 4, 9).$$

Putting $a = -2, b = 4, c = 9$ in (1), we get

$$x = \frac{35}{4}, y = \frac{15}{4}, z = \frac{-7}{2}.$$

$\therefore$ $$T(\alpha_2) = \frac{35}{4}\alpha_1 + \frac{15}{4}\alpha_2 - \frac{7}{2}\alpha_3.$$

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 5: *Let T be a linear operator on $\mathbf{R}^3$ 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 a formula for T^{-1}.

Solution: Suppose B is the standard ordered basis for $\mathbf{R}^3$. 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

$$T(1, 0, 0) = (3, -2, -1), T(0, 1, 0) = (0, 1, 2) \text{ and } T(0, 0, 1) = (1, 0, 4).$$

$\therefore$ $$A = [T]_B = \begin{bmatrix} 3 & 0 & 1 \\ -2 & 1 & 0 \\ -1 & 2 & 4 \end{bmatrix}.$$

Now T will be invertible if the matrix $[T]_B$ is invertible. [See theorem 7 of article 2]

We have $$\det A = |A| = \begin{vmatrix} 3 & 0 & 1 \\ -2 & 1 & 0 \\ -1 & 2 & 4 \end{vmatrix}$$

$$= 3(4 - 0) + 1(-4 + 1) = 9.$$

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

$$\begin{vmatrix} 1 & 0 \\ 2 & 4 \end{vmatrix}, -\begin{vmatrix} -2 & 0 \\ -1 & 4 \end{vmatrix}, \begin{vmatrix} -2 & 1 \\ -1 & 2 \end{vmatrix} \text{ i.e., } 4, 8, -3.$$

The cofactors of the elements of the second row of A are

$$-\begin{vmatrix} 0 & 1 \\ 2 & 4 \end{vmatrix}, \begin{vmatrix} 3 & 1 \\ -1 & 4 \end{vmatrix}, -\begin{vmatrix} 3 & 0 \\ -1 & 2 \end{vmatrix} \text{ i.e., } 2, 13, -6.$$

The cofactors of the elements of the third row of A are

$$\begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix}, -\begin{vmatrix} 3 & 1 \\ -2 & 0 \end{vmatrix}, \begin{vmatrix} 3 & 0 \\ -2 & 1 \end{vmatrix} \text{ i.e., } -1, -2, 3.$$

$$\therefore \qquad \text{Adj. } A = \text{transpose of the matrix} \begin{bmatrix} 4 & 8 & -3 \\ 2 & 13 & -6 \\ -1 & -2 & 3 \end{bmatrix}$$

$$= \begin{bmatrix} 4 & 2 & -1 \\ 8 & 13 & -2 \\ -3 & -6 & 3 \end{bmatrix}.$$

$$\therefore \qquad A^{-1} = \frac{1}{\det A} \text{Adj. } A = \frac{1}{9} \begin{bmatrix} 4 & 2 & -1 \\ 8 & 13 & -2 \\ -3 & -6 & 3 \end{bmatrix}.$$

Now $\qquad [T^{-1}]_B = ([T]_B)^{-1} = A^{-1}.$ [See theorem 7 of article 2]

We shall now find a formula for T^{-1}. Let $\alpha = (a, b, c)$ be any vector belonging to $\mathbf{R}^3$. Then

$$[T^{-1}(\alpha)]_B = [T^{-1}]_B \, [\alpha]_B \qquad \text{[See Note of Theorem 2, article 2]}$$

$$= \frac{1}{9} \begin{bmatrix} 4 & 2 & -1 \\ 8 & 13 & -2 \\ -3 & -6 & 3 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

$$= \frac{1}{9} \begin{bmatrix} 4a + 2b - c \\ 8a + 13b - 2c \\ -3a - 6b + 3c \end{bmatrix}$$

Since B is the standard ordered basis for $\mathbf{R}^3$,

$$\therefore \qquad T^{-1}(\alpha) = T^{-1}(a, b, c)$$

$$= \frac{1}{9}(4a + 2b - c, 8a + 13b - 2c, -3a - 6b + 3c).$$

Example 6: *Let T be the linear operator on $\mathbf{R}^3$ defined by*

$$T(x_1, x_2, x_3) = (3x_1 + x_3, -2x_1 + x_2, -x_1 + 2x_2 + 4x_3).$$

(i) What is the matrix of T in the standard ordered basis B for $\mathbf{R}^3$? **(Kumaun 2014)**

(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

$$A = \begin{bmatrix} 3 & 0 & 1 \\ -2 & 1 & 0 \\ -1 & 2 & 4 \end{bmatrix}.$$ [For calculation work see Example 5]

(ii) Since B is the standard ordered basis, therefore the transition matrix P from B to B' can be immediately written as

$$P = \begin{bmatrix} 1 & -1 & 2 \\ 0 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$$

Now $\quad [T]_{B'} = P^{-1} [T]_B \, P.$ [See note of theorem 6, article 3]

In order to compute the matrix P^{-1}, we find that det $P = -4$.

Therefore $$P^{-1} = \frac{1}{\det P} \text{Adj. } P = -\frac{1}{4} \begin{bmatrix} 1 & 3 & -5 \\ 1 & -1 & -1 \\ -2 & -2 & 2 \end{bmatrix}.$$

$\therefore$ $$[T]_{B'} = -\frac{1}{4} \begin{bmatrix} 1 & 3 & -5 \\ 1 & -1 & -1 \\ -2 & -2 & 2 \end{bmatrix} \begin{bmatrix} 3 & 0 & 1 \\ -2 & 1 & 0 \\ -1 & 2 & 4 \end{bmatrix} \begin{bmatrix} 1 & -1 & 2 \\ 0 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

$$= -\frac{1}{4} \begin{bmatrix} 2 & -7 & -19 \\ 6 & -3 & -3 \\ -4 & 2 & 6 \end{bmatrix} \begin{bmatrix} 1 & -1 & 2 \\ 0 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

$$= -\frac{1}{4} \begin{bmatrix} -17 & -35 & -22 \\ 3 & -15 & 6 \\ 2 & 14 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} \frac{17}{4} & \frac{35}{4} & \frac{11}{2} \\ -\frac{3}{4} & \frac{15}{4} & -\frac{3}{2} \\ -\frac{1}{2} & -\frac{7}{2} & 0 \end{bmatrix}.$$

[Note that this result tallies with that of Example 4.]

Example 7: *Let T be a linear operator on $\mathbf{R}^2$ defined by : $T(x, y) = (2y, 3x - y)$. Find the matrix representation of T relative to the basis $\{(1, 3), (2, 5)\}$.*

(Kumaun 2011, 15)

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

$$p = -5a + 2b, q = 3a - b. \qquad \ldots(1)$$

Putting $a = 6, b = 0$ in (1), we get $p = -30, q = 18$.

$$\therefore \qquad T(\alpha_1) = (6, 0) = -30\alpha_1 + 18\alpha_2. \qquad \ldots(2)$$

Again putting $a = 10, b = 1$ in (1), we get

$$p = -48, q = 29.$$

$$\therefore \qquad T(\alpha_2) = (10, 1) = -48\alpha_1 + 29\alpha_2. \qquad \ldots(3)$$

From the relations (2) and (3), we see that the matrix of T relative to the basis

$$\{\alpha_1, \alpha_2\} \text{ is } = \begin{bmatrix} -30 & -48 \\ 18 & 29 \end{bmatrix}.$$

Example 8: *Consider the vector space $V(\mathbf{R})$ of all 2×2 matrices over the field $\mathbf{R}$ 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*

$$\alpha_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \alpha_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \alpha_3 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \alpha_4 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

Solution: We have

$$T(\alpha_1) = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$$

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

$$T(\alpha_2) = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$$

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

$$T(\alpha_3) = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$$

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

and

$$T(\alpha_4) = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$$

$$= 0\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + 1\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + 0\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + 1\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

$$\therefore \qquad [T]_B = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}.$$

Example 9: *If the matrix of a linear transformation T on $V_2(\mathbf{C})$, 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 $B' = \{(1,1),(1,-1)\}$?*

Solution: Let us first define T explicitly. It is given that

$$[T]_B = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}.$$

$\therefore \qquad T(1,0) = 1(1,0) + 1(0,1) = (1,1),$

and $\qquad T(0,1) = 1(1,0) + 1(0,1) = (1,1).$

If $(a,b) \in V_2(\mathbf{C})$, then we can write $(a,b) = a(1,0) + b(0,1)$.

$$\therefore \qquad T(a,b) = aT(1,0) + bT(0,1)$$
$$= a(1,1) + b(1,1)$$
$$= (a+b, a+b).$$

This is the explicit expression for T.

Now let us find the matrix of T with respect to B'.

We have $\qquad T(1,1) = (2,2).$

Let $\qquad (2,2) = x(1,1) + y(1,-1) = (x+y, x-y).$

Then $\qquad x + y = 2, x - y = 2$

$\Rightarrow \qquad x = 2, y = 0.$

$\therefore \qquad (2,2) = 2(1,1) + 0(1,-1).$

Also $\qquad T(1,-1) = (0,0)$
$$= 0(1,1) + 0(1,-1).$$

$$\therefore \qquad [T]_{B'} = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}.$$

Note: If P is the transition matrix from the basis B to the basis B', then

$$P = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}.$$

We can compute $[T]_{B'}$ by using the formula

$$[T]_{B'} = P^{-1}[T]_B P.$$

Example 10: *Show that the vectors $\alpha_1 = (1,0,-1), \alpha_2 = (1,2,1), \alpha_3 = (0,-3,2)$ form a basis for $\mathbf{R}^3$. Express each of the standard basis vectors as a linear combination of $\alpha_1, \alpha_2, \alpha_3$.*

(Kumaun 2015)

Solution: Let a, b, c be scalars *i.e.*, real numbers such that

$$a\alpha_1 + b\alpha_2 + c\alpha_3 = \mathbf{0}$$

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., $$\left.\begin{aligned} a + b + 0c &= 0,\\ 0a + 2b - 3c &= 0,\\ -a + b + 2c &= 0.\end{aligned}\right\} \qquad \ldots(1)$$

The coefficient matrix A of these equations 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 equations (1). Therefore the vectors $\alpha_1, \alpha_2, \alpha_3$ are linearly independent over $\mathbf{R}$. Since $\dim \mathbf{R}^3 = 3$, therefore the set $\{\alpha_1, \alpha_2, \alpha_3\}$ containing three linearly independent vectors forms a basis for $\mathbf{R}^3$.

Now let $B = \{e_1, e_2, e_3\}$ be the standard ordered basis for $\mathbf{R}^3$.

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
$$\alpha_1 = (1,0,-1) = 1\,e_1 + 0e_2 - 1e_3$$
$$\alpha_2 = (1,2,1) = 1e_1 + 2e_2 + 1e_3$$
$$\alpha_3 = (0,-3,2) = 0e_1 - 3e_2 + 2e_3.$$

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

$$\begin{bmatrix} 2 & -3 \\ 1 & 2 \end{bmatrix}, -\begin{bmatrix} 0 & -3 \\ -1 & 2 \end{bmatrix}, \begin{bmatrix} 0 & 2 \\ -1 & 1 \end{bmatrix}, \text{ i.e., } 7, 3, 2.$$

The cofactors of the elements of the second row of P are

$$-\begin{bmatrix} 1 & 0 \\ 1 & 2 \end{bmatrix}, \begin{vmatrix} 1 & 0 \\ -1 & 2 \end{vmatrix}, -\begin{vmatrix} 1 & 1 \\ -1 & 1 \end{vmatrix} \text{ i.e., } -2, 2, -2.$$

The cofactors of the elements of the third row of P are

$$\begin{vmatrix} 1 & 0 \\ 2 & -3 \end{vmatrix}, -\begin{vmatrix} 1 & 0 \\ 0 & -3 \end{vmatrix}, \begin{vmatrix} 1 & 1 \\ 0 & 2 \end{vmatrix}, \text{ i.e., } -3, 3, 2$$

$\therefore$ $$\text{Adj } P = \text{transpose of the matrix} \begin{bmatrix} 7 & 3 & 2 \\ -2 & 2 & -2 \\ -3 & 3 & 2 \end{bmatrix} = \begin{bmatrix} 7 & -2 & -3 \\ 3 & 2 & 3 \\ 2 & -2 & 2 \end{bmatrix}.$$

$\therefore$ $$P^{-1} = \frac{1}{\det P} \text{Adj } P = \frac{1}{10} \begin{bmatrix} 7 & -2 & -3 \\ 3 & 2 & 3 \\ 2 & -2 & 2 \end{bmatrix}.$$

Now $e_1 = 1 e_1 + 0 e_2 + 0 e_3$.

$\therefore$ Coordinate matrix of e_1 relative to the basis B

$$= \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}.$$

$\therefore$ Co-ordinate 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 $$[e_2]_B = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \text{ and } [e_3]_B = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

$\therefore$ $$[e_2]_{B'} = P^{-1} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \text{ and } [e_3]_{B'} = P^{-1} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Thus $$[e_2]_{B'} = \frac{1}{10} \begin{bmatrix} -2 \\ 2 \\ -2 \end{bmatrix}, [e_3]_{B'} = \frac{1}{10} \begin{bmatrix} -3 \\ 3 \\ 2 \end{bmatrix}.$$

$\therefore \qquad e_2 = -\frac{2}{10}\alpha_1 + \frac{2}{10}\alpha_2 - \frac{2}{10}\alpha_3$

and $\qquad e_3 = -\frac{3}{10}\alpha_1 + \frac{3}{10}\alpha_2 + \frac{2}{10}\alpha_3.$

Example 11: *Let A be an $m \times n$ matrix with real entries. Prove that $A = 0$ (null matrix) if and only if trace $(A^t A) = 0$.*

Solution: Let $A = [a_{ij}]_{m \times n}$.

Then $A^t = [b_{ij}]_{n \times m}$, where $b_{ij} = a_{ji}$.

Now $A^t A$ is a matrix of the type $n \times n$.

Let $A^t A = [c_{ij}]_{n \times n}$. Then

c_{ii} = the sum of the products of the corresponding elements of the i^{th} row of A^t and the i^{th} column of A

$$= b_{i1}\, a_{1i} + b_{i2}\, a_{2i} + \ldots + b_{im}\, a_{mi}$$

$$= a_{1i}\, a_{1i} + a_{2i}\, a_{2i} + \ldots + a_{mi}\, a_{mi} \qquad [\because \; b_{ij} = a_{ji}]$$

$$= a_{1i}^2 + a_{2i}^2 + \ldots + a_{mi}^2.$$

Now $\qquad \text{trace}\,(A^t A) = \sum_{i=1}^{n} c_{ii} = \sum_{i=1}^{n} \{a_{1i}^2 + a_{2i}^2 + \ldots + a_{mi}^2\}$

= the sum of the squares of all the elements of A.

Now the elements of A are all real numbers.

Therefore $\qquad \text{trace}\,(A^t A) = 0$

$\Rightarrow$ the sum of the squares of all the elements of A is zero

$\Rightarrow$ each element of A is zero

$\Rightarrow$ A is a null matrix.

Conversely if A is a null matrix, then $A^t A$ is also a null matrix and so trace $(A^t A) = 0$.

Hence trace $(A^t A) = 0$ iff $A = 0$.

Example 12: *Show that the only matrix similar to the identity matrix I is I itself.*

Solution: The identity matrix I is invertible and we can write $I = I^{-1} I I$. Therefore I is similar to I. Further let B be a matrix similar to I. Then there exists an invertible matrix P such that

$$B = P^{-1} I P$$

$\Rightarrow \qquad B = P^{-1} P \qquad [\because \; P^{-1} I = P^{-1}]$

$\Rightarrow \qquad B = I.$

Hence the only matrix similar to I is I itself.

Example 13: *If T and S are similar linear transformations on a finite dimensional vector space $V(F)$, then det T = det S.*

Solution: Since T and S are similar, therefore there exists an invertible linear transformation P on V such that $T = PSP^{-1}$.

Therefore $\det T = \det (PSP^{-1}) = (\det P)(\det S)(\det P^{-1})$

$$= (\det P)(\det P^{-1})(\det S)$$

$$= [\det (PP^{-1})](\det S)$$

$$= (\det I)(\det S)$$

$$= 1 (\det S) = \det S.$$

Comprehensive Exercise 1

1. Let $V = \mathbf{R}^3$ and $T : V \to V$ be a linear mapping defined by

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

What is matrix of T relative to the basis $B = \{(1, 0, 1), (-1, 1, 1), (0, 1, 1)\}$?

2. Find matrix representation of linear mapping $T : \mathbf{R}^3 \to \mathbf{R}^3$ given by

$T(x, y, z) = (z, y + z, x + y + z)$ relative to the basis

$B = \{(1, 0, 1), (-1, 2, 1), (2, 1, 1)\}$. **(Gorakhpur 2010)**

3. Find the coordinates of the vector $(2, 1, 3, 4)$ of $\mathbf{R}^4$ relative to the basis vectors

$$\alpha_1 = (1, 1, 0, 0), \alpha_2 = (1, 0, 1, 1), \alpha_3 = (2, 0, 0, 2), \alpha_4 = (0, 0, 2, 2).$$

4. Let T be the linear operator on $\mathbf{R}^2$ defined by

$$T(x, y) = (4x - 2y, 2x + y).$$

Compute the matrix of T relative to the basis $\{\alpha_1, \alpha_2\}$ where $\alpha_1 = (1, 1)$, $\alpha_2 = (-1, 0)$. **(Kumaun 2008)**

5. Let T be the linear operator on $\mathbf{R}^2$ defined by

$$T(x, y) = (4x - 2y, 2x + y).$$

(i) What is the matrix of T in the standard ordered basis B for $\mathbf{R}^2$?

(ii) Find the transition matrix P from the ordered basis B to the ordered basis $B' = \{\alpha_1, \alpha_2\}$ where $\alpha_1 = (1, 1), \alpha_2 = (-1, 0)$. Hence find the matrix of T relative to the ordered basis B'.

6. Let T be the linear operator on $\mathbf{R}^2$ defined by

$$T(a, b) = (a, 0).$$

Write the matrix of T in the standard ordered basis $B = \{(1, 0), (0, 1)\}$.If $B' = \{(1, 1), (2, 1)\}$ is another ordered basis for $\mathbf{R}^3$, find the transition matrix P from the basis B to the basis B'. Hence find the matrix of T relative to the basis B'.

7. The matrix of a linear transformation T on V_3 (**C**) relative to the basis

$$B = \{(1,0,0),(0,1,0),(0,0,1)\} \text{ is}$$

$$\begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & -1 \\ -1 & -1 & 0 \end{bmatrix}.$$

What is the matrix of T relative to the basis

$$B' = \{(0,1,-1),(1,-1,1),(-1,0,1)\}?$$

8. 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 $\mathbf{R}^3$, of the linear transformation $T : \mathbf{R}^3 \to \mathbf{R}^3$ whose matrix relative to the standard ordered basis is

$$\begin{bmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 3 \end{bmatrix}.$$

9. Find the matrix representation of the linear mappings relative to the usual bases for $\mathbf{R}^n$.

(i) $T : \mathbf{R}^3 \to \mathbf{R}^2$ defined by $T(x, y, z) = (2x - 4y + 9z, 5x + 3y - 2z)$.

(ii) $T : \mathbf{R} \to \mathbf{R}^2$ defined by $T(x) = (3x, 5x)$.

(iii) $T : \mathbf{R}^3 \to \mathbf{R}^3$ defined by $T(x, y, z) = (x, y, 0)$.

(iv) $T : \mathbf{R}^3 \to \mathbf{R}^3$ defined by $T(x, y, z) = (z, y + z, x + y + z)$.

10. Let $B = \{(1,0),(0,1)\}$ and $B' = \{(1,2),(2,3)\}$ be any two bases of $\mathbf{R}^2$ and $T(x, y) = (2x - 3y, x + y)$.

(i) Find the transition matrices P and Q from B to B' and from B' to B respectively.

(ii) Verify that $[\alpha]_B = P[\alpha]_{B'}\ \forall\, \alpha \in \mathbf{R}^2$

(iii) Verify that $P^{-1}[T]_B P = [T]_{B'}$.

11. Let V be the space of all 2×2 matrices over the field F and let P be a fixed 2×2 matrix over F. Let T be the linear operator on V defined by

$$T(A) = PA, \forall\, A \in V.$$

Prove that trace $(T) = 2$ trace (P).

12. Let V be the space of 2×2 matrices over $\mathbf{R}$ and let $M = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$.

Let T be the linear operator on V defined by $T(A) = MA$. Find the trace of T.

13. Find the trace of the operator T on $\mathbf{R}^3$ defined by

$$T(x, y, z) = (a_1 x + a_2 y + a_3 z, b_1 x + b_2 y + b_3 z, c_1 x + c_2 y + c_3 z).$$

14. Show that the only matrix similar to the zero matrix is the zero matrix itself.

Answers 1

1. $\begin{bmatrix} 2 & 1 & 2 \\ 0 & 1 & 1 \\ -2 & 2 & 0 \end{bmatrix}$

2. $\begin{bmatrix} \frac{3}{2} & 0 & \frac{13}{4} \\ \frac{1}{2} & 1 & \frac{5}{4} \\ 0 & 1 & -\frac{1}{2} \end{bmatrix}$

3. $\left(1, 0, \frac{1}{2}, \frac{3}{2}\right)$

4. $\begin{bmatrix} 3 & -2 \\ 1 & 2 \end{bmatrix}$

5. (i) $\begin{bmatrix} 4 & -2 \\ 2 & 1 \end{bmatrix}$ (ii) $\begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix}; \begin{bmatrix} 3 & -2 \\ 1 & 2 \end{bmatrix}$

6. $[T]_B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$; $P = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}$; $[T]_{B'} = \begin{bmatrix} -1 & -2 \\ 1 & 2 \end{bmatrix}$

7. $[T]_{B'} = \begin{bmatrix} 1 & 0 & -4 \\ 0 & 0 & -2 \\ 0 & 0 & -3 \end{bmatrix}$

8. $\begin{bmatrix} 3 & -2/3 & -2/3 \\ -2/3 & 10/3 & 0 \\ -2/3 & 0 & 8/3 \end{bmatrix}$

9. (i) $\begin{bmatrix} 2 & -4 & 9 \\ 5 & 3 & -2 \end{bmatrix}$ (ii) $\begin{bmatrix} 3 \\ 5 \end{bmatrix}$

(iii) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ (iv) $\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$

10. (i) $P = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}; Q = \begin{bmatrix} -3 & 2 \\ 2 & -1 \end{bmatrix}$

12. trace $(T) = 10$

13. trace $(T) = a_1 + b_2 + c_3$

Objective Type Questions

Multiple Choice Questions

Indicate the correct answer for each question by writing the corresponding letter from (a), (b), (c) *and* (d).

1. Let T be a linear transformation on the vector space $V_2(F)$ defined by

$$T(a, b) = (a, 0).$$

The matrix of T relative to the ordered basis $\{(1, 0), (0, 1)\}$ of $V_2(F)$ is

(a) $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ (b) $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$

(c) $\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$ (d) $\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$.

2. The transition matrix P from the standard ordered basis to the ordered basis $\{(1, 1), (-1, 0)\}$ is

(a) $\begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix}$

(c) $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ (d) $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

3. Let V be the vector space of 2×2 matrices over $\mathbf{R}$ and let

$$M = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}.$$

Let T be the linear operator on V defined by

$$T(A) = MA.$$

Then the trace of T is

(a) 5 (b) 10

(c) 0 (d) none of these.

4. Let T be a linear operator on $\mathbf{R}^2$ defined by $T(a, b) = (b, a)$. Then the matrix of T with respect to basis $\{e_1 = (1, 0) \text{ and } e_2 = (0, 1)\}$:

(a) $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$

(c) $\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$ (d) $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

Fill in the Blank(s)

Fill in the blanks "......" so that the following statements are complete and correct.

1. If T is a linear operator on $\mathbf{R}^2$ defined by

$$T(x, y) = (x - y, y),$$

then $T^2(x, y) = \ldots\ldots$

2. Let A be a square matrix of order n over a field F. The sum of the elements of A lying along the principal diagonal is called the ...... of A.
3. Let T and S be similar linear operators on the finite dimensional vector space $V(F)$, then det (T) ...... det (S).
4. Let $V(F)$ be an n-dimensional vector space and B be any ordered basis for V. If I be the identity transformation on V, then $[I ; B] = \ldots\ldots$
5. Let T be a linear operator on an n-dimensional vector space $V(F)$ and let B and B' be two ordered bases for V. Then the matrix of T relative to B' is ...... to the matrix of T relative to B.

True or False

Write 'T' for true and 'F' for false statement.

1. The relation of similarity is an equivalence relation in the set of all linear transformations on a vector space $V(F)$.
2. Similar matrices have the same trace.

Answers

Multiple Choice Questions

1. (a) 2. (b) 3. (b) 4. (a)

Fill in the Blank(s)

1. $(x - 2y, y)$ 2. trace 3. $=$
4. unit matrix of order n 5. similar

True or False

1. T 2. T

□

Chapter

3

Linear Functionals and Dual Space

1 Linear Functionals

(Kumaun 2015)

Let $V(F)$ be a vector space. We know that the field F can be regarded as a vector space over F. This is the vector space $F(F)$ or F^1. We shall simply denote it by F. A linear transformation from V into F is called a **linear functional** on V. We shall now give independent definition of a linear functional.

Linear Functionals: Definition: *Let $V(F)$ be a vector space. A function f 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.**

Illustration 1: Let $V_n(F)$ be the vector space of ordered n-tuples of the elements of the field F.

Let $x_1, x_2, \ldots, x_n$ be n field elements of F. If

$$\alpha = (a_1, a_2, \ldots, a_n) \in V_n(F),$$

let f be a function from $V_n(F)$ into F defined by

$$f(\alpha) = x_1 a_1 + x_2 a_2 + \ldots + x_n a_n.$$

Let $\beta = (b_1, b_2, \ldots, b_n) \in V_n(F)$. If $a, b \in F$, we have

$$\begin{aligned} f(a\alpha + b\beta) &= f[a(a_1, \ldots, a_n) + b(b_1, \ldots, b_n)] \\ &= f(aa_1 + bb_1, \ldots, aa_n + bb_n) \\ &= x_1(aa_1 + bb_1) + \ldots + x_n(aa_n + bb_n) \\ &= a(x_1 a_1 + \ldots + x_n a_n) + b(x_1 b_1 + \ldots + x_n b_n) \\ &= af(a_1, \ldots, a_n) + bf(b_1, \ldots, b_n) = af(\alpha) + bf(\beta). \end{aligned}$$

$\therefore$ f is a linear functional on $V_n(F)$.

Illustration 2: Now we shall give a very important example 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 F 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

$$\text{tr } A = a_{11} + a_{22} + \ldots + a_{nn} = \sum_{i=1}^{n} a_{ii}.$$

Thus the trace of A is the scalar obtained by adding the elements of A lying along the principal diagonal.

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

$$\begin{aligned} \text{tr}(aA + bB) &= \text{tr}(a[a_{ij}]_{n \times n} + b[b_{ij}]_{n \times n}) = \text{tr}([aa_{ij} + bb_{ij}]_{n \times n}) \\ &= \sum_{i=1}^{n}(aa_{ii} + bb_{ii}) = a\sum_{i=1}^{n} a_{ii} + b\sum_{i=1}^{n} b_{ii} \\ &= a(\text{tr } A) + b(\text{tr } B). \end{aligned}$$

Illustration 3: Now we shall give another important example of a linear functional. *Let V be a finite-dimensional vector space over the field F and let B be an ordered basis for V. The function f_i which assigns to each vector α in V the i^{th} coordinate of α relative to the ordered basis B is a linear functional on V.*

Let $B = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$.

If $\alpha = a_1\alpha_1 + a_2\alpha_2 + \ldots + a_n\alpha_n \in V$, then by definition of f_i, we have

$$f_i(\alpha) = a_i.$$

Similarly if $\beta = b_1\alpha_1 + \ldots + b_n\alpha_n \in V$, then $f_i(\beta) = b_i$.

If $a, b \in F$, we have

$$\begin{aligned} f_i(a\alpha + b\beta) &= f_i[a(a_1\alpha_1 + \ldots + a_n\alpha_n) + b(b_1\alpha_1 + \ldots + b_n\alpha_n)] \\ &= f_i[(aa_1 + bb_1)\alpha_1 + \ldots + (aa_n + bb_n)\alpha_n] \\ &= aa_i + bb_i = af_i(\alpha) + bf_i(\beta). \end{aligned}$$

Hence f_i is a linear functional on V.

Illustration 4: *Let V be the vector space of polynomials in x over $\mathbf{R}$. Let $T : V \to \mathbf{R}$ be the integral operator defined by $T(f(x)) = \int_0^1 f(x)\,dx$. Then T is linear and hence it is a linear functional on V.*

2 Some Particular Linear Functionals

1. Zero Functional: *Let V be a vector space over the field F. The function f from V into F defined by*

$$f(\alpha) = 0 \text{ (zero of } F) \ \forall\ \alpha \in V$$

is a linear functional on V.

Proof: Let $\alpha, \beta \in \forall$ and $a, b \in F$. We have

$$f(a\alpha + b\beta) = 0 \qquad \text{[By def. of } f\text{]}$$
$$= a0 + b0 = af(\alpha) + bf(\beta).$$

$\therefore$ f is a linear functional on V. It is called the zero functional and we shall in future denote it by $\hat{\mathbf{0}}$

2. Negative of a 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. **(Kumaun 2013)**

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

$$(-f)(a\alpha + b\beta) = -[f(a\alpha + b\beta)] \qquad \text{[By def. of } -f\text{]}$$
$$= -[af(\alpha) + bf(\beta)] \qquad [\because f \text{ is a linear functional}]$$
$$= a[-f(\alpha)] + b[-f(\beta)]$$
$$= a[(-f)(\alpha)] + b[(-f)(\beta)].$$

$\therefore$ $-f$ is a linear functional on V.

Properties of a Linear Functional:

Theorem 1: *Let f be a linear functional on a vector space V(F). Then*

(i) $f(\mathbf{0}) = 0$ *where* $\mathbf{0}$ *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: (i) Let $\alpha \in V$. Then $f(\alpha) \in F$.

We have
$$f(\alpha) + 0 = f(\alpha) \qquad [\because 0 \text{ is zero element of } F]$$
$$= f(\alpha + \mathbf{0}) \qquad [\because \mathbf{0} \text{ is zero element of } V]$$
$$= f(\alpha) + f(\mathbf{0}) \qquad [\because f \text{ is a linear functional}]$$

Now F is a field. Therefore

$$f(\alpha) + 0 = f(\alpha) + f(\mathbf{0})$$

$\Rightarrow$ $f(\mathbf{0}) = 0$, by left cancellation law for addition in F.

(ii) We have $f[\alpha + (-\alpha)] = f(\alpha) + f(-\alpha)$ $\quad [\because f$ is a linear functional]

But $f[\alpha + (-\alpha)] = f(\mathbf{0}) = 0$ $\quad$ [By (i)]

Thus in F, we have

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

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

3 Dual Spaces

(Gorakhpur 2015)

Let V' be the set of all linear functionals on a vector space $V(F)$. Sometimes we denote this set by V^*. Now our aim is to impose a vector space structure on the set V' over the same field F. For this purpose we shall have to suitably define addition in V' and scalar multiplication in V' over F.

Theorem: *Let V be a vector space over the field F. Let f_1 and f_2 be linear functionals 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' of all linear functionals on V, together with the addition and scalar multiplication defined as above is a vector space over the field F.

Proof: Suppose f_1 and f_2 are linear functionals on V and we define $f_1 + f_2$ as follows :

$$(f_1 + f_2)(\alpha) = f_1(\alpha) + f_2(\alpha) \ \forall \ \alpha \in V. \qquad \ldots(1)$$

Since $f_1(\alpha) + f_2(\alpha) \in F$, therefore $f_1 + f_2$ is a function from V into F.

Let $a, b \in F$ and $\alpha, \beta \in V$. Then

$$\begin{aligned}
(f_1 + f_2)(a\alpha + b\beta) &= f_1(a\alpha + b\beta) + f_2(a\alpha + b\beta) && \text{[By (1)]} \\
&= [af_1(\alpha) + bf_1(\beta)] + [af_2(\alpha) + bf_2(\beta)] \\
& && [\because f_1 \text{ and } f_2 \text{ are linear functionals}] \\
&= a[f_1(\alpha) + f_2(\alpha)] + b[f_1(\beta) + f_2(\beta)] \\
&= a[(f_1 + f_2)(\alpha)] + b[(f_1 + f_2)(\beta)] && \text{[By (1)]}
\end{aligned}$$

$\therefore$ $(f_1 + f_2)$ is a linear functional on V. Thus

$$f_1, f_2 \in V' \Rightarrow f_1 + f_2 \in V'.$$

Therefore V' is closed with respect to addition defined in it.

Again let $f \in V'$ and $c \in F$. Let us define cf as follows :

$$(cf)(\alpha) = cf(\alpha) \ \forall \ \alpha \in V. \qquad \ldots(2)$$

Since $cf(\alpha) \in F$, therefore cf is a function from V into F.

Let $a, b \in F$ and $\alpha, \beta \in V$. Then

$$\begin{aligned}
(cf)(a\alpha + b\beta) &= cf(a\alpha + b\beta) && \text{[By (2)]} \\
&= c[af(\alpha) + bf(\beta)] && [\because f \text{ is linear functional}] \\
&= c[af(\alpha)] + c[bf(\beta)] && [\because F \text{ is a field}] \\
&= (ca)f(\alpha) + (cb)f(\beta) \\
&= (ac)f(\alpha) + (bc)f(\beta) \\
&= a[cf(\alpha)] + b[cf(\beta)] \\
&= a[(cf)(\alpha)] + b[(cf)(\beta)].
\end{aligned}$$

$\therefore$ cf is a linear functional on V. Thus

$$f \in V' \text{ and } c \in F \Rightarrow cf \in V'.$$

Therefore V' is closed with respect to scalar multiplication defined in it.

Associativity of addition in V'.

Let $f_1, f_2, f_3 \in V'$. If $\alpha \in V$, then

$$[f_1 + (f_2 + f_3)](\alpha) = f_1(\alpha) + (f_2 + f_3)(\alpha) \qquad \text{[By (1)]}$$
$$= f_1(\alpha) + [f_2(\alpha) + f_3(\alpha)] \qquad \text{[By (1)]}$$
$$= [f_1(\alpha) + f_2(\alpha)] + f_3(\alpha) \qquad [\because \text{ addition in } F \text{ is associative}]$$
$$= (f_1 + f_2)(\alpha) + f_3(\alpha) \qquad \text{[By (1)]}$$
$$= [(f_1 + f_2) + f_3](\alpha) \qquad \text{[By (1)]}$$

$\therefore$ $f_1 + (f_2 + f_3) = (f_1 + f_2) + f_3$

[By def. of equality of two functions]

Commutativity of addition in V': Let $f_1, f_2 \in V'$. If α is any element of V, then

$$(f_1 + f_2)(\alpha) = f_1(\alpha) + f_2(\alpha) \qquad \text{[By (1)]}$$
$$= f_2(\alpha) + f_1(\alpha) \qquad [\because \text{ Addition in } F \text{ is commutative}]$$
$$= (f_2 + f_1)(\alpha) \qquad \text{[By (1)]}$$

$\therefore$ $f_1 + f_2 = f_2 + f_1$.

Existence of additive identity in V': Let $\hat{\mathbf{0}}$ be the zero linear functional in V *i.e.*,

$$\hat{\mathbf{0}}(\alpha) = 0 \ \ \forall \ \ \alpha \in V.$$

Then $\hat{\mathbf{0}} \in V'$. If $f \in V'$ and $\alpha \in V$, we have

$$(\hat{\mathbf{0}} + f)(\alpha) = \hat{\mathbf{0}}(\alpha) + f(\alpha) \qquad \text{[By (1)]}$$
$$= 0 + f(\alpha) \qquad [\text{By def. of } \hat{\mathbf{0}}]$$
$$= f(\alpha) \qquad [0 \text{ being additive identity in } F]$$

$\therefore$ $\hat{\mathbf{0}} + f = f \ \forall \ f \in V'$.

$\therefore$ $\hat{\mathbf{0}}$ is the additive identity in V'.

Existence of additive inverse of each element in V':

Let $f \in V'$. Let us define $-f$ as follows :

$$(-f)(\alpha) = -f(\alpha) \ \ \forall \ \ \alpha \in V.$$

Then $-f \in V'$. If $\alpha \in V$, we have

$$(-f + f)(\alpha) = (-f)(\alpha) + f(\alpha) \qquad \text{[By (1)]}$$
$$= -f(\alpha) + f(\alpha) \qquad [\text{By def. of } -f]$$
$$= 0$$
$$= \hat{\mathbf{0}}(\alpha) \qquad [\text{By def. of } \hat{\mathbf{0}}]$$

$\therefore$ $-f + f = \hat{\mathbf{0}}$ for every $f \in V'$.

Thus each element in V' possesses additive inverse. Therefore V' is an **abelian group** with respect to addition defined in it.

Further we make the following observations :

(i) Let $c \in F$ and $f_1, f_2 \in V'$. If α is any element in V, we have

$$[c\,(f_1 + f_2)](\alpha) = c\,[(f_1 + f_2)(\alpha)] \qquad \text{[By (2)]}$$
$$= c\,[f_1(\alpha) + f_2(\alpha)] \qquad \text{[By (1)]}$$
$$= c f_1(\alpha) + c f_2(\alpha)$$
$$= (c f_1)(\alpha) + (c f_2)(\alpha) \qquad \text{[By (2)]}$$
$$= (c f_1 + c f_2)(\alpha) \qquad \text{[By (1)]}$$

$\therefore \qquad c\,(f_1 + f_2) = c f_1 + c f_2.$

(ii) Let $a, b \in F$ and $f \in V'$. If $\alpha \in V$, we have

$$[(a + b) f](\alpha) = (a + b) f(\alpha) \qquad \text{[By (2)]}$$
$$= af(\alpha) + bf(\alpha) \qquad [\because F \text{ is a field}]$$
$$= (af)(\alpha) + (bf)(\alpha) \qquad \text{[By (2)]}$$
$$= (af + bf)(\alpha) \qquad \text{[By (1)]}$$

$\therefore \qquad (a + b) f = af + bf.$

(iii) Let $a, b \in F$ and $f \in V'$. If $\alpha \in V$, we have

$$[(ab) f](\alpha) = (ab) f(\alpha) \qquad \text{[By (2)]}$$
$$= a\,[bf(\alpha)] \qquad [\because \text{multiplication in } F \text{ is associative}]$$
$$= a\,[(bf)(\alpha)] \qquad \text{[By (2)]}$$
$$= [a\,(bf)](\alpha) \qquad \text{[By (2)]}$$

$\therefore \qquad (ab) f = a\,(bf).$

(iv) Let 1 be the multiplicative identity of F and $f \in V'$. If $\alpha \in V$, we have

$$(1f)(\alpha) = 1f(\alpha) \qquad \text{[By (2)]}$$
$$= f(\alpha) \qquad [\because F \text{ is a field}]$$

$\therefore \qquad 1f = f.$

Hence V' is a vector space over the field F.

Dual Space:

Definition: *Let V be a vector space over the field F. Then the set V' of all linear functionals on V is also a vector space over the field F. The vector space V' is called the dual space of V.*

Sometimes V^* and $\hat{V}$ are also used to denote the dual space of V. The dual space of V is also called the **conjugate space** of V.

4 Dual Bases

Theorem 1: *Let V be an n-dimensional vector space over the field F and let $B = \{\alpha_1, \ldots, \alpha_n\}$ be an ordered basis for V. If $\{x_1, \ldots, x_n\}$ is any ordered set of n scalars, then there exists a unique linear functional f on V such that*

$$f(\alpha_i) = x_i\,, i = 1, 2, \ldots, n.$$

Proof: Existence of f. Let $\alpha \in V$.

Since $B = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ is a basis for V, therefore there exist unique scalars $a_1, a_2, \ldots, a_n$ such that

$$\alpha = a_1\alpha_1 + \ldots + a_n \alpha_n.$$

For this vector α, let us define $f(\alpha) = a_1 x_1 + \ldots + a_n x_n$.

Obviously $f(\alpha)$ as defined above is a unique element of F. Therefore f is a well-defined rule for associating with each vector α in V a unique scalar $f(\alpha)$ in F. Thus f is a function from V into F.

The unique representation of $\alpha_i \in V$ as a linear combination of the vectors belonging to the basis B is

$$\alpha_i = 0\alpha_1 + 0\alpha_2 + \ldots + 1\alpha_i + 0\alpha_{i+1} + \ldots + 0\alpha_n.$$

Therefore according to our definition of f, we have

$$f(\alpha_i) = 0x_1 + 0x_2 + \ldots + 1x_i + 0x_{i+1} + \ldots + 0x_n$$

i.e., $$f(\alpha_i) = x_i, \; i = 1, 2, \ldots, n.$$

Now to show that f is a linear functional.

Let $a, b \in F$ and $\alpha, \beta \in V$. Also let

$$\alpha = a_1\alpha_1 + \ldots + a_n \alpha_n, \text{ and } \beta = b_1\alpha_1 + \ldots + b_n \alpha_n.$$

Then
$$\begin{aligned} f(a\alpha + b\beta) &= f[a(a_1\alpha_1 + \ldots + a_n\alpha_n) + b(b_1\alpha_1 + \ldots + b_n\alpha_n)] \\ &= f[(aa_1 + bb_1)\alpha_1 + \ldots + (aa_n + bb_n)\alpha_n] \\ &= (aa_1 + bb_1)x_1 + \ldots + (aa_n + bb_n)x_n \quad \text{[By def. of } f\text{]} \\ &= a(a_1x_1 + \ldots + a_n x_n) + b(b_1x_1 + \ldots + b_n x_n) \\ &= af(\alpha) + bf(\beta). \end{aligned}$$

$\therefore$ f is a linear functional on V.

Thus there exists a linear functional f on V such that $f(\alpha_i) = x_i, i = 1, 2, \ldots, n.$

Uniqueness of f: Let g be a linear functional on V such that

$$g(\alpha_i) = x_i, \; i = 1, 2, \ldots, n.$$

For any vector $\alpha = a_1\alpha_1 + \ldots + a_n\alpha_n \in V$, we have

$$\begin{aligned} g(\alpha) &= g(a_1\alpha_1 + \ldots + a_n\alpha_n) \\ &= a_1 g(\alpha_1) + \ldots + a_n g(\alpha_n) \quad [\because g \text{ is linear}] \\ &= a_1x_1 + \ldots + a_n x_n \quad \text{[By def. of } g\text{]} \\ &= f(\alpha). \quad \text{[By def. of } f\text{]} \end{aligned}$$

Thus $$g(\alpha) = f(\alpha) \; \forall \; \alpha \in V.$$

$\therefore$ $$g = f.$$

This shows the uniqueness of f.

Remark: From this theorem we conclude that if f is a linear functional on a finite dimensional vector space V, then f is completely determined if we mention under f the images of the elements of a basis set of V. If f and g are two linear functionals on V such that $f(\alpha_i) = g(\alpha_i)$ for all α_i belonging to a basis of V, then

$$f(\alpha) = g(\alpha) \; \forall \; \alpha \in V \; \textit{i.e.,} \; f = g.$$

Thus two linear functionals of V are equal if they agree on a basis of V.

Theorem 2: *Let V be an n-dimensional vector space over the field F and let $B = \{\alpha_1, \ldots, \alpha_n\}$ be a basis for V. Then there is a uniquely determined basis $B' = \{f_1, \ldots, f_n\}$ for V' such that $f_i(\alpha_j) = \delta_{ij}$. Consequently the dual space of an n-dimensional space is n-dimensional.*

(Gorakhpur 2015)

The basis B' is called the **dual basis** of B.

Proof: $B = \{\alpha_1, \ldots, \alpha_n\}$ is an ordered basis for V. Therefore by theorem 1, there exists a unique linear functional f_1 on V such that

$$f_1(\alpha_1) = 1, f_1(\alpha_2) = 0, \ldots, f_1(\alpha_n) = 0$$

where $\{1, 0, \ldots, 0\}$ is an ordered set of n scalars.

In fact, for each $i = 1, 2, \ldots, n$ there exists a unique linear functional f_i on V such that

$$f_i(\alpha_j) = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases}$$

i.e., $$f_i(\alpha_j) = \delta_{ij}, \qquad \ldots(1)$$

where $\delta_{ij} \in F$ is Kronecker delta *i.e.,* $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ if $i \neq j$.

Let $B' = \{f_1, \ldots, f_n\}$. Then B' is a subset of V' containing n distinct elements of V'. We shall show that B' is a basis for V'.

First we shall show that B' is linearly independent.

Let $$c_1 f_1 + c_2 f_2 + \ldots + c_n f_n = \hat{\mathbf{0}}$$

$\Rightarrow$ $(c_1 f_1 + \ldots + c_n f_n)(\alpha) = \hat{\mathbf{0}}(\alpha) \ \forall \alpha \in V$

$\Rightarrow$ $c_1 f_1(\alpha) + \ldots + c_n f_n(\alpha) = 0 \ \forall \alpha \in V$ $\quad [\because \hat{\mathbf{0}}(\alpha) = 0]$

$\Rightarrow$ $\sum_{i=1}^{n} c_i f_i(\alpha) = 0 \ \forall \alpha \in V$

$\Rightarrow$ $\sum_{i=1}^{n} c_i f_i(\alpha_j) = 0, j = 1, 2, \ldots, n$ $\quad$ [Putting $\alpha = \alpha_j$ where $j = 1, 2, \ldots, n$]

$\Rightarrow$ $\sum_{i=1}^{n} c_i \delta_{ij} = 0, j = 1, 2, \ldots, n$

$\Rightarrow$ $c_j = 0, j = 1, 2, \ldots, n$

$\Rightarrow$ $f_1, f_2, \ldots, f_n$ are linearly independent.

In the second place, we shall show that the linear span of B' is equal to V'.

Let f be any element of V'. The linear functional f will be completely determined if we define it on a basis for V. So let

$$f(\alpha_i) = a_i, i = 1, 2, \ldots, n. \qquad \ldots(2)$$

We shall show that

$$f = a_1 f_1 + \ldots + a_n f_n = \sum_{i=1}^{n} a_i f_i.$$

We know that two linear functionals on V are equal if they agree on a basis of V. So let $\alpha_j \in B$ where $j = 1, \ldots, n$. Then

$$\left[\sum_{i=1}^{n} a_i f_i\right](\alpha_j) = \sum_{i=1}^{n} a_i f_i(\alpha_j)$$

$$= \sum_{i=1}^{n} a_i \delta_{ij} \qquad \text{[From (1)]}$$

$$= a_j, \qquad \text{[On summing with respect to } i \text{ and remembering that } \delta_{ij} = 1 \text{ when } i = j \text{ and } \delta_{ij} = 0 \text{ when } i \neq j\text{]}$$

$$= f(\alpha_j) \qquad \text{[From (2)]}$$

Thus $\left[\sum_{i=1}^{n} a_i f_i\right](\alpha_j) = f(\alpha_j) \ \forall \alpha_j \in B.$

Therefore $f = \sum_{i=1}^{n} a_i f_i$. Thus every element f in V' can be expressed as a linear combination of $f_1, \ldots, f_n$.

$\therefore$ V' = linear span of B'. Hence B' is a basis for V'.

Now dim V' = number of distinct elements in $B' = n$.

Corollary: *If V is an n-dimensional vector space over the field F, then V is isomorphic to its dual space V'.*

Proof: We have dim V' = dim $V = n$.

$\therefore$ V is isomorphic to V'.

Theorem 3: *Let V be an n-dimensional vector space over the field F and let $B = \{\alpha_1, \ldots, \alpha_n\}$ be a basis for V. Let $B' = \{f_1, \ldots, f_n\}$ be the dual basis of B. Then for each linear functional f on V, we have*

$$f = \sum_{i=1}^{n} f(\alpha_i) f_i$$

and for each vector α in V we have

$$\alpha = \sum_{i=1}^{n} f_i(\alpha) \alpha_i.$$

Proof: Since B' is dual basis of B, therefore

$$f_i(\alpha_j) = \delta_{ij}. \qquad \ldots(1)$$

If f is a linear functional on V, then $f \in V'$ for which B' is basis. Therefore f can be expressed as a linear combination of $f_1, \ldots, f_n$.

Let $f = \sum_{i=1}^{n} c_i f_i$.

Then $f(\alpha_j) = \left(\sum_{i=1}^{n} c_i f_i\right)(\alpha_j) = \sum_{i=1}^{n} c_i f_i(\alpha_j)$

$$= \sum_{i=1}^{n} c_i \delta_{ij} \qquad \text{[From (1)]}$$

$$= c_j, \ j = 1, 2, \ldots, n.$$

$\therefore \qquad f = \sum_{i=1}^{n} f(\alpha_i) f_i .$

Now let α be any vector in V. Let

$$\alpha = x_1\alpha_1 + \ldots + x_n \alpha_n . \qquad \ldots(2)$$

Then $\qquad f_i(\alpha) = f_i\left(\sum_{j=1}^{n} x_j \alpha_j\right) \qquad \left[\text{From (2)}, \alpha = \sum_{j=1}^{n} x_j \alpha_j\right]$

$$= \sum_{j=1}^{n} x_j f_i(\alpha_j) \qquad [\because f_i \text{ is linear functional}]$$

$$= \sum_{j=1}^{n} x_j \delta_{ij} \qquad [\text{From (1)}]$$

$$= x_i .$$

$\therefore \qquad \alpha = f_1(\alpha)\alpha_1 + \ldots + f_n(\alpha)\alpha_n$

$$= \sum_{i=1}^{n} f_i(\alpha)\alpha_i .$$

Important: It should be noted that if $B = \{\alpha_1, \ldots, \alpha_n\}$ is an ordered basis for V and $B' = \{f_1, \ldots, f_n\}$ is the dual basis, then f_i is precisely the function which assigns to each vector α in V the i^{th} coordinate of α relative to the ordered basis B.

Theorem 4: *Let V be an n-dimensional vector space over the field F. If α is a non-zero vector in V, there exists a linear functional f on V such that $f(\alpha) \neq 0$.*

Proof: Since $\alpha \neq \mathbf{0}$, therefore $\{\alpha\}$ is a linearly independent subset of V. So it can be extended to form a basis for V. Thus there exists a basis $B = \{\alpha_1, \ldots, \alpha_n\}$ for V such that $\alpha_1 = \alpha$.

If $B' = \{f_1, \ldots, f_n\}$ is the dual basis, then

$$f_1(\alpha) = f_1(\alpha_1) = 1 \neq 0.$$

Thus there exists linear functional f_1 such that

$$f_1(\alpha) \neq 0.$$

Corollary: *Let V be an n-dimensional vector space over the field F. If*

$$f(\alpha) = 0 \;\forall\; f \in V', \textit{then } \alpha = \mathbf{0}.$$

Proof: Suppose $\alpha \neq \mathbf{0}$. Then there is a linear functional f on V such that $f(\alpha) \neq 0$. This contradicts the hypothesis that

$$f(\alpha) = 0 \;\forall\; f \in V'.$$

Hence we must have $\alpha = \mathbf{0}$.

Theorem 5: *Let V be an n-dimensional vector space over the field F. If α, β are any two different vectors in V, then there exists a linear functional f on V such that $f(\alpha) \neq f(\beta)$.*

Proof: We have $\alpha \neq \beta \Rightarrow \alpha - \beta \neq \mathbf{0}$.

Now $\alpha - \beta$ is a non-zero vector in V. Therefore by theorem 4, there exists a linear functional f on V such that

$$f(\alpha - \beta) \neq 0 \quad \Rightarrow \quad f(\alpha) - f(\beta) \neq 0 \quad \Rightarrow \quad f(\alpha) \neq f(\beta)$$

Hence the result.

5 Reflexivity (Second Dual Space or Bi-dual Space or Double Dual Space)

Second dual space (or Bi-dual space): We know that every vector space V possesses a dual space V' consisting of all linear functionals on V. Now V' is also a vector space. Therefore it will also possess a dual space $(V')'$ consisting of all linear functionals on V'. This dual space of V' is called the **Second dual space** or **Bi-dual space** of V and for the sake of simplicity we shall denote it by V''.

If V is finite-dimensional, then

$$\dim V = \dim V' = \dim V''$$

showing that they are **isomorphic** to each other.

Natural Isomorphism:

Theorem 1: *Let V be a finite dimensional vector space over the field F. If α is any vector in V, the function 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 $\alpha \Rightarrow L_\alpha$ is an isomorphism of V onto V''.

Proof: If $\alpha \in V$ and $f \in V'$, then $f(\alpha)$ is a unique element of F. Therefore the correspondence L_α defined by

$$L_\alpha(f) = f(\alpha) \ \forall f \in V' \qquad \ldots(1)$$

is a function from V' into F.

Let $a, b \in F$ and $f, g \in V'$, then

$$L_\alpha(af + bg) = (af + bg)(\alpha) \qquad \text{[From (1)]}$$
$$= (af)(\alpha) + (bg)(\alpha)$$
$$= a f(\alpha) + b g(\alpha)$$

[By scalar multiplication of linear functionals]

$$= a[L_\alpha(f)] + b[L_\alpha(g)]. \qquad \text{[From (1)]}$$

Therefore L_α is a linear functional on V' and thus $L_\alpha \in V''$.

Now let ψ be the function from V into V'' defined by

$$\psi(\alpha) = L_\alpha \ \forall \ \alpha \in V.$$

ψ is one-one: If $\alpha, \beta \in V$, then

$$\psi(\alpha) = \psi(\beta)$$
$$\Rightarrow \quad L_\alpha = L_\beta \ \Rightarrow \ L_\alpha(f) = L_\beta(f) \ \forall \ f \in V'$$
$$\Rightarrow \quad f(\alpha) = f(\beta) \ \forall f \in V' \qquad \text{[From (1)]}$$
$$\Rightarrow \quad f(\alpha) - f(\beta) = 0 \ \forall \ f \in V'$$
$$\Rightarrow \quad f(\alpha - \beta) = 0 \ \forall f \in V'$$
$$\Rightarrow \quad \alpha - \beta = \mathbf{0}$$

[$\because$ By theorem 4 of article 4, if $\alpha - \beta \neq \mathbf{0}$, then $\exists$ a linear functional f on V such that $f(\alpha - \beta) \neq 0$. Here we have $f(\alpha - \beta) = 0 \ \forall \ f \in V'$ and so $\alpha - \beta$ must be $\mathbf{0}$]

$\Rightarrow \qquad \alpha = \beta.$

$\therefore$ ψ is one-one.

ψ is a linear transformation:

Let $a, b \in F$ and $\alpha, \beta \in V$. Then

$$\psi(a\alpha + b\beta) = L_{a\alpha + b\beta} \qquad \text{[By def. of } \psi]$$

For every $f \in V'$, we have

$$L_{a\alpha + b\beta}(f) = f(a\alpha + b\beta)$$
$$= af(\alpha) + bf(\beta) \qquad \text{[From (1)]}$$
$$= aL_\alpha(f) + bL_\beta(f) \qquad \text{[From (1)]}$$
$$= (aL_\alpha)(f) + (bL_\beta)(f) = (aL_\alpha + bL_\beta)(f).$$

$$\therefore \qquad L_{a\alpha + b\beta} = aL_\alpha + bL_\beta = a\,\psi(\alpha) + b\,\psi(\beta).$$

Thus $\psi(a\alpha + b\beta) = a\,\psi(\alpha) + b\,\psi(\beta)$.

$\therefore$ ψ is a linear transformation from V into V''. We have $\dim V = \dim V''$. Therefore ψ is one-one implies that ψ must also be onto.

Hence ψ is an isomorphism of V onto V''.

Note: The correspondence $\alpha \to L_\alpha$ as defined in the above theorem is called the **natural correspondence** between V and V''. It is important to note that the above theorem shows not only that V and V'' are isomorphic—this much is obvious from the fact that they have the same dimension — but that the natural correspondence is an isomorphism. This property of vector spaces is called **reflexivity.** Thus in the above theorem we have proved that *every finite-dimensional vector space is reflexive.*

In future we shall identify V'' with V through the natural isomorphism $\alpha \leftrightarrow L_\alpha$. We shall say that the element L of V'' is the same as the element α of V iff $L = L_\alpha$ *i.e.*, iff

$$L(f) = f(\alpha) \;\forall\; f \in V'.$$

It will be in this sense that we shall regard $V'' = V$.

Theorem 2: *Let V be a finite dimensional vector space over the field F. If L is a linear functional on the dual space V' of V, then there is a unique vector α in V such that*

$$L(f) = f(\alpha) \;\forall\; f \in V'.$$

Proof: This theorem is an immediate corollary of theorem 1. We should first prove theorem 1. Then we should conclude like this :

The correspondence $\alpha \to L_\alpha$ is a one-to-one correspondence between V and V''. Therefore if $L \in V''$, there exists a unique vector α in V such that $L = L_\alpha$ *i.e.*, such that

$$L(f) = f(\alpha) \;\forall\; f \in V'.$$

Theorem 3: *Let V be a finite dimensional vector space over the field F. Each basis for V' is the dual of some basis for V.*

Proof: Let $B' = \{f_1, f_2, \ldots, f_n\}$ be a basis for V.

Then there exists a dual basis $(B')' = \{L_1, L_2, \ldots, L_n\}$ for V'' such that

$$L_i(f_j) = \delta_{ij}. \qquad \ldots(1)$$

By previous theorem, for each i there is a vector α_i in V such that

$$L_i = L_{\alpha_i} \text{ where } L_{\alpha_i}(f) = f(\alpha_i) \ \forall f \in V'. \quad \ldots(2)$$

The correspondence $\alpha \leftrightarrow L_\alpha$ is an isomorphism of V onto V''. Under an isomorphism a basis is mapped onto a basis. Therefore $B = \{\alpha_1, \ldots, \alpha_n\}$ is a basis for V because it is the image set of a basis for V'' under the above isomorphism.

Putting $f = f_j$ in (2), we get

$$f_j(\alpha_i) = L_{\alpha_i}(f_j) = L_i(f_j)$$

$$= \delta_{ij}. \quad \text{[From (1)]}$$

$\therefore$ $B' = \{f_1, \ldots, f_n\}$ is the dual of the basis B.

Hence the result.

Theorem 4: *Let V be a finite dimensional vector space over the field F. Let B be a basis for V and B' be the dual basis of B. Then show that*

$$B'' = (B')' = B. \quad \textbf{(Kumaun 2007)}$$

Proof: Let $B = \{\alpha_1, \ldots, \alpha_n\}$ be a basis for V,

$B' = \{f_1, \ldots, f_n\}$ be the dual basis of B in V'

and $B'' = (B')' = \{L_1, \ldots, L_n\}$ be the dual basis of B in V''.

Then $f_i(\alpha_j) = \delta_{ij}$,

and $L_i(f_j) = \delta_{ij}, i = 1, \ldots, n; j = 1, \ldots, n.$

If $\alpha \in V$, then there exists $L_\alpha \in V''$ such that

$$L_\alpha(f) = f(\alpha) \ \forall f \in V'.$$

Taking α_i in place of α, we see that for each $j = 1, \ldots, n$,

$$L_{\alpha_i}(f_j) = f_j(\alpha_i) = \delta_{ij} = L_i(f_j).$$

Thus L_{α_i} and L_i agree on a basis for V'. Therefore $L_{\alpha_i} = L_i$.

If we identify V'' with V through the natural isomorphism $\alpha \leftrightarrow L_\alpha$, then we consider L_α as the same element as α.

So $L_i = L_{\alpha_i} = \alpha_i$ where $i = 1, 2, \ldots, n$.

Thus $B'' = B$.

Illustrative Examples

Example 1: *Find the dual basis of the basis set*

$$B = \{(1, -1, 3), (0, 1, -1), (0, 3, -2)\} \textit{ for } V_3(\mathbf{R}).$$

(Kumaun 2008, 09, 11, 13)

Solution: Let $\alpha_1 = (1, -1, 3), \alpha_2 = (0, 1, -1), \alpha_3 = (0, 3, -2)$.

Then $B = \{\alpha_1, \alpha_2, \alpha_3\}$.

If $B' = \{f_1, f_2, f_3\}$ is dual basis of B, then

$$f_1(\alpha_1) = 1, f_1(\alpha_2) = 0, f_1(\alpha_3) = 0,$$

$$f_2(\alpha_1) = 0, f_2(\alpha_2) = 1, f_2(\alpha_3) = 0,$$

and $f_3(\alpha_1) = 0, f_3(\alpha_2) = 0, f_3(\alpha_3) = 1.$

Now to find explicit expressions for f_1, f_2, f_3.

Let $(a, b, c) \in V_3(\mathbf{R})$.

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

$= x\alpha_1 + y\alpha_2 + z\alpha_3$.

Then $f_1(a, b, c) = x, f_2(a, b, c) = y,$ and $f_3(a, b, c) = z$.

Now to find the values of x, y, z.

From (1), we have

$$x = a, -x + y + 3z = b, 3x - y - 2z = c.$$

Solving these equations, we have

$$x = a, y = 7a - 2b - 3c, z = b + c - 2a.$$

Hence $f_1(a, b, c) = a, \ f_2(a, b, c) = 7a - 2b - 3c,$

and $f_3(a, b, c) = -2a + b + c$.

Therefore $B' = \{f_1, f_2, f_3\}$ is a dual basis of B where f_1, f_2, f_3 are as defined above.

Example 2: *The vectors* $\alpha_1 = (1, 1, 1), \alpha_2 = (1, 1, -1)$, *and* $\alpha_3 = (1, -1, -1)$ *form a basis of* $V_3(\mathbf{C})$. *If* $\{f_1, f_2, f_3\}$ *is the dual basis and if* $\alpha = (0, 1, 0)$, *find* $f_1(\alpha), f_2(\alpha)$ *and* $f_3(\alpha)$.

Solution: Let $\alpha = a_1\alpha_1 + a_2\alpha_2 + a_3\alpha_3$. Then

$$f_1(\alpha) = a_1, f_2(\alpha) = a_2, f_3(\alpha) = a_3.$$

Now $\alpha = a_1\alpha_1 + a_2\alpha_2 + a_3\alpha_3$

$\Rightarrow$ $(0, 1, 0) = a_1(1, 1, 1) + a_2(1, 1, -1) + a_3(1, -1, -1)$

$\Rightarrow$ $(0, 1, 0) = (a_1 + a_2 + a_3, a_1 + a_2 - a_3, a_1 - a_2 - a_3)$

$\Rightarrow$ $a_1 + a_2 + a_3 = 0, a_1 + a_2 - a_3 = 1, a_1 - a_2 - a_3 = 0$

$\Rightarrow$ $a_1 = 0, a_2 = \frac{1}{2}, a_3 = -\frac{1}{2}$.

$\therefore$ $f_1(\alpha) = 0, f_2(\alpha) = \frac{1}{2}, f_3(\alpha) = -\frac{1}{2}$.

Example 3: *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* $f(\alpha) = x$?

Solution: f is a non-zero linear functional on V. Therefore there must be some non-zero vector β in V such that

$f(\beta) = y$ where y is a non-zero element of F.

If x is any element of F, then

$$x = (xy^{-1})\, y = (xy^{-1})\, f(\beta)$$

$= f[(xy^{-1})\beta]$ [$\because$ f is linear functional]

Thus there exists $\alpha = (xy^{-1})\beta \in V$ such that $f(\alpha) = x$.

Note 1: If f is a non-zero linear functional on $V(F)$, then f is necessarily a function from V onto F.

Note 2: In some books $f(\alpha)$ is written as $[\alpha, f]$.

Example 4: *Prove that if f is a linear functional on an n-dimensional vector space $V(F)$, then the set of all those vectors α for which $f(\alpha) = 0$ is a subspace of V, what is the dimension of that subspace ?*

Solution: Let $N = \{\alpha \in V : f(\alpha) = 0\}$.

N is not empty because at least $\mathbf{0} \in N$. Remember that

$$f(\mathbf{0}) = 0.$$

Let $\alpha, \beta \in N$. Then $f(\alpha) = 0, f(\beta) = 0$.

If $a, b \in F$, we have

$$f(a\alpha + b\beta) = af(\alpha) + bf(\beta) = a0 + b0 = 0.$$

$\therefore \quad a\alpha + b\beta \in N.$

Thus $a, b \in F$ and $\alpha, \beta \in N$

$\Rightarrow \quad a\alpha + b\beta \in N.$

$\therefore$ N is a subspace of V. This subspace N is the null space of f.

We know that dim V = dim N + dim (range of f).

(i) If f is zero linear functional, then range of f consists of zero element of F alone. Therefore dim (range of f) $= 0$ in this case.

$\therefore$ In this case, we have

$$\dim V = \dim N + 0$$

$\Rightarrow \quad n = \dim N.$

(ii) If f is a non-zero linear functional on V, then f is onto F. So range of f consists of all F in this case. The dimension of the vector space F' is 1.

$\therefore$ In this case we have

$$\dim V = \dim N + 1$$

$\Rightarrow \quad \dim N = n - 1.$

Example 5: *Let V be a vector space over the field F. Let f be a non-zero linear functional on V and let N be the null space of f. Fix a vector 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.*

Solution: Since f is a non-zero linear functional on V, therefore there exists a non-zero vector α_0 in V such that $f(\alpha_0) \neq 0$. Consequently $\alpha_0 \notin N$. Let $f(\alpha_0) = y \neq 0$.

Let α be any element of V and let $f(\alpha) = x$.

We have $\quad f(\alpha) = x$

$\Rightarrow \quad f(\alpha) = (x\, y^{-1})\, y \qquad [\because 0 \neq y \in F \Rightarrow y^{-1} \text{ exists}]$

$\Rightarrow \quad f(\alpha) = cy$ where $c = x\, y^{-1} \in F$

$\Rightarrow \quad f(\alpha) = c\, f(\alpha_0)$

$\Rightarrow \quad f(\alpha) = f(c\, \alpha_0) \qquad [\because f \text{ is a linear functional}]$

$\Rightarrow \quad f(\alpha) - f(c\, \alpha_0) = 0$

$\Rightarrow \quad f(\alpha - c\, \alpha_0) = 0$

$\Rightarrow \quad \alpha - c\, \alpha_0 \in N$

$\Rightarrow \quad \alpha - c\,\alpha_0 = \beta$ for some $\beta \in N$.

$\Rightarrow \quad \alpha = c\,\alpha_0 + \beta$.

If possible, let

$$\alpha = c'\alpha_0 + \beta' \text{ where } c' \in F \text{ and } \beta' \in N.$$

Then $\quad c\,\alpha_0 + \beta = c'\alpha_0 + \beta' \qquad \ldots(1)$

$\Rightarrow \quad (c - c')\,\alpha_0 + (\beta - \beta') = \mathbf{0}$

$\Rightarrow \quad f\,[(c - c')\,\alpha_0 + (\beta - \beta')] = f\,(\mathbf{0})$

$\Rightarrow \quad (c - c')\,f\,(\alpha_0) + f\,(\beta - \beta') = 0$

$\Rightarrow \quad (c - c')\,f\,(\alpha_0) = 0$

$[\because\ \beta, \beta' \in N \ \Rightarrow\ \beta - \beta' \in N$ and thus $f\,(\beta - \beta') = 0]$

$\Rightarrow \quad (c - c') = 0 \qquad [\because\ f\,(\alpha_0)$ is a non-zero element of $F\,]$

$c = c'$.

Putting $c = c'$ in (1), we get $c\alpha_0 + \beta = c\alpha_0 + \beta' \ \Rightarrow\ \beta = \beta'$.

Hence c and β are unique.

Example 6: *If f and g are in V' such that $f\,(\alpha) = 0 \Rightarrow g\,(\alpha) = 0$, prove that $g = kf$ for some $k \in F$.*

Solution: It is given that $f\,(\alpha) = 0 \Rightarrow g\,(\alpha) = 0$. Therefore if α belongs to the null space of f, then α also belongs to the null space of g. Thus null space of f is a subset of the null space of g.

(i) If f is zero linear functional, then null space of f is equal to V. Therefore in this case V is a subset of null space of g. Hence null space of g is equal to V. So g is also zero linear functional. Hence we have

$$g = k\,f \ \forall\ k \in F.$$

(ii) Let f be non-zero linear functional on V. Then there exists a non-zero vector $\alpha_0 \in V$ such that $f\,(\alpha_0) = y$ where y is a non-zero element of F.

Let $\quad k = \dfrac{g\,(\alpha_0)}{f\,(\alpha_0)}$.

If $\alpha \in V$, then we can write

$$\alpha = c\,\alpha_0 + \beta \text{ where } c \in F \text{ and } \beta \in \text{null space of } f.$$

We have $\quad g\,(\alpha) = g\,(c\,\alpha_0 + \beta) = cg\,(\alpha_0) + g\,(\beta)$

$= cg\,(\alpha_0)$

$[\because\ \beta \in$ null space of $f \ \Rightarrow\ f\,(\beta) = 0$ and so $g\,(\beta) = 0]$

Also $\quad (k\,f\,)\,(\alpha) = k\,f\,(\alpha) = k\,f\,(c\alpha_0 + \beta)$

$= k\,[cf\,(\alpha_0) + f\,(\beta)]$

$= kc\,f\,(\alpha_0) \qquad [\because f\,(\beta) = 0]$

$= \dfrac{g\,(\alpha_0)}{f\,(\alpha_0)}\,cf\,(\alpha_0) = cg\,(\alpha_0)$.

Thus $\quad g\,(\alpha) = (k\,f\,)\,(\alpha)\ \forall\ \alpha \in V$.

$\therefore \quad g = k\,f$.

Comprehensive Exercise 1

1. Let $f : \mathbf{R}^2 \to \mathbf{R}$ and $g : \mathbf{R}^2 \to \mathbf{R}$ be the linear functionals defined by

$$f(x, y) = x + 2y \text{ and } g(x, y) = 3x - y.$$

Find (i) $f + g$ (ii) $4f$ (iii) $2f - 5g$.

2. Let $f : \mathbf{R}^2 \to \mathbf{R}$ and $g : \mathbf{R}^2 \to \mathbf{R}$ be the linear functionals defined by

$$f(x, y, z) = 2x - 3y + z \text{and } g(x, y, z) = 4x - 2y + 3z.$$

Find (i) $f + g$ (ii) $3f$ (iii) $2f - 5g$.

3. Find the dual basis of the basis set $\{(2, 1), (3, 1)\}$ for $\mathbf{R}^2$.

4. Find the dual basis of the basis set

$B = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ for V_3 **(R)**. **(Gorakhpur 2013)**

5. Find the dual basis of the basis set

$B = \{(1, -2, 3), (1, -1, 1), (2, -4, 7)\}$ of V_3**(R)**. **(Kumaun 2010)**

6. If $B = \{(-1, 1, 1), (1, -1, 1), (1, 1, -1)\}$ is a basis of V_3 **(R)**, then find the dual basis of B. **(Kumaun 2011, 15)**

7. Let V be the vector space of polynomials over $\mathbf{R}$ of degree ≤ 1 *i.e.*,

$V = \{a + bx : a, b \in \mathbf{R}\}$.

Let ϕ_1 and ϕ_2 be linear functionals on V defined by

$$\phi_1 [f(x)] = \int_0^1 f(x)\, dx \text{ and } \phi_2 [f(x)] = \int_0^2 f(x)\, dx.$$

Find the basis $\{f_1, f_2\}$ of V which is dual to $\{\phi_1, \phi_2\}$.

8. Let V be the vector space of polynomials over $\mathbf{R}$ of degree ≤ 2. Let ϕ_1, ϕ_2 and ϕ_3 be the linear functionals on V defined by

$$\phi_1 [f(x)] = \int_0^1 f(x)\, dx, \ \phi_2 [f(x)] = f'(1), \phi_3 [f(x)] = f(0).$$

Here $f(x) = a + bx + cx^2 \in V$ and $f'(x)$ denotes the derivative of $f(x)$. Find the basis $\{f_1(x), f_2(x), f_3(x)\}$ of V which is dual to $\{\phi_1, \phi_2, \phi_3\}$.

9. Prove that every finite dimensional vector space V is isomorphic to its second conjugate space V^{**} under an isomorphism which is independent of the choice of a basis in V.

10. Define a non-zero linear functional f on $\mathbf{C}^3$ such that $f(\alpha) = 0 = f(\beta)$ where $\alpha = (1, 1, 1)$ and $\beta\, (1, 1, -1)$.

Answers 1

1. (i) $4x + y$ (ii) $4x + 8y$ (iii) $-13x + 9y$
2. (i) $6x - 5y + 4z$ (ii) $6x - 9y + 3z$ (iii) $-16x + 4y - 13z$
3. $B' = \{f_1, f_2\}$ where $f_1(a, b) = 3b - a$, $f_2(a, b) = a - 2b$
4. $B' = \{f_1, f_2, f_3\}$ where $f_1(a, b, c) = a$, $f_2(a, b, c) = b$, $f_3(a, b, c) = c$
5. $B' = \{f_1, f_2, f_3\}$ where $f_1(a, b, c) = -3a - 5b - 2c$, $f_2(a, b, c) = 2a + b$, $f_3(a, b, c) = a + 2b + c$
6. $B' = \{f_1, f_2, f_3\}$, where $f_1(a, b, c) = \frac{1}{2}(b + c)$, $f_2(a, b, c) = \frac{1}{2}(a + c)$, $f_3(a, b, c) = \frac{1}{2}(a + b)$
7. $f_1(x) = 2 - 2x$, $f_2(x) = -\frac{1}{2} + x$
8. $f_1(x) = 3x - \frac{3}{2}x^2$, $f_2(x) = -\frac{1}{2}x + \frac{3}{4}x^2$, $f_3(x) = 1 - 3x + \frac{3}{2}x^2$

10. $f(a, b, c) = x(a - b)$, where x is any non-zero scalar.

6 Annihilators

(Kumaun 2008)

Definition: *If V is a vector space over the field F and S is a subset of V, the annihilator of S is the set S^0 of all linear functionals f on V such that*

$$f(\alpha) = 0 \;\; \forall \;\; \alpha \in S.$$

Sometimes $A(S)$ is also used to denote the annihilator of S.

Thus $S^0 = \{f \in V' : f(\alpha) = 0 \;\; \forall \, \alpha \in S\}$.

It should be noted that we have defined the annihilator of S which is simply a subset of V. S should not necessarily be a subspace of V.

If S = zero subspace of V, then $S^0 = V'$.

If $S = V$, then $S^0 = V^0$ = zero subspace of V'.

If V is finite dimensional and S contains a non-zero vector, then $S^0 \neq V'$. If $\mathbf{0} \neq \alpha \in S$, then there is a linear functional f on V such that $f(\alpha) \neq 0$. Thus there is $f \in V'$ such that $f \notin S^0$. Therefore $S^0 \neq V'$.

Theorem 1: *If S is any subset of a vector space $V(F)$, then S^0 is a subspace of V'.*

Proof: First we see that S^0 is a non-empty subset of V' because at least $\hat{\mathbf{0}} \in S^0$.

We have $\hat{\mathbf{0}}(\alpha) = 0 \ \forall \ \alpha \in S$.

Let $f, g \in S^0$. Then $f(\alpha) = 0 \ \forall \ \alpha \in S$, and $g(\alpha) = 0 \ \forall \ \alpha \in S$.

If $a, b \in F$, then

$$(af + bg)(\alpha) = (af)(\alpha) + (bg)(\alpha) = af(\alpha) + bg(\alpha) = a0 + b0 = 0.$$

$\therefore \qquad af + bg \in S^0$.

Thus $a, b \in F$ and $f, g \in S^0 \Rightarrow af + bg \in S^0$.

$\therefore \quad S^0$ is a subspace of V'.

Dimension of Annihilator:

Theorem 2: *Let V be a finite dimensional vector space over the field F, and let W be a subspace of V. Then*

$$\dim W + \dim W^0 = \dim V.$$

Proof: If W is zero subspace of V, then $W^0 = V'$.

$\therefore \qquad \dim W^0 = \dim V' = \dim V$.

Also in this case $\dim W = 0$. Hence the result.

Similarly the result is obvious when $W = V$.

Let us now suppose that W is a proper subspace of V.

Let $\dim V = n$, and $\dim W = m$ where $0 < m < n$.

Let $B_1 = \{\alpha_1, \ldots, \alpha_m\}$ be a basis for W. Since B_1 is a linearly independent subset of V also, therefore it can be extended to form a basis for V. Let $B = \{\alpha_1, \ldots, \alpha_m, \alpha_{m+1}, \ldots, \alpha_n\}$ be a basis for V.

Let $B' = \{f_1, \ldots, f_m, f_{m+1}, \ldots, f_n\}$ be the dual basis of B. Then B' is a basis for V' such that $f_i(\alpha_j) = \delta_{ij}$.

We claim that $S = \{f_{m+1}, \ldots, f_n\}$ is a basis for W^0.

Since $S \subset B'$, therefore S is linearly independent because B' is linearly independent. So S will be a basis for W^0, if W^0 is equal to the subspace of V' spanned by S *i.e.*, if $W^0 = L(S)$.

First we shall show that $W^0 \subseteq L(S)$. Let $f \in W^0$. Then $f \in V'$. So let

$$f = \sum_{i=1}^{n} x_i f_i. \qquad \ldots(1)$$

Now $f \in W^0 \Rightarrow f(\alpha) = 0 \ \forall \ \alpha \in W$

$\Rightarrow \qquad f(\alpha_j) = 0$ for each $j = 1, \ldots, m$ $\qquad$ [$\because \alpha_1, \ldots, \alpha_m$ are in W]

$\Rightarrow \qquad \left(\sum_{i=1}^{n} x_i f_i\right)(\alpha_j) = 0$ $\qquad$ [From (1)]

$\Rightarrow \quad \sum_{i=1}^{n} x_i f_i(\alpha_j) = 0 \quad \Rightarrow \quad \sum_{i=1}^{n} x_i \delta_{ij} = 0$

$\Rightarrow \quad x_j = 0$ for each $j = 1, \ldots, m$.

Putting $x_1 = 0, x_2 = 0, \ldots, x_m = 0$ in (1), we get

$$f = x_{m+1} f_{m+1} + \ldots + x_n f_n$$
$$= \text{a linear combination of the elements of } S.$$

$\therefore \quad f \in L(S)$.

Thus $f \in W^0 \quad \Rightarrow \quad f \in L(S)$.

$\therefore \quad W^0 \subseteq L(S)$.

Now we shall show that $L(S) \subseteq W^0$.

Let $g \in L(S)$. Then g is a linear combination of $f_{m+1}, \ldots, f_n$. Let

$$g = \sum_{k=m+1}^{n} y_k f_k. \qquad \ldots(2)$$

Let $\alpha \in W$. Then α is a linear combination of $\alpha_1, \ldots, \alpha_m$. Let

$$\alpha = \sum_{j=1}^{m} c_j \alpha_j. \qquad \ldots(3)$$

We have

$$g(\alpha) = g\left(\sum_{j=1}^{m} c_j \alpha_j\right) \qquad \text{[From (3)]}$$

$$= \sum_{j=1}^{m} c_j g(\alpha_j) \qquad [\because g \text{ is linear functional}]$$

$$= \sum_{j=1}^{m} c_j \left(\sum_{k=m+1}^{n} y_k f_k\right)(\alpha_j) \qquad \text{[From (2)]}$$

$$= \sum_{j=1}^{m} c_j \sum_{k=m+1}^{n} y_k f_k(\alpha_j) = \sum_{j=1}^{m} c_j \sum_{k=m+1}^{n} y_k \delta_{kj}$$

$$= \sum_{j=1}^{m} c_j \, 0 \qquad [\because \delta_{kj} = 0 \text{ if } k \neq j \text{ which is so for each } k = m+1, \ldots, n \text{ and for each } j = 1, \ldots, m]$$

$$= 0.$$

Thus $g(\alpha) = 0 \; \forall \; \alpha \in W$. Therefore $g \in W^0$.

Thus $g \in L(S) \quad \Rightarrow \quad g \in W^0$.

$\therefore \quad L(S) \subseteq W^0$.

Hence $W^0 = L(S)$ and S is a basis for W^0.

$\therefore \quad \dim W^0 = n - m = \dim V - \dim W$

or $\quad \dim V = \dim W + \dim W^0$.

Corollary: *If V is finite-dimensional and W is a subspace of V, then W' is isomorphic to V'/W^0.*

Proof: Let $\dim V = n$ and $\dim W = m$. W' is dual space of W,

so $\qquad \dim W' = \dim W = m.$

Now $\qquad \dim V'/W^0 = \dim V' - \dim W^0$

$$= \dim V - (\dim V - \dim W) = \dim W = m.$$

Since $\dim W' = \dim V'/W^0$, therefore

$$W' \cong V'/W^0.$$

Annihilator of an Annihilator: Let V be a vector space over the field F. If S is any subset of V, then S^0 is a subspace of V'. By definition of an annihilator, we have

$$(S^0)^0 = S^{00} = \{L \in V'' : L(f) = 0 \ \forall f \in S^0\}.$$

Obviously S^{00} is a subspace of V''. But if V is finite dimensional, then we have identified V'' with V through the natural isomorphism $\alpha \leftrightarrow L_\alpha$. Therefore we may regard S^{00} as a subspace of V. Thus

$$S^{00} = \{\alpha \in V : f(\alpha) = 0 \ \forall f \in S^0\}.$$

Theorem 3: *Let V be a finite dimensional vector space over the field F and let W be a subspace of V. Then $W^{00} = W$.*

Proof: We have

$$W^0 = \{f \in V' : f(\alpha) = 0 \ \forall \alpha \in W\} \qquad \ldots(1)$$

and $$W^{00} = \{\alpha \in V : f(\alpha) = 0 \ \forall f \in W^0\}. \qquad \ldots(2)$$

Let $\alpha \in W$. Then from (1), $f(\alpha) = 0 \ \forall f \in W^0$ and so from (2), $\alpha \in W^{00}$.

$\therefore \qquad \alpha \in W \Rightarrow \alpha \in W^{00}.$

Thus $W \subseteq W^{00}$. Now W is a subspace of V and W^{00} is also a subspace of V. Since $W \subseteq W^{00}$, therefore W is a subspace of W^{00}.

Now $\qquad \dim W + \dim W^0 = \dim V. \qquad$ [By theorem (2)]

Applying the same theorem for vector space V' and its subspace W^0, we get

$$\dim W^0 + \dim W^{00} = \dim V' = \dim V.$$

$\therefore \qquad \dim W = \dim V - \dim W^0 = \dim V - [\dim V - \dim W^{00}]$

$$= \dim W^{00}.$$

Since W is a subspace of W^{00} and $\dim W = \dim W^{00}$, therefore $W = W^{00}$.

7 Hyperspace

Definition : *Let V be an n dimensional vector space, then any subspace of dimension $(n-1)$ is called a hyperspace. These are also called hyperspaces or subspaces of codimension 1.*

A hyperspace of V can be represented as the null space of a non zero linear functional on V.

Let f be a non-zero functional on an n dimensional vector space V.

Since f is a linear transformation from V into the scalar field F, therefore range of f is a non-zero subspace of F.

$\therefore$ $\dim F = 1$ and rank $f = 1$.

By nullity theorem, we have

$$\Leftrightarrow \quad \text{nullity } f = \dim V - \text{rank } f = n - 1.$$

Thus nullspace of f is a hyperspace of V.

Remark: If U is a k-dimensional subspace of an n-dimensional vector space V, then U is the intersection of $(n-k)$ hyperspaces in V.

Illustrative Examples

Example 7: *If S_1 and S_2 are two subsets of a vector space V such that $S_1 \subseteq S_2$, then show that $S_2^0 \subseteq S_1^0$.* **(Kumaun 2008)**

Solution: Let $f \in S_2^0$.

Then $f(\alpha) = 0 \ \forall \alpha \in S_2$

$\Rightarrow f(\alpha) = 0 \ \forall \alpha \in S_1$ $\quad [\because S_1 \subseteq S_2]$

$\Rightarrow f \in S_1^0$.

$\therefore S_2^0 \subseteq S_1^0$.

Example 8: *Let V be a vector space over the field F. If S is any subset of V, then show that $S^0 = [L(S)]^0$.*

Solution: We know that $S \subseteq L(S)$.

$\therefore \quad [L(S)]^0 \subseteq S^0$. ...(1)

Now let $f \in S^0$. Then $f(\alpha) = 0 \ \forall \ \alpha \in S$.

If β is any element of $L(S)$, then $\beta = \sum_{i=1}^{n} x_i \alpha_i$ where each $\alpha_i \in S$.

We have $f(\beta) = \sum_{i=1}^{n} x_i f(\alpha_i) = 0$, since each $f(\alpha_i) = 0$.

Thus $f(\beta) = 0 \;\; \forall \;\; \beta \in L(S)$.

$\therefore$ $f \in (L(S))^0$.

Therefore $S^0 \subseteq (L(S))^0$ ...(2)

From (1) and (2), we conclude that $S^0 = (L(S))^0$.

Example 9: *Let V be a finite-dimensional vector space over the field F. If S is any subset of V, then $S^{00} = L(S)$.*

Solution: We have $S^0 = (L(S))^0$. [See Example 8]

$\therefore$ $S^{00} = (L(S))^{00}$. ...(1)

But V is finite-dimensional and $L(S)$ is a subspace of V. Therefore by theorem 3,

$$(L(S))^{00} = L(S).$$

$\therefore$ from (1), we have $S^{00} = L(S)$.

Example 10: *Let W_1 and W_2 be subspaces of a finite dimensional vector space V.*

(i) Prove that $(W_1 + W_2)^0 = W_1^0 \cap W_2^0$. **(Kumaun 2011, 13)**

(ii) Prove that $(W_1 \cap W_2)^0 = W_1^0 + W_2^0$. **(Kumaun 2013)**

Solution: (i) First we shall prove that

$$W_1^0 \cap W_2^0 \subseteq (W_1 + W_2)^0.$$

Let $f \in W_1^0 \cap W_2^0$. Then $f \in W_1^0$, $f \in W_2^0$.

Suppose α is any vector in $W_1 + W_2$. Then

$$\alpha = \alpha_1 + \alpha_2 \text{ where } \alpha_1 \in W_1, \; \alpha_2 \in W_2.$$

We have $f(\alpha) = f(\alpha_1 + \alpha_2) = f(\alpha_1) + f(\alpha_2)$

$= 0 + 0$ [$\because$ $\alpha_1 \in W_1$ and $f \in W_1^0 \Rightarrow f(\alpha_1) = 0$ and similarly $f(\alpha_2) = 0$]

$= 0$.

Thus $f(\alpha) = 0 \;\; \forall \; \alpha \in W_1 + W_2$.

$\therefore$ $f \in (W_1 + W_2)^0$.

$\therefore$ $W_1^0 \cap W_2^0 \subseteq (W_1 + W_2)^0$. ...(1)

Now we shall prove that

$$(W_1 + W_2)^0 \subseteq W_1^0 \cap W_2^0.$$

We have $W_1 \subseteq W_1 + W_2$.

$\therefore$ $(W_1 + W_2)^0 \subseteq W_1^0$. ...(2)

Similarly, $W_2 \subseteq W_1 + W_2$.

$\therefore \qquad (W_1 + W_2)^0 \subseteq W_2^{\,0}$. ...(3)

From (2) and (3), we have

$$(W_1 + W_2)^0 \subseteq W_1^{\,0} \cap W_2^{\,0}. \qquad \text{...(4)}$$

From (1) and (4), we have

$$(W_1 + W_2)^0 = W_1^{\,0} \cap W_2^{\,0}.$$

(ii) Let us use the result (a) for the vector space V' in place of the vector space V. Thus replacing W_1 by $W_1^{\,0}$ and W_2 by $W_2^{\,0}$ in (a), we get

$$(W_1^{\,0} + W_2^{\,0})^0 = W_1^{\,00} \cap W_2^{\,00}$$

$\Rightarrow \qquad (W_1^{\,0} + W_2^{\,0})^0 = W_1 \cap W_2 \qquad [\because W_1^{\,00} = W_1 \text{ etc.}]$

$\Rightarrow \qquad (W_1^{\,0} + W_2^{\,0})^{00} = (W_1 \cap W_2)^0$

$\Rightarrow \qquad W_1^{\,0} + W_2^{\,0} = (W_1 \cap W_2)^0$.

Example 11: *If W_1 and W_2 are subspaces of a vector space V, and if $V = W_1 \oplus W_2$, then $V' = W_1^{\,0} \oplus W_2^{\,0}$.*

Solution: To prove that $V' = W_1^{\,0} \oplus W_2^{\,0}$, we are to prove that

(i) $W_1^{\,0} \cap W_2^{\,0} = \{\hat{\mathbf{0}}\}$

and (ii) $V' = W_1^{\,0} + W_2^{\,0}$ *i.e.*, each $f \in V'$ can be written as $f_1 + f_2$

where $f_1 \in W_1^{\,0}, f_2 \in W_2^{\,0}$.

(i) First to prove that $W_1^{\,0} \cap W_2^{\,0} = \{\hat{\mathbf{0}}\}$.

Let $f \in W_1^{\,0} \cap W_2^{\,0}$.

Then $f \in W_1^{\,0}$ and $f \in W_2^{\,0}$.

If α is any vector in V, then, V being the direct sum of W_1 and W_2, we can write

$$\alpha = \alpha_1 + \alpha_2 \text{ where } \alpha_1 \in W_1, \ \alpha_2 \in W_2.$$

We have $\quad f(\alpha) = f(\alpha_1 + \alpha_2)$

$= f(\alpha_1) + f(\alpha_2) \qquad [\because f \text{ is linear functional}]$

$= 0 + 0 \qquad [\because f \in W_1^{\,0} \text{ and } \alpha_1 \in W_1 \Rightarrow f(\alpha_1) = 0$ and similarly $f(\alpha_2) = 0]$

$= 0$.

Thus $f(\alpha) = 0 \ \forall \alpha \in V$.

$\therefore \qquad f = \hat{\mathbf{0}}$.

$\therefore \qquad W_1^{\,0} \cap W_2^{\,0} = \{\hat{\mathbf{0}}\}$.

(ii) Now to prove that $V' = W_1^{\,0} + W_2^{\,0}$.

Let $f \in V'$.

If $\alpha \in V$, then α can be uniquely written as

$$\alpha = \alpha_1 + \alpha_2 \text{ where } \alpha_1 \in W_1, \alpha_2 \in W_2.$$

For each f, let us define two functions f_1 and f_2 from V into F such that

$$f_1(\alpha) = f_1(\alpha_1 + \alpha_2) = f(\alpha_2) \qquad \ldots(1)$$

and
$$f_2(\alpha) = f_2(\alpha_1 + \alpha_2) = f(\alpha_1). \qquad \ldots(2)$$

First we shall show that f_1 is a linear functional on V. Let $a, b \in F$ and $\alpha = \alpha_1 + \alpha_2, \beta = \beta_1 + \beta_2 \in V$ where $\alpha_1, \beta_1 \in W_1$ and $\alpha_2, \beta_2 \in W_2$. Then

$$\begin{aligned} f_1(a\alpha + b\beta) &= f_1[a(\alpha_1 + \alpha_2) + b(\beta_1 + \beta_2)] \\ &= f_1[(a\alpha_1 + b\beta_1) + (a\alpha_2 + b\beta_2)] \\ &= f(a\alpha_2 + b\beta_2) && [\because a\alpha_1 + b\beta_1 \in W_1, a\alpha_2 + b\beta_2 \in W_2] \\ &= af(\alpha_2) + bf(\beta_2) && [\because f \text{ is linear functional}] \\ &= af_1(\alpha) + bf_1(\beta) && [\text{From } (1)] \end{aligned}$$

$\therefore$ f_1 is linear functional on V *i.e.*, $f_1 \in V'$.

Now we shall show that $f_1 \in W_1^0$.

Let α_1 be any vector in W_1. Then α_1 is also in V. We can write

$$\alpha_1 = \alpha_1 + \mathbf{0}, \text{ where } \alpha_1 \in W_1, \mathbf{0} \in W_2.$$

$\therefore$ From (1), we have $f_1(\alpha_1) = f_1(\alpha_1 + \mathbf{0}) = f(\mathbf{0}) = 0$.

Thus $f_1(\alpha_1) = 0 \ \forall \ \alpha_1 \in W_1$.

$\therefore$ $\qquad f_1 \in W_1^0$.

Similarly we can show that f_2 is a linear functional on V and $f_2 \in W_2^0$.

Now we claim that $f = f_1 + f_2$.

Let α be any element in V. Let $\alpha = \alpha_1 + \alpha_2$, where $\alpha_1 \in W_1, \alpha_2 \in W_2$.

Then $\quad (f_1 + f_2)(\alpha) = f_1(\alpha) + f_2(\alpha) = f(\alpha_2) + f(\alpha_1)$ [From (1) and (2)]

$$= f(\alpha_1) + f(\alpha_2) = f(\alpha_1 + \alpha_2)$$

[$\because f$ is linear functional]

$$= f(\alpha).$$

Thus $\quad (f_1 + f_2)(\alpha) = f(\alpha) \ \forall \ \alpha \in V$.

$\therefore$ $\quad f = f_1 + f_2$.

Thus $f \in V' \Rightarrow f = f_1 + f_2$ where $f_1 \in W_1^0, f_2 \in W_2^0$.

$\therefore$ $\quad V' = W_1^0 + W_2^0$.

Hence $\quad V' = W_1^0 \oplus W_2^0$.

8 Invariant Direct-Sum Decompositions

Let T be a linear operator on a vector space $V(F)$. If S is a non-empty subset of V, then by $T(S)$ we mean the set of those elements of V which are images under T of the elements in S. Thus

$$T(S) = \{T(\alpha) \in V : \alpha \in S\}.$$

Obviously $T(S) \subseteq V$. We call it the **image** of S under T.

Invariance : Definition. *Let V be a vector space and T a linear operator on V. If W is a subspace of V, we say that W is invariant under T if*

$$\alpha \in W \Rightarrow T(\alpha) \in W.$$

Illustration 1: *If T is any linear operator on V, then V is invariant under T. If $\alpha \in V$, then $T(\alpha) \in V$ because T is a linear operator on V. Thus V is invariant under T.*

The zero subspace of V is also invariant under T. The zero subspace contains only one vector *i.e.*, $\mathbf{0}$ and we know that $T(\mathbf{0}) = \mathbf{0}$ which is in zero subspace.

Illustration 2: *Let $V(F)$ be the vector space of all polynomials in x over the field F and let D be the differentiation operator on V. Let W be the subspace of V consisting of all polynomials of degree not greater than n.*

If $f(x) \in W$, then $D[f(x)] \in W$ because differentiation operator D is degree decreasing. Therefore W is invariant under D.

Let W be a subspace of the vector space V and let W be invariant under the linear operator T on V *i.e.*, let

$$\alpha \in W \Rightarrow T(\alpha) \in W.$$

We know that W itself is a vector space. If we ignore the fact that T is defined outside W, then we may regard T as a linear operator on W. Thus the linear operator T induces a linear operator T_W on the vector space W defined by

$$T_W(\alpha) = T(\alpha) \ \forall \ \alpha \in W.$$

It should be noted that T_W is quite a different object from T because the domain of T_W is W while the domain of T is V.

Invariance can be considered for several linear transformations also. Thus W is invariant under a set of linear transformations if it is invariant under each member of the set.

9 Matrix Interpretation of Invariance

Let V be a finite dimensional vector space over the field F and let T be a linear operator on V. Suppose V has a subspace W which is invariant under T. Then we can choose suitable ordered basis B for V so that the matrix of T with respect to B takes some particular simple form.

Let $B_1 = \{\alpha_1, \ldots, \alpha_m\}$ be an ordered basis for W where $\dim W = m$. We can extend B_1 to form a basis for V. Let

$$B = \{\alpha_1, \ldots, \alpha_m, \alpha_{m+1}, \ldots, \alpha_n\}$$

be an ordered basis for V where $\dim V = n$.

Let $A = [a_{ij}]_{n \times n}$ be the matrix of T with respect to the ordered basis B. Then

$$T(\alpha_j) = \sum_{i=1}^{n} a_{ij} \alpha_i, \ j = 1, 2, \ldots, n. \qquad \ldots(1)$$

If $1 \le j \le m$, then α_j is in W. But W is invariant under T. Therefore if $1 \le j \le m$, then $T(\alpha_j)$ is in W and so it can be expressed as a linear combination of the vectors $\alpha_1, \ldots, \alpha_m$, which form a basis for W. This means that

$$T(\alpha_j) = \sum_{i=1}^{m} a_{ij}\, \alpha_i, 1 \le j \le m. \qquad \ldots(2)$$

In other words in the relation (1), the scalars a_{ij} are all zero if $1 \le j \le m$ and $m + 1 \le i \le n$. Therefore the matrix A takes the simple form

$$A = \begin{bmatrix} M & C \\ O & D \end{bmatrix}$$

where M is an $m \times m$ matrix, C is an $m \times (n-m)$ matrix, O is the null matrix of the type $(n-m) \times m$ and D is an $(n-m) \times (n-m)$ matrix.

From the relation (2) it is obvious that the matrix M is nothing but the matrix of the induced operator T_W on W relative to the ordered basis B_1 for W.

10 Reducibility

Definition: *Let W_1 and W_2 be two subspaces of a vector space V and let T be a linear operator on V. Then T is said to be reduced by the pair (W_1, W_2) if*

(i) $V = W_1 \oplus W_2$,

(ii) both W_1 and W_2 are invariant under T.

It should be noted that if a subspace W_1 of V is invariant under T, then there are many ways of finding a subspace W_2 of V such that $V = W_1 \oplus W_2$, but it is not necessary that some W_2 will also be invariant under T. In other words among the collection of all subspaces invariant under T we may not be able to select any two other than V and the zero subspace with the property that V is their direct sum .

The definition of reducibility can be extended to more than two subspaces. Thus let $W_1, \ldots, W_k$ be k subspaces of a vector space V and let T be a linear operator on V. Then T is said to be reduced by $(W_1, \ldots, W_k)$ if

(i) V is the direct sum of the subspaces $W_1, \ldots, W_k$,

and (ii) Each of the subspaces W_i is invariant under T.

11 Direct Sum of Linear Operators

Definition: Suppose T is a linear operator on the vector space V. Let

$$V = W_1 \oplus \ldots \oplus W_k$$

be a direct sum decomposition of V in which each subspace W_i is invariant under T. Then T induces a linear operator T_i on each W_i by restricting its domain from V to W_i. If $\alpha \in V$, then there exist unique vectors $\alpha_1, \ldots, \alpha_k$ with α_i in W_i such that

$$\alpha = \alpha_1 + \ldots + \alpha_k$$

$\Rightarrow \quad T(\alpha) = T(\alpha_1 + \ldots + \alpha_k)$

$\Rightarrow \quad T(\alpha) = T(\alpha_1) + \ldots + T(\alpha_k)$ [$\because$ T is linear]

$\Rightarrow \quad T(\alpha) = T_1(\alpha_1) + \ldots + T_k(\alpha_k)$ [$\because$ If $\alpha_i \in W_i$, then by def. of T_i, we have $T(\alpha_i) = T_i(\alpha_i)$]

Thus we can find the action of T on V with the help of independent action of the operators T_i on the subspaces W_i. In such situation we say that the operator T is the **direct sum of the operators $T_1, \ldots, T_k$. It should be noted carefully that T is a linear operator on V, while the T_i are linear operators on the various subspaces W_i.**

12 Matrix Representation of Reducibility

If T is a linear operator on a finite dimensional vector space V and is reduced by the pair (W_1, W_2), then by choosing a suitable basis B for V, we can give a particularly simple form to the matrix of T with respect to B.

Let $\dim V = n$ and $\dim W_1 = m$. Then $\dim W_2 = n - m$ since V is the direct sum of W_1 and W_2.

Let $B_1 = \{\alpha_1, \ldots, \alpha_m\}$ be a basis for W_1 and $B_2 = \{\alpha_{m+1}, \ldots, \alpha_n\}$ be a basis for W_2. Then $B = B_1 \cup B_2 = \{\alpha_1, \ldots, \alpha_m, \alpha_{m+1}, \ldots, \alpha_n\}$ is a basis for V.

It can be easily seen, as in the case of invariance, that

$$[T]_B = \begin{bmatrix} M & O \\ O & N \end{bmatrix}$$

where M is an $m \times m$ matrix, N is an $(n - m) \times (n - m)$ matrix and O are null matrices of suitable sizes.

Also if T_1 and T_2 are linear operators induced by T on W_1 and W_2 respectively, then

$$M = [T_1]_{B_1}, \text{ and } N = [T_2]_{B_2}.$$

Illustrative Examples

Example 12: *If T is a linear operator on a vector space V and if W is any subspace of V, then $T(W)$ is a subspace of V. Also W is invariant under T iff $T(W) \subseteq W$.*

Solution: We have, by definition

$$T(W) = \{T(\alpha) : \alpha \in W\}.$$

Since $\mathbf{0} \in W$ and $T(\mathbf{0}) = \mathbf{0}$, therefore $T(W)$ is not empty because at least $\mathbf{0} \in T(W)$.

Now let $T(\alpha_1), T(\alpha_2)$ be any two elements of $T(W)$ where α_1, α_2 are any two elements of W.

If $a, b \in F$, then

$$aT(\alpha_1) + bT(\alpha_2) = T(a\alpha_1 + b\alpha_2), \text{ because } T \text{ is linear.}$$

But W is a subspace of V.

Therefore $\alpha_1, \alpha_2 \in W$ and $a, b \in F \Rightarrow a\alpha_1 + b\alpha_2 \in W$. Consequently

$$T(a\alpha_1 + b\alpha_2) \in T(W).$$

Thus $\quad a, b \in F$ and $T(\alpha_1), T(\alpha_2) \in T(W)$

$\Rightarrow \quad aT(\alpha_1) + bT(\alpha_2) \in T(W)$.

$\therefore$ $T(W)$ is a subspace of V.

Second Part: Suppose W is invariant under T.

Let $T(\alpha)$ be any element of $T(W)$ where $\alpha \in W$.

Since $\alpha \in W$ and W is invariant under T, therefore

$$T(\alpha) \in W. \text{ Thus } T(\alpha) \in T(W) \Rightarrow T(\alpha) \in W.$$

Therefore $\quad T(W) \subseteq W$.

Conversely suppose that $T(W) \subseteq W$.

Then $T(\alpha) \in W \ \forall \ \alpha \in W$. Therefore W is invariant under T.

Example 13: *If T is any linear operator on a vector space V, then the range of T and the null space of T are both invariant under T.*

Solution: Let $N(T)$ be the null space of T. Then

$$N(T) = \{\alpha \in V : T(\alpha) = \mathbf{0}\}.$$

If $\beta \in N(T)$, then $T(\beta) = \mathbf{0} \in N(T)$ because $N(T)$ is a subspace.

$\therefore$ $N(T)$ is invariant under T.

Again let $R(T)$ be the range of T. Then

$$R(T) = \{T(\alpha) \in V : \alpha \in V\}.$$

Since $R(T)$ is a subset of V, therefore $\beta \in R(T) \Rightarrow \beta \in V$.

Now $\quad \beta \in V \Rightarrow T(\beta) \in R(T)$.

Thus $\beta \in R(T) \Rightarrow T(\beta) \in R(T)$. Therefore $R(T)$ is invariant under T.

Example 14: *Give an example of a linear transformation T on a finite-dimensional vector space V such that V and the zero subspace are the only subspaces invariant under T.*

Solution: Let T be the linear operator on $V_2(\mathbf{R})$ which is represented in the standard ordered basis by the matrix

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

Let W be a proper subspace of $V_2(\mathbf{R})$ which is invariant under T. Then W must be of dimension 1. Let W be the subspace spanned by some non-zero vector α. Now $\alpha \in W$ and W is invariant under T. Therefore $T(\alpha) \in W$.

$\therefore$ $T(\alpha) = c\alpha$ for some $c \in \mathbf{R}$

$\Rightarrow \quad T(\alpha) = cI(\alpha)$ where I is identity operator on V

$\Rightarrow \quad [T - cI](\alpha) = \mathbf{0}$

$\Rightarrow \quad T - cI$ is singular $\quad [\because \ \alpha \neq \mathbf{0}]$

$\Rightarrow \quad T - cI$ is not invertible.

If B denotes the standard ordered basis for $V_2(\mathbf{R})$, then

$$[T - cI]_B = [T]_B - c\,[I]_B$$

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

Now det $\begin{bmatrix} -c & -1 \\ 1 & -c \end{bmatrix} = \begin{vmatrix} -c & -1 \\ 1 & -c \end{vmatrix} = c^2 + 1 \neq 0$ for any real number c.

$\therefore \quad \begin{bmatrix} -c & -1 \\ 1 & -c \end{bmatrix}$ *i.e.*, $[T - cI]_B$ is invertible.

Consequently $T - cI$ is invertible which is contradictory to the result that $T - cI$ is not invertible.

Hence no proper subspace W of V_2 (**R**) can be invariant under T.

13 The Adjoint or the Transpose of a Linear Transformation

In order to bring some simplicity in our work we shall introduce a few changes in our notation of writing the image of an element of a vector space under a linear transformation and that under a linear functional. If T is a linear transformation on a vector space V and $\alpha \in V$, then in place of writing $T(\alpha)$ we shall simply write $T\alpha$ *i.e.*, we shall omit the brackets. Thus $T\alpha$ will mean the image of α under T. If T_1 and T_2 are two linear transformations of V, then in our new notation $T_1 T_2 \alpha$ will stand for $T_1 [T_2 (\alpha)]$.

Let f be a linear functional on V. If $\alpha \in V$, then in place of writing $f(\alpha)$ we shall write $[\alpha, f]$. This is the square brackets notation to write the image of a vector under a linear functional. Thus $[\alpha, f]$ will stand for $f(\alpha)$. If $a, b \in F$ and $\alpha, \beta \in V$, then in this new notation the linearity property of f

i.e., $\quad f(a\alpha + b\beta) = af(\alpha) + bf(\beta)$

will be written as

$$[a\alpha + b\beta, f] = a[\alpha, f] + b[\beta, f].$$

Also if f and g are two linear functionals on V and $a, b \in F$, then the property defining addition and scalar multiplication of linear functionals *i.e.*, the property

$$(af + bg)(\alpha) = af(\alpha) + bg(\alpha)$$

will be written as

$$[\alpha, af + bg] = a[\alpha, f] + b[\alpha, g].$$

Note that in this new notation, we get

$$[\alpha, f] = f(\alpha), [\alpha, g] = g(\alpha).$$

Theorem 1: *Let U and V be vector spaces over the field F. For each linear transformation T from U into V, there is a unique linear transformation T' from V' into U' such that*

$$[T'(g)](\alpha) = g[T(\alpha)] \qquad \textit{(in old notation)}$$

or

$$[\alpha, T'g] = [T\alpha, g] \qquad \textit{(in new notation)}$$

for every g in V' and α in U.

The linear transformation T' is called the **adjoint** or the **transpose** or the **dual** of T. In some books it is denoted by T^t or by T^*.

Proof : T is a linear transformation from U to V. U' is the dual space of U and V' is the dual space of V. Suppose $g \in V'$ *i.e.*, g is a linear functional on V. Let us define

$$f(\alpha) = g[T(\alpha)] \;\forall\; \alpha \in U \qquad \ldots(1)$$

Then f is a function from U into F. We see that f is nothing but the product or composite of the two functions T and g where $T: U \to V$ and $g: V \to F$. Since both T and g are linear therefore f is also linear. Thus f is a linear functional on U *i.e.*, $f \in U'$. In this way T provides us with a rule T' which associates with each functional g on V a linear functional

$$f = T'(g) \text{ on } U, \text{ defined by (1).}$$

Thus $\quad T': V' \to U'$ such that

$$T'(g) = f \;\forall\; g \in V' \text{ where } f(\alpha) = g[T(\alpha)] \;\forall\; \alpha \in U.$$

Putting $f = T'(g)$ in (1), we see that T' is a function from V' into U' such that

$$[T'(g)](\alpha) = g[T(\alpha)]$$

or in square brackets notation

$$[\alpha, T'g] = [T\alpha, g] \;\forall\; g \in V' \text{ and } \forall\; \alpha \in U. \qquad \ldots(2)$$

Now we shall show that T' is a linear transformation from V' into U'. Let $g_1, g_2 \in V'$ and $a, b \in F$.

Then we are to prove that

$$T'(ag_1 + bg_2) = aT'g_1 + bT'g_2 \qquad \ldots(3)$$

where $T'g_1$ stands for $T'(g_1)$ and $T'g_2$ stands for $T'(g_2)$.

We see that both the sides of (3) are elements of U' *i.e.*, both are linear functionals on U. So if α is any element of U, we have

$[\alpha, T'(ag_1 + bg_2)] = [T\alpha, ag_1 + bg_2]$ [From (2) because $ag_1 + bg_2 \in V'$]

$= [T\alpha, ag_1] + [T\alpha, bg_2]$ [By def. of addition in V']

$= a[T\alpha, g_1] + b[Ta, g_2]$

[by def. of scalar multiplication in V']

$= a[\alpha, T'g_1] + b[\alpha_1, T'g_2]$ [From (2)]

$= [\alpha, aT'g_1] + [\alpha, bT'g_2]$

[By def. of scalar multiplication in U'. Note that $T'g_1, T'g_2 \in U'$]

$= [\alpha, aT'g_1 + bT'g_2]$ [By addition in U']

Thus $\forall\; \alpha \in U$, we have

$$[\alpha, T'(ag_1 + bg_2)] = [\alpha, aT'g_1 + bT'g_2].$$

$\therefore \quad T'(ag_1 + bg_2) = aT'g_1 + bT'g_2$ [By def. of equality of two functions]

Hence T' is a linear transformation from V' into U'.

Now let us show that T' is uniquely determined for a given T. If possible, let T_1 be a linear transformation from V' into U' such that

$$[\alpha, T_1 g] = [T\alpha, g] \;\forall\; g \in V' \text{ and } \alpha \in U. \qquad \ldots(4)$$

Then from (2) and (4), we get

$$[\alpha, T_1\, g] = [\alpha, T'\, g] \;\; \forall\, \alpha \in U, \forall\; g \in V'$$

$\Rightarrow \qquad T_1\, g = T'\, g \;\; \forall\; g \in V'$

$\Rightarrow \qquad T_1 = T'.$

$\therefore$ T' is uniquely determined for each T. Hence the theorem.

Note: If T is a linear transformation on the vector space V, then in the proof of the above theorem we should simply replace U by V.

Theorem 2 : *If T is a linear transformation from a vector space U into a vector space V, then*

(i) the annihilator of the range of T is equal to the null space of T' i.e.,

$$[R(T)]^0 = N(T').$$

If in addition U and V are finite dimensional, then

(ii) $\qquad \rho(T') = \rho(T)$

and (iii) the range of T' is the annihilator of the null space of T i.e.,

$$R(T') = [N(T)]^0.$$

Proof : (i) If $g \in V'$, then by definition of T', we have

$$[\alpha, T'\, g] = [T\alpha, g] \;\; \forall\, \alpha \in U. \qquad \ldots(1)$$

Let $g \in N(T')$ which is a subspace of V'. Then

$T'\, g = \hat{\mathbf{0}}$ where $\hat{\mathbf{0}}$ is zero element of U' *i.e.*, $\hat{\mathbf{0}}$ is zero functional on U. Therefore from (1), we get

$$[T\alpha, g] = [\alpha, \hat{\mathbf{0}}] \;\; \forall\; \alpha \in U$$

$\Rightarrow \qquad [T\alpha, g] = 0 \;\; \forall\; \alpha \in U \qquad [\because \hat{\mathbf{0}}(\alpha) = 0 \;\; \forall\; \alpha \in U]$

$\Rightarrow \qquad g(\beta) = 0 \;\; \forall\; \beta \in R(T) \qquad [\because R(T) = \{\beta \in V : \beta = T(\alpha) \text{ for some } \alpha \in U\}]$

$\Rightarrow \qquad g \in [R(T)]^0.$

$\therefore \qquad N(T') \subseteq [R(T)]^0.$

Now let $g \in [R(T)]^0$ which is a subspace of V'. Then

$$g(\beta) = 0 \;\; \forall\; \beta \in R(T)$$

$\Rightarrow \qquad [T\alpha, g] = 0 \;\; \forall\; \alpha \in U \qquad [\because \forall\; \alpha \in U, T\alpha \in R(T)]$

$\Rightarrow \qquad [\alpha, T'\, g] = 0 \;\; \forall\; \alpha \in U \qquad$ [From (1)]

$\Rightarrow \qquad T'\, g = \hat{\mathbf{0}}$ (zero functional on U) $\Rightarrow \; g \in N(T').$

$\therefore \qquad [R(T)]^0 \subseteq N(T').$

Hence $\qquad [R(T)]^0 = N(T').$

(ii) Suppose U and V are finite dimensional. Let $\dim U = n$, $\dim V = m$. Let $r = \rho(T) =$ the dimension of $R(T)$.

Now $R(T)$ is a subspace of V. Therefore

$$\dim R(T) + \dim [R(T)]^0 = \dim V. \qquad \text{[See Th. 2 of article 6]}$$

$\therefore \qquad \dim [R(T)]^0 = \dim V - \dim R(T)$

$\qquad\qquad = \dim V - r = m - r.$

By part (i) of this theorem $[R(T)]^0 = N(T')$.

$\therefore \qquad \dim N(T') = m - r \quad \Rightarrow \quad$ nullity of $T' = m - r$.

But T' is a linear transformation from V' into U'.

$\therefore \qquad \rho(T') + \nu(T') = \dim V'$

or $\qquad \rho(T') = \dim V' - \nu(T')$

$\qquad\qquad = \dim V -$ nullity of $T' = m - (m - r) = r.$

$\therefore \qquad \rho(T) = \rho(T') = r.$

(iii) T' is a linear transformation from V' into U'. Therefore $R(T')$ is a subspace of U'. Also $[N(T)]^0$ is a subspace of U' because $N(T)$ is a subspace of U. First we shall show that $R(T') \subseteq [N(T)]^0$.

Let $\qquad f \in R(T')$.

Then $\qquad f = T'g$ for some $g \in V'$.

If α is any vector in $N(T)$, then $T\alpha = \mathbf{0}$.

We have

$$[\alpha, f] = [\alpha, T'g] = [T\alpha, g] = [\mathbf{0}, g] = 0.$$

Thus $\qquad f(\alpha) = 0 \ \forall \ \alpha \in N(T)$.

Therefore $f \in [N(T)]^0$.

$\therefore \qquad R(T') \subseteq [N(T)]^0$

$\Rightarrow \qquad R(T')$ is a subspace of $[N(T)]^0$.

Now $\qquad \dim N(T) + \dim [N(T)]^0 = \dim U. \qquad$ [Theorem 2 of article 6]

$\therefore \qquad \dim [N(T)]^0 = \dim U - \dim N(T)$

$\qquad\qquad = \dim R(T) \qquad [\because \dim U = \dim R(T) + \dim N(T)]$

$\qquad\qquad = \rho(T) = \rho(T') = \dim R(T').$

Thus $\qquad \dim R(T') = \dim [N(T)]^0$

and $\qquad R(T') \subseteq [N(T)]^0$.

$\therefore \qquad R(T') = [N(T)]^0$.

Note: If T is a linear transformation on a vector space V, then in the proof of the above theorem we should replace U by V and m by n.

Theorem 3: *Let U and V be finite-dimensional vector spaces over the field F. Let B be an ordered basis for U with dual basis B', and let B_1 be an ordered basis for V with dual basis B_1'. Let T be a linear transformation from U into V. Let A be the matrix of T relative to B, B_1 and let C be the matrix of T' relative to B_1', B'. Then $C = A'$ i.e., the matrix C is the transpose of the matrix A.*

Proof : Let $\dim U = n$, $\dim V = m$.

Let $\qquad B = \{\alpha_1, \ldots, \alpha_n\}, B' = \{f_1, \ldots, f_n\},$

$$B_1 = \{\beta_1, \dots, \beta_m\},\ B_1' = \{g_1, \dots, g\}.$$

Now T is a linear transformation from U into V and T' is that from V' into U'. The matrix A of T relative to B, B_1 will be of the type $m \times n$. If $A = [a_{ij}]_{m \times n}$, then by definition

$$T(\alpha_j) \text{ or simply } T\alpha_j = \sum_{i=1}^{m} a_{ij}\, \beta_i,\ j = 1, 2 \dots, n. \qquad \dots(1)$$

The matrix C of T' relative to B_1', B' will be of the type $n \times m$. If $C = [c_{ji}]_{n \times m}$, then by definition

$$T'(g_i) \text{ or simply } T' g_i = \sum_{j=1}^{n} c_{j\,i}\, f_j,\ i = 1, 2, \dots, m \qquad \dots(2)$$

Now $T' g_i$ is an element of U' *i.e.*, $T' g_i$ is a linear functional on U. If f is any linear functional on U, then we know that

$$f = \sum_{j=1}^{n} f(\alpha_j)\, f_j. \qquad \text{[See theorem 3, article 4]}$$

Applying this formula for $T' g_i$ in place of f, we get

$$T' g_i = \sum_{j=1}^{n} \{(T' g_i)(\alpha_j)\}\, f_j. \qquad \dots(3)$$

Now let us find $(T' g_i)(\alpha_j)$.

We have $(T' g_i)(\alpha_j) = g_i\, T(\alpha_j)$ [By def. of T']

$= g_i \left(\sum_{k=1}^{m} a_{k\,j}\, \beta_k \right)$ [From (1), replacing the suffix i by k which is immaterial]

$= \sum_{k=1}^{m} a_{kj}\, g_i(\beta_k)$ [$\because g_i$ is linear]

$= \sum_{k=1}^{m} a_{kj}\, \delta_{ik}$ [$\because g_i \in B_1'$ which is dual basis of B_1]

$= a_{i\,j}$

[On summing with respect to k and remembering that $\delta_{ik} = 1$ when $k = i$ and $\delta_{ik} = 0$ when $k \neq i$]

Putting this value of $(T' g_i)(\alpha_j)$ in (3), we get

$$T' g_i = \sum_{j=1}^{m} a_{ij}\, f_j. \qquad \dots(4)$$

Since $f_1, \dots, f_n$ are linearly independent, therefore from (2) and (4), we get

$$c_{ji} = a_{ij}.$$

Hence by definition of transpose of a matrix, we have

$$C = A'.$$

Note: If T is a linear transformation on a finite-dimensional vector space V, then in the above theorem we put $U = V$ and $m = n$. Also according to our convention we take $B_1 = B$. The students should write the complete proof themselves.

Theorem 4: *Let A be any $m \times n$ matrix over the field F. Then the row rank of A is equal to the column rank of A.*

Proof : Let $A = [a_{ij}]_{m \times n}$. Let

$$B = \{\alpha_1, \dots, \alpha_n\} \text{ and } B_1 = \{\beta_1, \dots, \beta_m\}$$

be the standard ordered bases for $V_n(F)$ and $V_m(F)$ respectively. Let T be the linear transformation from $V_n(F)$ into $V_m(F)$ whose matrix is A relative to ordered bases B and B_1. Then obviously the vectors $T(\alpha_1), \dots, T(a_n)$ are nothing but the column vectors of the matrix A. Also these vectors span the range of T because $\alpha_1, \dots, \alpha_n$ form a basis for the domain of T *i.e.*, $V_n(F)$.

$\therefore$ the range of T = the column space of A

$\Rightarrow$ the dimension of the range of T = the dimension of the column space of A

$\Rightarrow$ $\rho(T)$ = the column rank of A. ...(1)

If T' is the adjoint of the linear transformation T, then the matrix of T' relative to the dual bases B_1' and B' is the matrix A' which is the transpose of the matrix A. The columns of the matrix A' are nothing but the rows of the matrix A. By the same reasoning as given in proving the result (1), we have

$\rho(T')$ = the column rank of A'

= the row rank of A ...(2)

Since $\rho(T) = \rho(T')$, therefore from (1) and (2), we get the result that

the column rank of A = the row rank of A.

Theorem 5 : *Prove the following properties of adjoints of linear operators on a vector space $V(F)$:*

(i) $\hat{\mathbf{0}}' = \hat{\mathbf{0}}$;

(ii) $I' = I$; **(Kumaun 2007, 09)**

(iii) $(T_1 + T_2)' = T_1' + T_2'$; **(Kumaun 2010)**

(iv) $(T_1 T_2)' = T_2' T_1'$; **(Kumaun 2007, 09, 10)**

(v) $(aT)' = aT'$ *where* $a \in F$;

(vi) $(T^{-1})' = (T')^{-1}$ *if* T *is invertible ;*

(vii) $(T')' = T'' = T$ *if* V *is finite-dimensional.*

Proof : (i) If $\hat{\mathbf{0}}$ is the zero transformation on V, then by the definition of the adjoint of a linear transformation, we have

$[\alpha, \hat{\mathbf{0}}' g] = [\hat{\mathbf{0}} \alpha, g]$ for every g in V' and α in V

= for every g in V' $[\because \hat{\mathbf{0}}(\alpha) = \mathbf{0} \ \forall \ \alpha \in V]$

$= 0$ $[\because g(\mathbf{0}) = 0]$

$= [\alpha, \hat{\mathbf{0}}] \ \forall \ \alpha \in V$ [Here $\hat{\mathbf{0}} \in V'$ and $\hat{\mathbf{0}}(\alpha) = 0$]

$= [\alpha, \hat{\mathbf{0}} g] \ \forall \ g \in V'$ and $\alpha \in V$

[Here $\hat{\mathbf{0}}$ is the zero transformation on V']

Thus we have

$$[\alpha, \hat{\mathbf{0}}' g] = [\alpha, \hat{\mathbf{0}} g] \text{ for all } g \text{ in } V' \text{ and } \alpha \text{ in } V.$$

$$\therefore \quad \hat{\mathbf{0}}' = \hat{\mathbf{0}}.$$

(ii) If I is the identity transformation on V, then by the definition of the adjoint of a linear transformation, we have

$$[\alpha, I' g] = [I\alpha, g] \text{ for every } g \text{ in } V' \text{ and } \alpha \text{ in } V$$

$$= [\alpha, g] \text{ for every } g \text{ in } V' \text{ and } \alpha \text{ in } V \qquad [\because I(\alpha) = \alpha \ \forall \ \alpha \in V]$$

$$= [\alpha, Ig] \text{ for every } g \text{ in } V' \text{ and } \alpha \text{ in } V$$

[Here I is the identity operator on V']

$$\therefore \quad I' = I.$$

(iii) If T_1, T_2 are linear operators on V, then $T_1 + T_2$ is also a linear operator on V. By the definition of adjoint, we have

$$[\alpha, (T_1 + T_2)' g] = [(T_1 + T_2)\alpha, g] \text{ for every } g \text{ in } V' \text{ and } \alpha \text{ in } V$$

$$= [T_1 \alpha + T_2 \alpha, g]$$

[By def. of addition of linear transformations]

$$= [T_1 \alpha, g] + [T_2 \alpha, g] \qquad \text{[By linearity property of } g]$$

$$= [\alpha, T_1' g] + [\alpha, T_2' g] \qquad \text{[By def. of adjoint]}$$

$$= [\alpha, T_1' g + T_2' g]$$

[By def. of addition of linear functionals. Note that $T_1' g, T_2' g$ are elements of V']

$$= [\alpha, (T_1' + T_2') g].$$

Thus we have

$$[\alpha, (T_1 + T_2)' g] = [a, (T_1' + T_2') g] \text{ for every } g \text{ in } V' \text{ and } \alpha \text{ in } V.$$

$$\therefore \quad (T_1 + T_2)' g = (T_1 + T_2') g \ \forall \ g \in V'.$$

$$\therefore \quad (T_1 + T_2)' = T_1' + T_2'.$$

(iv) If T_1, T_2 are linear operators on V, then $T_1 T_2$ is also a linear operator on V. By the definition of adjoint, we have

$$[\alpha, (T_1 T_2)' g] = [(T_1 T_2)\alpha, g] \text{ for every } g \text{ in } V' \text{ and } \alpha \text{ in } V$$

$$= [(T_1) T_2 \alpha, g]$$

[By def. of product of linear transformations]

$$= [T_2 \alpha, T_1' g] \qquad \text{[By def. of adjoint]}$$

$$= [\alpha, T_2' T_1' g] \qquad \text{[By def. of adjoint]}$$

Thus we have

$$[\alpha, (T_1 T_2)' g] = [a, T_2' T_1' g] \text{ for every } g \text{ in } V' \text{ and } \alpha \text{ in } V.$$

$$\therefore \quad (T_1 T_2)' = T_2' T_1'.$$

Note: This is called the reversal law for the adjoint of the product of two linear transformations.

(v) If T is a linear operator on V and $a \in F$, then aT is also a linear operator on V. By the definition of the adjoint, we have

$$[\alpha, (aT)' g] = [(aT)\alpha, g] \text{ for every } g \text{ in } V' \text{ and } \alpha \text{ in } V$$

$= [a(T\alpha), g]$ [By def. of scalar multiplication of a linear transformation]

$= a[T\alpha, g]$ [$\because$ g is linear]

$= a[\alpha, T' g]$ [By def. of adjoint]

$= [\alpha, a(T' g)]$ [By def. of scalar multiplication in V'. Note that $T' g \in V'$]

$= [\alpha, (aT') g]$

[By def. of scalar multiplication of T' by a]

$\therefore$ $(aT)' = aT'$.

(vi) Suppose T is an invertible linear operator on V. If T^{-1} is the inverse of T, we have

$$T^{-1} T = I = T T^{-1}$$

$\Rightarrow$ $(T^{-1} T)' = I' = (T T^{-1})'$

$\Rightarrow$ $T'(T^{-1})' = I = (T^{-1})' T'$ [Using results (ii) and (iv)]

$\therefore$ T' is invertible and $(T')^{-1} = (T^{-1})'$.

(vii) V is a finite dimensional vector space. T is a linear operator on V, T' is a linear operator on V' and $(T')'$ or T'' is a linear operator on V''. We have identified V'' with V through natural isomorphism $\alpha \leftrightarrow L_\alpha$ where $\alpha \in V$ and $L_\alpha \in V''$. Here L_α is a linear functional on V' and is such that

$$L_\alpha(g) = g(\alpha) \ \forall \ g \in V'. \quad \ldots(1)$$

Through this natural isomorphism we shall take $\alpha = L_\alpha$ and thus T'' will be regarded as a linear operator on V.

Now T' is a linear operator on V'. Therefore by the definition of adjoint, we have

$$[g, T'' L_\alpha] = [g, (T')' L_\alpha] = [T' g, L_\alpha] \text{ for every } g \in V' \text{ and } \alpha \in V.$$

Now $T' g$ is an element of V'. Therefore from (1), we have

$$[T' g, L_\alpha] = [\alpha, T' g]$$

[Note that from (1), $L_\alpha(T' g) = (T' g)\alpha$]

$= [T\alpha, g]$. [By def. of adjoint]

Again $T'' L_\alpha$ is an element of V''. Therefore from (1), we have

$[g, T'' L_\alpha] = [\beta, g]$ where $\beta \in V$ and $\beta \leftrightarrow T'' L_\alpha$ under natural isomorphism

$= [T''\alpha, g]$ $[\because \beta = T'' L_\alpha = T''\alpha$ when we regard T'' as linear operator on V in place of $V'']$

Thus, we have

$[T\alpha, g] = [T''\alpha, g]$ for every g in V' and α in V

$\Rightarrow$ $g(T\alpha) = g(T''\alpha)$ for every g in V' and α in V

$\Rightarrow$ $g(T\alpha - T''\alpha) = 0$ for every g in V' and α in V

$\Rightarrow$ $T\alpha - T''\alpha = \mathbf{0}$ for every α in V

$\Rightarrow$ $(T - T'')\alpha = \mathbf{0}$ for every α in V

$\Rightarrow$ $T - T'' = \hat{\mathbf{0}}$

$\Rightarrow$ $T = T''$.

Illustrative Examples

Example 15: *Let f be the linear functional on $\mathbf{R}^2$ defined by $f(x, y) = 2x - 5y$. For each linear mapping*

$T : \mathbf{R}^3 \to \mathbf{R}^2$ *find* $[T'(f)](x, y, z)$, *where*

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

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

(iii) $T(x, y, z) = (x + y, 0)$.

Solution: By definition of the transpose mapping,

$T'(f) = f \circ T$ i.e., $[T'(f)]\alpha = f[T(\alpha)]$ for every $\alpha \in \mathbf{R}^3$.

(i) $[T'(f)](x, y, z) = f[T(x, y, z)]$

$= f(x - y, y + z) = 2(x - y) - 5(y + z) = 2x - 7y - 5z$.

(ii) $[T'(f)](x, y, z) = f[T(x, y, z)]$

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

$= 2(x + y + 2z) - 5(2x + y) = -8x - 3y + 4z$.

(iii) $[T'(f)](x, y, z) = f[T(x, y, z)]$

$= f(x + y, 0) = 2(x + y) - 5.0 = 2x + 2y$.

Example 16: *If A and B are similar linear transformations on a vector space V, then so also are A' and B'.*

Solution: A is similar to B means that there exists an invertible linear transformation C on V such that

$A = CBC^{-1}$

$\Rightarrow$ $A = (CBC^{-1})'$

$\Rightarrow$ $A' = (C^{-1})' B' C'$.

Now C is invertible implies that C' is also invertible and

$$(C')^{-1} = (C^{-1})'.$$

$\therefore$ $\quad A' = (C')^{-1} B' C' \quad \Rightarrow \quad C' A' (C')^{-1} = B'$

[Multiplying on right by $(C')^{-1}$ and on left by C']

$\Rightarrow$ B' is similar to A'

$\Rightarrow$ A' and B' are similar.

Example 17: *Let V be a finite dimensional vector space over the field F. Show that $T \to T'$ is an isomorphism of $L(V, V)$ onto $L(V', V')$.*

Solution: Let $\dim V = n$.

Then $\quad \dim V' = n.$

Also $\quad \dim L(V, V) = n^2, \dim L(V', V') = n^2.$

Let $\quad \psi : L(V, V) \to L(V', V')$ such that

$$\psi(T) = T' \ \forall \ T \in L(V, V).$$

(i) ψ is linear transformation:

Let $\quad a, b \in F$ and $T_1, T_2 \in L(V, V)$. Then

$\psi(aT_1 + bT_2) = (aT_1 + bT_2)'$ [By def. of ψ]

$= (aT_1)' + (bT_2)'$ [$\because (A + B)' = A' + B'$]

$= aT_1' + bT_2'$ [$\because (aA)' = aA'$]

$= a\psi(T_1) + b\psi(T_2)$ [By def. of ψ]

$\therefore$ ψ is a linear transformation from $L(V, V)$ into $L(V', V')$.

(ii) ψ is one-one:

Let $\quad T_1, T_2 \in L(V, V)$.

Then $\quad \psi(T_1) = \psi(T_2)$

$\Rightarrow \quad T_1' = T_2'$

$\Rightarrow \quad T_1'' = T_2''$

$\Rightarrow \quad T_1 = T_2$ [$\because V$ is finite-dimensional]

$\therefore$ ψ is one-one.

(iii) ψ is onto:

We have $\dim L(V, V) = \dim L(V', V') = n^2$.

Since ψ is a linear transformation from $L(V, V)$ into $L(V', V')$ therefore ψ is one-one implies that ψ must be onto.

Hence ψ is an isomorphism of $L(V, V)$ onto $L(V', V')$.

Example 18: *If A and B are linear transformations on an n-dimensional vector space V, then prove that*

(i) $\rho(AB) \geq \rho(A) + \rho(B) - n.$

(ii) $\nu(AB) \leq \nu(A) + \nu(B).$ **(Sylvester's law of nullity)**

Solution: (i) First we shall prove, that if T is a linear transformation on V and W_1 is an h-dimensional subspace of V, then the dimension of $T(W_1)$ is $\geq h - \nu(T)$.

Since V is finite-dimensional, therefore the subspace W_1 will possess complement. Let $V = W_1 \oplus W_2$. Then

$$\dim W_2 = n - h = k \text{ (say).}$$

Since $V = W_1 + W_2$, therefore

$$T(V) = T(W_1) + T(W_2), \text{ as can be easily seen.}$$

$\therefore$ $\dim T(V) = \dim [T(W_1) + T(W_2)]$

$\leq \dim T(W_1) + \dim T(W_2)$

[$\because$ The dimension of a sum is $\leq$ the sum of the dimensions]

But $T(V) =$ the range of T.

$\therefore$ $\dim T(V) = \rho(T)$.

Thus $\dim T(W_1) + \dim T(W_2) \geq \rho(T)$. ...(1)

Now $T(W_2)$ is a subspace of W_2. Therefore

$$\dim W_2 \geq \dim T(W_2). \quad ...(2)$$

From (1) and (2), we get

$$\dim T(W_1) + \dim W_2 \geq \rho(T)$$

$\Rightarrow$ $\dim T(W_1) \geq \rho(T) - \dim W_2$

$\Rightarrow$ $\dim T(W_1) \geq n - \nu(T) - k$ [$\because$ $\rho(T) + \nu(T) = n$]

$\Rightarrow$ $\dim T(W_1) \geq n - k - \nu(T)$

$\Rightarrow$ $\dim T(W_1) \geq h - \nu(T)$. ...(3)

Now taking $T = A$ and $W_1 = B(V)$ in (3), we get

$$\dim A[B(V)] \geq \dim B(V) - \nu(A)$$

$\Rightarrow$ $\dim (AB)(V) \geq \rho(B) - \nu(A)$ [$\because$ $B(V) =$ the range of B]

$\Rightarrow$ $\rho(AB) \geq \rho(B) - [n - \rho(A)]$

$\Rightarrow$ $\rho(AB) \geq \rho(A) + \rho(B) - n$.

(ii) We have $\rho(AB) + \nu(AB) = n$.

$\therefore$ $\rho(AB) = n - \nu(AB)$.

But $\rho(AB) \geq \rho(A) + \rho(B) - n$.

$\therefore$ $n - \nu(AB) \geq \rho(A) + \rho(B) - n$

$\Rightarrow$ $\nu(AB) \leq [n - \rho(A)] + [n - \rho(B)]$

$\Rightarrow$ $\nu(AB) \leq \nu(A) + \nu(B)$ [$\because$ $\rho(A) + \nu(A) = n$]

Comprehensive Exercise 2

1. Let V be a finite dimensional vector space over the field F. If W_1 and W_2 are subspaces of V, then $W_1^0 = W_2^0$ iff $W_1 = W_2$.

2. If W_1 and W_2 are subspaces of a finite-dimensional vector space V and if $V = W_1 \oplus W_2$, then
   (i) W_1' is isomorphic to W_2^0. (ii) W_2' is isomorphic to W_1^0.
3. Let W be the subspace of $\mathbf{R}^3$ spanned by $(1, 1, 0)$ and $(0, 1, 1)$. Find a basis of the annihilator of W.
4. Let W be the subspace of $\mathbf{R}^4$ spanned by $(1, 2, -3, 4), (1, 3, -2, 6)$ and $(1, 4, -1, 8)$. Find a basis of the annihilator of W.
5. If the set $S = \{W_i\}$ is the collection of subspaces of a vector space V which are invariant under T, then show that $W = \bigcap_i W_i$ is also invariant under T.
6. Prove that the subspace spanned by two subspaces each of which is invariant under some linear operator T, is itself invariant under T.
7. Let V be a vector space over the field F, and let T be a linear operator on V and let $f(t)$ be a polynomial in the indeterminate t over the field F. If W is the null space of the operator $f(T)$, then W is invariant under T.
8. Let T be the linear operator on $\mathbf{R}^2$, the matrix of which in the standard ordered basis is

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

   (a) Prove that the only subspaces of $\mathbf{R}^2$ invariant under T are $\mathbf{R}^2$ and the zero subspace.
   (b) 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 one dimensional invariant subspaces.
9. Let T be the linear operator on $\mathbf{R}^2$, the matrix of which in the standard ordered basis is

$$\begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}.$$

   If W_1 is the subspace of $\mathbf{R}^2$, spanned by the vector $(1, 0)$, prove that W_1 is invariant under T.
10. Show that the space generated by $(1, 1, 1)$ and $(1, 2, 1)$ is an invariant subspace of $\mathbf{R}^3$ under T, where $T(x, y, z) = (x + y - z,\ x + y, x + y - z)$.
11. If A and B are linear transformations on a finite-dimensional vector space V, then prove that
    (i) $\rho(A + B) \le \rho(A) + \rho(B)$
    (ii) $\rho(AB) \le \min\{\rho(A), \rho(B)\}$.
    (iii) If B is invertible, then $\rho(AB) = \rho(BA) = \rho(A)$.

Answers 2

3. Basis is $\{ f(x, y, z) = x - y + z \}$
4. Basis is $\{ f_1, f_2 \}$ where $f_1(x, y, z, w) = 5x - y + z$, $f_2(x, y, z, w) = 2y - w$

Objective Type Questions

Multiple Choice Questions

Indicate the correct answer for each question by writing the corresponding letter from (a), (b), (c) *and* (d).

1. Let V be an n-dimensional vector space over the field F. The dimension of the dual space of V is
   (a) n (b) n^2
   (c) $\frac{1}{2} n$ (d) none of these
   (Kumaun 2014)

2. If the dual basis of the basis set $B = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ for $V_3(\mathbf{R})$ is $B' = \{ f_1, f_2, f_3 \}$, then
   (a) $f_1(a, b, c) = a$, $f_2(a, b, c) = b$, $f_3(a, b, c) = c$
   (b) $f_1(a, b, c) = b$, $f_2(a, b, c) = c$, $f_3(a, b, c) = a$
   (c) $f_1(a, b, c) = c$, $f_2(a, b, c) = a$, $f_3(a, b, c) = b$
   (d) None of these.

3. If $V(F)$ is a vector space and f be a linear functional from $V \to F$, then
   (a) $f(\mathbf{0}) = 0$ (b) $f(\mathbf{0}) \neq 0$
   (c) $f(0) = \mathbf{0}$ (d) $f(0) \neq \mathbf{0}$

4. If $V(F)$ is a vector space and f is a linear functional from $V \to F$, then
   (a) $f(-x) = f(x)$ (b) $f(-x) = -f(x)$
   (c) $f(-x) \neq -f(x)$ (d) $f(-x) \neq f(x)$

5. Let V be a vector space over a field F. A linear functional on V is a linear mapping from
   (a) F into V (b) V into F
   (c) V into itself (d) none of these

6. If S is any subset of a vector space $V(F)$, then S^0 is a subspace of
   (a) V (b) V'
   (c) V'' (d) none of these.

7. Let V be a finite dimensional vector space over the field F. If S is any subset of V, then S^{00} =

(a) S (b) $L(S)$
(c) $[L(S)]^0$ (d) none of these.

8. Let $V(F)$ be a vector space over the field F. Then the set V' of all linear functional on V is also a vector space over the field F. The vector space V' is called :

(a) Dual space of V (b) Null space of V
(c) Subspace of V (d) none of these

9. Let V be a finite dimensional vector space of dimension n. If V' is the dual space of V then dimension of V' will be:

(a) n^2 (b) $2n$
(c) $\frac{n}{2}$ (d) n

10. Let $V(F)$ be a vector space. A function f from V to F which satisfies the condition $f(a\alpha + b\beta) = af(\alpha) + bf(\beta)\ \forall\ a, b \in F$ and $\alpha, \beta \in V$ is called :

(a) linear transformation (b) linear operator
(c) linear functional (d) none of these

11. If W_1 and W_2 are subspaces of a vector space $V(F)$ and $W_2 \subset W_1$ then

(a) $W_1^0 \subset W_2^0$ (b) $W_1^0 \supset W_2^0$
(c) $W_1^0 = W_2^0$ (d) none of these

12. If S^0 denote the annihilator of subspace S of vector space $V(F)$, then :

(a) $\dim S + \dim S^0 < \dim V$ (b) If $S = V$, $S^0 = V^0$
(c) If S is the zero subspace of V, then $S^0 = V$
(d) $S \subset S^{00}$

13. Let $\{f_1, f_2\}$ be the dual basis for the basis set $\{(2,1), (3,1)\}$ for $\mathbf{R}^2$. Then:

(a) $f_1(a,b) = 3b - a$, $f_2(a,b) = a - 2b$
(b) $f_1(a,b) = a - 3b$, $f_2(a,b) = a - 2b$
(c) $f_1(a,b) = a - 3b$, $f_2(a,b) = 2b - a$
(d) $f_1(a,b) = a - 3b$, $f_2(a,b) = a + 2b$

14. If W_1 and W_2 are subspaces of a vector space $V(F)$, then $(W_1 \cap W_2)^0$ =

(a) $W_1^0 \cup W_2^0$ (b) $W_1^0 \cap W_2^0$
(c) $W_1^0 + W_2^0$ (d) $(W_1 + W_2)^0$

Fill in the Blank(s)

Fill in the blanks "......" so that the following statements are complete and correct.

1. Let V be a vector space over the field F. The dual space of V is also called the ...... space of V.
2. If V is an n-dimensional vector space over the field F, then V is ...... to its dual space V'.
3. Let V be a finite dimensional vector space over the field F, and let W be a subspace of V. Then $\dim W + \dim W^0 = \ldots\ldots$.
4. Let V be a finite dimensional vector space over the field F and let W be a subspace of V. Then $W^{00} = \ldots\ldots$.

True or False

Write 'T' for true and 'F' for false statement.

1. Let $V(F)$ be a vector space. A linear functional on V is a vector valued function.
2. Every finite dimensional vector space is not reflexive.
3. If $V(F)$ is a vector space, then mapping $f : V \to F : f(\alpha) = 0$, then f is a linear functional.
4. If V is finite-dimensional and W is a subspace of V then W' is isomorphic to V'/W^0.
5. If T is a linear transformation from a vector space U into a vector space V then $\rho(T') \neq \rho(T)$.

Answers

Multiple Choice Questions

1.	(a)	2.	(a)	3.	(a)	4.	(b)	5.	(b)
6.	(b)	7.	(b)	8.	(a)	9.	(d)	10.	(c)
11.	(b)	12.	(b)	13.	(a)	14.	(c)		

Fill in the Blank(s)

1. conjugate 2. isomorphic 3. $\dim V$ 4. W

True or False

1. F 2. F 3. T 4. T 5. F

□

Chapter 4

Eigenvalues and Eigenvectors

1 Matric Polynomials

Definition: *An expression of the form*

$$F(\lambda) = \mathbf{A}_0 + \mathbf{A}_1\lambda + \mathbf{A}_2\lambda^2 + \ldots + \mathbf{A}_{m-1}\lambda^{m-1} + \mathbf{A}_m\lambda^m,$$

where $\mathbf{A}_0, \mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_m$ *are all square matrices of the same order, is called a* **Matric polynomial** of *degree m provided* $\mathbf{A}_m$ *is not a null matrix. The symbol* λ is called *indeterminate*. If the order of each of the matric coefficients $\mathbf{A}_0, \mathbf{A}_1, \ldots, \mathbf{A}_m$ is *n*, then we say that the matric polynomial is *n*-rowed. According to this definition of a matric polynomial, each square matrix can be expressed as a matric polynomial with zero degree. For example, if **A** be any square matrix, we can write

$$\mathbf{A} = \lambda^0 \mathbf{A}.$$

Equality of Polynomials: Two matric polynomials are equal iff (if and only if), the coefficients of the like powers of λ are the same.

Theorem: *Every square matrix whose elements are ordinary polynomials in λ, can essentially be expressed as a matric polynomial in λ of degree m, where m is the highest power of λ occurring in any element of the matrix.* We shall illustrate this theorem by the following example.

Consider the matrix

$$\mathbf{A} = \begin{bmatrix} 1 + 2\lambda + 3\lambda^2 & \lambda^2 & 4 - 6\lambda \\ 1 + \lambda^3 & 3 + 4\lambda^2 & 1 - 2\lambda + 4\lambda^3 \\ 2 - 3\lambda + 2\lambda^3 & 5 & 6 \end{bmatrix}$$

in which the highest power of λ occurring in any element is 3. Rewriting each element as a cubic in λ, supplying missing coefficients with zeros, we get

$$\mathbf{A} = \begin{bmatrix} 1 + 2.\lambda + 3.\lambda^2 + 0.\lambda^3 & 0 + 0.\lambda + 1.\lambda^2 + 0.\lambda^3 & 4 - 6.\lambda + 0.\lambda^2 + 0.\lambda^3 \\ 1 + 0.\lambda + 0.\lambda^2 + 1.\lambda^3 & 3 + 0.\lambda + 4.\lambda^2 + 0.\lambda^3 & 1 - 2.\lambda + 0.\lambda^2 + 4.\lambda^3 \\ 2 - 3.\lambda + 0.\lambda^2 + 2.\lambda^3 & 5 + 0.\lambda + 0.\lambda^2 + 0.\lambda^3 & 6 + 0.\lambda + 0.\lambda^2 + 0.\lambda^3 \end{bmatrix}$$

Obviously **A** can be written as the matric polynomial

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 4 \\ 1 & 3 & 1 \\ 2 & 5 & 6 \end{bmatrix} + \lambda \begin{bmatrix} 2 & 0 & -6 \\ 0 & 0 & -2 \\ -3 & 0 & 0 \end{bmatrix} + \lambda^2 \begin{bmatrix} 3 & 1 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \lambda^3 \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 4 \\ 2 & 0 & 0 \end{bmatrix}.$$

2 Characteristic Values and Characteristic Vectors of a Matrix

Let $\mathbf{A} = [a_{ij}]_{n \times n}$ be a given n-rowed square matrix. Let

$$\mathbf{A} = \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ \dots \\ x_n \end{bmatrix}$$

be a column vector. Consider the vector equation

$$\mathbf{AX} = \lambda \mathbf{X}. \qquad ..(1)$$

where λ is a scalar (*i.e.,* number).

It is obvious that the zero vector **X** = **O** is a solution of (1) for any value of λ. Now let us see whether there exist scalars λ and non-zero vectors **X** which satisfy (1).

If **I** denotes the unit matrix of order n, then the equation (1) may be written as

$$\mathbf{AX} = \lambda \mathbf{IX}$$

or $$(\mathbf{A} - \lambda \mathbf{I})\, \mathbf{X} = \mathbf{O}. \qquad \ldots(2)$$

The matrix equation (2) represents the following system of n homogeneous equations in n unknowns :

$$\left.\begin{array}{l} (a_{11} - \lambda) x_1 + a_{12} x_2 + \ldots + a_{1n} x_n = 0 \\ a_{21} x_1 + (a_{22} - \lambda)\, x_2 + \ldots + a_{2n} x_n = 0 \\ \ldots \ \ldots \ \ldots \ \ldots \ \ldots \ \ldots \ \ldots \ \ldots \\ a_{n1} x_1 + a_{n2} x_2 + \ldots + (a_{nn} - \lambda) x_n = 0 \end{array}\right\} \qquad \ldots(3)$$

The coefficient matrix of the equations (3) is $\mathbf{A} - \lambda \mathbf{I}$. The necessary and sufficient condition for equations (3) to possess a non-zero solution ($\mathbf{X} \neq \mathbf{O}$) is that the coefficient matrix $\mathbf{A} - \lambda \mathbf{I}$ should be of rank less than the number of unknowns n. But this will be so if and only if the matrix $\mathbf{A} - \lambda \mathbf{I}$ is singular *i.e.*, if and only if $|\mathbf{A} - \lambda \mathbf{I}| = 0$. Thus the scalars λ for which

$$|\mathbf{A} - \lambda \mathbf{I}| = 0$$

are of special importance.

Definitions: **(Lucknow 2009; Gorakhpur 11)**

Let $\mathbf{A} = [a_{ij}]_{n \times n}$ be any n-rowed square matrix and λ an indeterminate. The matrix $\mathbf{A} - \lambda \mathbf{I}$ is called the characteristic matrix of $\mathbf{A}$ where $\mathbf{I}$ is the unit matrix of order n.

Also the determinant

$$|\mathbf{A} - \lambda \mathbf{I}| = \begin{vmatrix} a_{11} - \lambda & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} - \lambda & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} - \lambda \end{vmatrix},$$

which is an ordinary polynomial in λ of degree n, is called the **characteristic polynomial of A**. The equation $|\mathbf{A} - \lambda \mathbf{I}| = 0$ is called the **characteristic equation** of **A** and the roots of this equation are called the **characteristic roots** or **characteristic values** or **eigenvalues** or **latent roots** or **proper values** of the matrix **A**. The set of the eigenvalues of **A** is called the **spectrum** of A.

If λ is a characteristic root of the matrix **A**, then

$$|\mathbf{A} - \lambda \mathbf{I}| = 0$$

and the matrix $\mathbf{A} - \lambda \mathbf{I}$ is singular. Therefore there exists a non-zero vector **X** such that

$$(\mathbf{A} - \lambda \mathbf{I})\, \mathbf{X} = \mathbf{O}$$

or $$\mathbf{AX} = \lambda \mathbf{X}.$$

Characteristic vectors. Definition: *If λ is a characteristic root of an $n \times n$ matrix* **A**, *then a non-zero vector* **X** *such that* $\mathbf{AX} = \lambda \mathbf{X}$ *is called a characteristic vector or eigenvector of* **A** *corresponding to the characteristic root* λ. **(Lucknow 2009)**

3 Certain Relations between Characteristic Roots and Characteristic Vectors

Theorem 1: *λ is a characteristic root of a matrix* **A** *if and only if there exists a non-zero vector* **X** *such that* $\mathbf{AX} = \lambda \mathbf{X}$.

Proof: Suppose λ is a characteristic root of the matrix **A**. Then $|\mathbf{A} - \lambda \mathbf{I}| = 0$ and the matrix $\mathbf{A} - \lambda \mathbf{I}$ is singular. Therefore, the matrix equation $(\mathbf{A} - \lambda \mathbf{I})\, \mathbf{X} = \mathbf{O}$ possesses a non-zero solution *i.e.*, there exists a non-zero vector **X** such that $(\mathbf{A} - \lambda \mathbf{I})\mathbf{X} = \mathbf{O}$ or $\mathbf{AX} = \lambda \mathbf{X}$.

Conversely suppose there exists a non-zero vector **X** such that $\mathbf{AX} = \lambda \mathbf{X}$ *i.e.*, $(\mathbf{A} - \lambda \mathbf{I})\, \mathbf{X} = \mathbf{O}$. Since the matrix equation $(\mathbf{A} - \lambda \mathbf{I})\, \mathbf{X} = \mathbf{O}$ possesses a non-zero solution, therefore the coefficient matrix $\mathbf{A} - \lambda \mathbf{I}$ must be singular *i.e.*, $|\mathbf{A} - \lambda \mathbf{I}| = 0$. Hence λ is a characteristic root of the matrix **A.**

Theorem 2: *If* **X** *is a characteristic vector of a matrix* **A** *corresponding to the characteristic value λ, then k***X** *is also a characteristic vector of* **A** *corresponding to the same characteristic value λ. Here k is any non-zero scalar.*

Proof: Suppose **X** is a characteristic vector of **A** corresponding to the characteristic value λ. Then $\mathbf{X} \neq \mathbf{O}$ and $\mathbf{AX} = \lambda \mathbf{X}$.

If k is any non-zero scalar, then $k\,\mathbf{x} \neq \mathbf{O}$. Also

$$\mathbf{A}\,(k\mathbf{X}) = k(\mathbf{AX}) = k(\lambda \mathbf{X}) = \lambda(k\mathbf{X}).$$

Now $k\mathbf{X}$ is a non-zero vector such that $\mathbf{A}(k\mathbf{X}) = \lambda(k\mathbf{X})$. Hence $k\mathbf{X}$ is a characteristic vector of **A** corresponding to the characteristic value λ. Thus corresponding to a characteristic value λ, there corresponds more than one characteristic vectors.

Theorem 3: *If* **X** *is a characteristic vector of a matrix* **A**, *then* **X** *cannot correspond to more than one characteristic values of* **A**.

Proof: Let **X** be a characteristic vector of a matrix **A** corresponding to two characteristic values λ_1 and λ_2. Then $\mathbf{AX} = \lambda_1 \mathbf{X}$ and $\mathbf{AX} = \lambda_2 \mathbf{X}$. Therefore

$$\lambda_1 \mathbf{X} = \lambda_2 \mathbf{X}$$

$$\Rightarrow \quad (\lambda_1 - \lambda_2)\, \mathbf{X} = \mathbf{O}$$

$$\Rightarrow \quad \lambda_1 - \lambda_2 = 0 \qquad [\because \ \mathbf{X} \neq \mathbf{O}]$$

$$\Rightarrow \quad \lambda_1 = \lambda_2.$$

Hence the result.

4 Nature of the Characteristic Roots of Special Types of Matrices

Theorem 1: *The characteristic roots of a Hermitian matrix are real.*

(Kumaun 2010, 11, 13; Meerut 11; Lucknow 05, 11)

Proof: Suppose **A** is a Hermitian matrix, λ a characteristic root of **A** and **X** a corresponding eigenvector. Then

$$\mathbf{AX} = \lambda \mathbf{X}. \quad ...(1)$$

Premultiplying both sides of (1) by X^θ, we get

$$\mathbf{X}^\theta \mathbf{AX} = \lambda \mathbf{X}^\theta \mathbf{X}. \quad ...(2)$$

Taking conjugate transpose of both sides of (2), we get

$$(\mathbf{X}^\theta \mathbf{AX})^\theta = (\lambda \mathbf{X}^\theta \mathbf{X})^\theta$$

or $$\mathbf{X}^\theta \mathbf{A}^\theta (\mathbf{X}^\theta)^\theta = \bar{\lambda} \mathbf{X}^\theta (\mathbf{X}^\theta)^\theta$$

or $$\mathbf{X}^\theta \mathbf{AX} = \bar{\lambda} \mathbf{X}^\theta \mathbf{X} \quad ...(3)$$

[$\because (\mathbf{X}^\theta)^\theta = \mathbf{X}$ and $\mathbf{A}^\theta = \mathbf{A}$, **A** being Hermitian]

From (2) and (3), we have

$$\lambda \mathbf{X}^\theta \mathbf{X} = \bar{\lambda} \mathbf{X}^\theta \mathbf{X}$$

or $$(\lambda - \bar{\lambda}) \mathbf{X}^\theta \mathbf{X} = \mathbf{O}.$$

But **X** is not a zero vector, therefore $\mathbf{X}^\theta \mathbf{X} \neq \mathbf{O}$.

Hence $\lambda - \bar{\lambda} = 0$, so that $\lambda = \bar{\lambda}$ and consequently λ is real.

Corollary 1: *The characteristic roots of a real symmetric matrix are all real.*

If the elements of a Hermitian matrix **A** are all real, then **A** is a real symmetric matrix. Thus a real symmetric matrix is Hermitian and therefore the result follows.

Corollary 2: *The characteristic roots of a skew-Hermitian matrix are either pure imaginary or zero.* **(Lucknow 2010)**

Suppose **A** is a skew-Hermitian matrix. Then $i\mathbf{A}$ is Hermitian. Let λ be a characteristic root of **A**. Then

$$\mathbf{AX} = \lambda \mathbf{X}$$

or $$(i\mathbf{A})\mathbf{X} = (i\lambda)\mathbf{X}.$$

From this it follows that $i\lambda$ is a characteristic root of $i\mathbf{A}$ which is Hermitian. Hence $i\lambda$ is real. Therefore either λ must be zero or pure imaginary.

Corollary 3: *The characteristic roots of a real skew-symmetric matrix are either pure imaginary or zero, for every such matrix is skew-Hermitian.*

Theorem 2: *The characteristic roots of a unitary matrix are of unit modulus.*

Proof: Suppose **A** is a unitary matrix. Then $\mathbf{A}^\theta \mathbf{A} = \mathbf{I}$.

Let λ be a characteristic root of **A** and **X** a corresponding eigenvector. Then

$$\mathbf{AX} = \lambda \mathbf{X}. \quad ...(1)$$

Taking conjugate transpose of both sides of (1), we get

$$(\mathbf{AX})^\theta = (\lambda \mathbf{X})^\theta$$

or $$\mathbf{X}^\theta \mathbf{A}^\theta = \bar{\lambda} \mathbf{X}^\theta. \quad ...(2)$$

From (1) and (2), we have

$$(X^\theta \mathbf{A}^\theta)(\mathbf{AX}) = \bar{\lambda}\lambda \mathbf{X}^\theta \mathbf{X}$$

or $\mathbf{X}^\theta (\mathbf{A}^\theta \mathbf{A})\, \mathbf{X} = \bar{\lambda}\lambda \mathbf{X}^\theta \mathbf{X}$ $[\because \mathbf{A}^\theta \mathbf{A} = \mathbf{I}]$

or $\mathbf{X}^\theta \mathbf{IX} = \bar{\lambda}\lambda \mathbf{X}^\theta \mathbf{X}$

or $\mathbf{X}^\theta \mathbf{X} = \bar{\lambda}\lambda \mathbf{X}^\theta \mathbf{X}$

or $\mathbf{X}^\theta \mathbf{X}(\lambda\bar{\lambda} - 1) = \mathbf{O}.$ ...(3)

Since $\mathbf{X}^\theta \mathbf{X} \neq \mathbf{O}$, therefore, (3) gives

$$\lambda\bar{\lambda} - 1 = 0$$

or $\lambda\, \bar{\lambda} = 1$

or $|\lambda|^2 = 1.$

Corollary: *The characteristic roots of an orthogonal matrix are of unit modulus.*

We know that if the elements of a unitary matrix **A** are all real, then **A** is said to be an orthogonal matrix. Hence the result follows.

5 The Process of Finding the Eigenvalues and Eigenvectors of a Matrix

Let $\mathbf{A} = [a_{ij}]_{n \times n}$ be a square matrix of order n. First we should write the characteristic equation of the matrix **A** *i.e.*, the equation $|\mathbf{A} - \lambda \mathbf{I}| = 0$. This equation will be of degree n in λ. So it will have n roots. These n roots will give us the eigenvalues of the matrix **A**. If λ_1 is an eigenvalue of **A**, then the corresponding eigenvectors of **A** will be given by the non-zero vectors

$$\mathbf{X} = [x_1, x_2, \ldots, x_n]'$$

satisfying the equation $\mathbf{AX} = \lambda_1 \mathbf{X}$

or $(\mathbf{A} - \lambda_1 \mathbf{I})\, \mathbf{X} = \mathbf{O}.$

Illustrative Examples

Example 1: *Determine the characteristic roots of the matrix*

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & -1 \\ 2 & -1 & 0 \end{bmatrix}.$$

(Meerut 2010)

Solution: The characteristic matrix of **A**

$$= \mathbf{A} - \lambda \mathbf{I}$$

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

$$= \begin{bmatrix} 0-\lambda & 1 & 2 \\ 1 & 0-\lambda & -1 \\ 2 & -1 & 0-\lambda \end{bmatrix}.$$

It should be noted that in order to obtain the characteristic matrix of a matrix **A**, we should simply subtract λ from each of its principal diagonal elements.

The characteristic polynomial of **A**

$$= |\mathbf{A} - \lambda \mathbf{I}|$$

$$= \begin{vmatrix} -\lambda & 1 & 2 \\ 1 & -\lambda & -1 \\ 2 & -1 & -\lambda \end{vmatrix}$$

$$= -\lambda(\lambda^2 - 1) - 1(-\lambda + 2) + 2(-1 + 2\lambda)$$

$$= -\lambda^3 + \lambda + \lambda - 2 - 2 + 4\lambda$$

$$= -\lambda^3 + 6\lambda - 4.$$

$\therefore$ The characteristic equation of **A** is

$$|\mathbf{A} - \lambda \mathbf{I}| = 0$$

i.e., $\lambda^3 - 6\lambda + 4 = 0$

i.e., $(\lambda - 2)(\lambda^2 + 2\lambda - 2) = 0.$

The roots of this equation are $\lambda = 2, -1 \pm \sqrt{3}$.

Hence the characteristic roots of the matrix **A** are $2, -1 \pm \sqrt{3}$.

Example 2: *Determine the eigenvalues of the matrix* $\mathbf{A} = \begin{bmatrix} a & h & g \\ 0 & b & 0 \\ 0 & c & c \end{bmatrix}$.

(Kumaun 2007; Kanpur 09; Meerut 10B)

Solution: Here $|\mathbf{A} - \lambda \mathbf{I}| = \begin{vmatrix} a-\lambda & h & g \\ 0 & b-\lambda & 0 \\ 0 & c & c-\lambda \end{vmatrix}$

$$= (a - \lambda)(b - \lambda)(c - \lambda).$$

The characteristic equation of **A** is

$$|\mathbf{A} - \lambda \mathbf{I}| = 0$$

i.e., $(a - \lambda)(b - \lambda)(c - \lambda) = 0.$

The roots of this equation are $\lambda = a, b, c$. Hence the eigenvalues of **A** are a, b, c.

Example 3: *Determine the eigenvalues and eigenvectors of the matrix*

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

(Bundelkhand 2006)

Solution: The characteristic equation of A is

$$|\mathbf{A} - \lambda \mathbf{I}| = 0$$

i.e., $\begin{vmatrix} 5-\lambda & 4 \\ 1 & 2-\lambda \end{vmatrix} = 0$

i.e., $(5-\lambda)(2-\lambda) - 4 = 0$

i.e., $\lambda^2 - 7\lambda + 6 = 0.$

The roots of this equation are $\lambda_1 = 6, \lambda_2 = 1$. Therefore the eigenvalues of **A** are 6, 1.

The eigenvectors $\mathbf{X} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ of **A** corresponding to the eigenvalue 6 are given by the non-zero solutions of the equation

$$(\mathbf{A} - 6\mathbf{I})\,\mathbf{X} = \mathbf{O} \quad \text{or} \quad \begin{bmatrix} 5-6 & 4 \\ 1 & 2-6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

or $\begin{bmatrix} -1 & 4 \\ 1 & -4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

or $\begin{bmatrix} -1 & 4 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, applying $R_2 \to R_2 + R_1$.

The coefficient matrix of these equations is of rank 1. Therefore these equations have 2 – 1 *i.e.,* 1 linearly independent solution. These equations reduce to the single equation $-x_1 + 4x_2 = 0$. Obviously $x_1 = 4, x_2 = 1$ is a solution of this equation. Therefore $\mathbf{X}_1 = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$ is an eigenvector of **A** corresponding to the eigenvalue 6. The set of all eigenvectors of **A** corresponding to the eigenvalue 6 is given by $c_1\mathbf{X}_1$ where c_1 is any non-zero scalar.

The eigenvectors **X** of **A** corresponding to the eigenvalue 1 are given by the non-zero solutions of the equation

$$(\mathbf{A} - 1\mathbf{I})\,\mathbf{X} = \mathbf{O}$$

or $\begin{bmatrix} 4 & 4 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

or $4x_1 + 4x_2 = 0, \quad x_1 + x_2 = 0.$

From these $x_1 = -x_2$. Let us take $x_1 = 1, x_2 = -1$. Then $\mathbf{X}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ is an eigenvector of **A** corresponding to the eigenvalue 1. Every non-zero multiple of the vector $\mathbf{X}_2$ is an eigenvector of **A** corresponding to the eigenvalue 1.

Example 4: *Determine the characteristic roots and the corresponding characteristic vectors of the matrix*

$$\mathbf{A} = \begin{bmatrix} 8 & -6 & 2 \\ -6 & 7 & -4 \\ 2 & -4 & 3 \end{bmatrix}.$$

(Purvanchal 2010; Bundelkhand 08; Rohilkhand 05; Agra 07; Kanpur 09; Avadh 05)

Solution: The characteristic equation of the matrix **A** is

$$|\mathbf{A} - \lambda\mathbf{I}| = 0$$

i.e.,
$$\begin{vmatrix} 8-\lambda & -6 & 2 \\ -6 & 7-\lambda & -4 \\ 2 & -4 & 3-\lambda \end{vmatrix} = 0$$

or $$(8-\lambda)\{(7-\lambda)(3-\lambda)-16\} + 6\{-6(3-\lambda)+8\} + 2\{24-2(7-\lambda)\} = 0$$

or $$\lambda^3 - 18\lambda^2 + 45\lambda = 0$$

or $$\lambda(\lambda-3)(\lambda-15) = 0.$$

Hence the characteristic roots of **A** are 0, 3, 15.

The eigenvectors $\mathbf{X} = [x_1, x_2, x_3]'$ of **A** corresponding to the eigenvalue 0 are given by the non-zero solutions of the equation

$$(\mathbf{A} - 0\mathbf{I})\mathbf{X} = \mathbf{O}$$

or $$\begin{bmatrix} 8 & -6 & 2 \\ -6 & 7 & -4 \\ 2 & -4 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

or $$\begin{bmatrix} 2 & -4 & 3 \\ -6 & 7 & -4 \\ 8 & -6 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \text{ by } R_1 \longleftrightarrow R_3$$

.or $$\begin{bmatrix} 2 & -4 & 3 \\ 0 & -5 & 5 \\ 0 & 10 & -10 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \text{ by } R_2 \to R_2 + 3R_1, R_3 \to R_3 - 4R_1$$

or $$\begin{bmatrix} 2 & -4 & 3 \\ 0 & -5 & 5 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \text{ by } R_3 \to R_3 + 2R_2.$$

The coefficient matrix of these equations is of rank 2. Therefore these equations have $3 - 2 = 1$ linearly independent solution. Thus there is only one linearly independent eigenvector corresponding to the eigenvalue 0. These equations can be written as

$$2x_1 - 4x_2 + 3x_3 = 0, -5x_2 + 5x_3 = 0.$$

From the last equation, we get $x_2 = x_3$. Let us take $x_2 = 1, x_3 = 1$. Then the first equation gives $x_1 = \frac{1}{2}$. Therefore $\mathbf{X}_1 = \begin{bmatrix} \frac{1}{2} & 1 & 1 \end{bmatrix}'$ is an eigenvector of **A** corresponding to the eigenvalue 0. If c_1 is any non-zero scalar, then $c_1\mathbf{X}_1$ is also an eigenvector of **A** corresponding to the eigenvalue 0.

The eigenvectors of **A** corresponding to the eigenvalue 3 are given by the non-zero solutions of the equation

$$(\mathbf{A} - 3\mathbf{I})\mathbf{X} = \mathbf{O}$$

or $$\begin{bmatrix} 5 & -6 & 2 \\ -6 & 4 & -4 \\ 2 & -4 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

or $$\begin{bmatrix} -1 & -2 & -2 \\ -6 & 4 & -4 \\ 2 & -4 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \text{ by } R_1 \to R_1 + R_2$$

or $$\begin{bmatrix} -1 & -2 & -2 \\ 0 & 16 & 8 \\ 0 & -8 & -4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \text{ by } R_2 \to R_2 - 6R_1,\ R_3 \to R_3 + 2R_1$$

or $$\begin{bmatrix} -1 & -2 & -2 \\ 0 & 16 & 8 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \text{ by } R_3 \to R_3 + \frac{1}{2} R_2.$$

The coefficient matrix of these equations is of rank 2. Therefore these equations have $3 - 2 = 1$ linearly independent solution. These equations can be written as

$$-x_1 - 2x_2 - 2x_3 = 0, 16x_2 + 8x_3 = 0.$$

From the second equation we get $x_2 = -\frac{1}{2} x_3$. Let us take $x_3 = 4, x_2 = -2$. Then the first equation gives $x_1 = -4$. Therefore $\mathbf{X}_2 = [-4 \ \ -2 \ \ 4]'$ is an eigenvector of **A** corresponding to the eigenvalue 3. Every non-zero multiple of the $\mathbf{X}_2$ is an eigenvector of **A** corresponding to the eigenvalue 3.

The eigenvectors of **A** corresponding to the eigenvalue 15 are given by the non-zero solutions of the equation

$$(\mathbf{A} - 15\mathbf{I})\,\mathbf{X} = \mathbf{O}$$

or $$\begin{bmatrix} 8-15 & -6 & 2 \\ -6 & 7-15 & -4 \\ 2 & -4 & 3-15 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

or $$\begin{bmatrix} -7 & -6 & 2 \\ -6 & -8 & -4 \\ 2 & -4 & -12 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

or $$\begin{bmatrix} -1 & 2 & 6 \\ -6 & -8 & -4 \\ 2 & -4 & -12 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \text{ by } R_1 \to R_1 - R_2$$

or $$\begin{bmatrix} -1 & 2 & 6 \\ 0 & -20 & -40 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \text{ by } R_2 \to R_2 - 6R_1, \ R_3 \to R_3 + 2R_1.$$

The coefficient matrix of these equations is of rank 2. Therefore these equations have $3 - 2 = 1$ linearly independent solution. These equations can be written as

$$-x_1 + 2x_2 + 6x_3 = 0, \ -20x_2 - 40x_3 = 0.$$

The last equation gives $x_2 = -2x_3$. Let us take $x_3 = 1, x_2 = -2$. Then the first equation gives $x_1 = 2$. Therefore

$$\mathbf{X}_3 = [2 \quad -2 \quad 1]'$$

is an eigenvector of **A** corresponding to the eigenvalue 15. If k is any non-zero scalar, then $k\mathbf{X}_3$ is also an eigenvector of **A** corresponding to the eigenvalue 15.

Example 5: *Determine the characteristic roots and the corresponding characteristic vectors of the matrix*

$$\mathbf{A} = \begin{bmatrix} 6 & -2 & 2 \\ -2 & 3 & -1 \\ 2 & -1 & 3 \end{bmatrix}.$$

(Meerut 2006B, 09; Purvanchal 07)

Solution: The characteristic equation of **A** is

$$|\mathbf{A} - \lambda \mathbf{I}| = 0$$

or $$\begin{bmatrix} 6-\lambda & -2 & 2 \\ -2 & 3-\lambda & -1 \\ 2 & -1 & 3-\lambda \end{bmatrix} = 0$$

or $$\begin{bmatrix} 6-\lambda & -2 & 0 \\ -2 & 3-\lambda & 2-\lambda \\ 2 & -1 & 2-\lambda \end{bmatrix} = 0, \text{ by } C_3 \to C_3 + C_2$$

or $$(2-\lambda)\begin{bmatrix} 6-\lambda & -2 & 0 \\ -2 & 3-\lambda & 1 \\ 2 & -1 & 1 \end{bmatrix} = 0$$

or $$(2-\lambda)\begin{bmatrix} 6-\lambda & -2 & 0 \\ -4 & 4-\lambda & 0 \\ 2 & -1 & 1 \end{bmatrix} = 0, \text{ by } R_2 \to R_2 - R_3$$

or $$(2-\lambda)\,[(6-\lambda)(4-\lambda) - 8] = 0$$

or $$(2-\lambda)(\lambda^2 - 10\lambda + 16) = 0$$

or $$(2-\lambda)(\lambda-2)(\lambda-8) = 0.$$

Therefore the characteristic roots of **A** are given by $\lambda = 2, 2, 8$.

The characteristic vectors of **A** corresponding to the characteristic root 8 are given by the non-zero solutions of the equation

$$(\mathbf{A} - 8\mathbf{I})\,\mathbf{X} = \mathbf{O}$$

or $$\begin{bmatrix} 6-8 & -2 & 2 \\ -2 & 3-8 & -1 \\ 2 & -1 & 3-8 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

or $$\begin{bmatrix} -2 & -2 & 2 \\ -2 & -5 & -1 \\ 2 & -1 & -5 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

or $$\begin{bmatrix} -2 & -2 & 2 \\ 0 & -3 & -3 \\ 0 & -3 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \text{ by } R_2 \to R_2 - R_1, R_3 \to R_3 + R_1$$

or $$\begin{bmatrix} -2 & -2 & 2 \\ 0 & -3 & -3 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \text{ by } R_3 \to R_3 - R_2.$$

The coefficient matrix of these equations is of rank 2. Therefore these equations possess $3 - 2 = 1$ linearly independent solution. These equations can be written as

$$-2x_1 - 2x_2 + 2x_3 = 0,$$
$$-3x_2 - 3x_3 = 0.$$

The last equation gives $x_2 = -x_3$. Let us take $x_3 = 1, x_2 = -1$. Then the first equation gives $x_1 = 2$. Therefore $\mathbf{X}_1 = \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}$ is an eigenvector of $\mathbf{A}$ corresponding to the eigenvalue 8. Every non-zero multiple of $\mathbf{X}_1$ is an eigenvector of $\mathbf{A}$ corresponding to the eigenvalue 8.

The eigenvectors of $\mathbf{A}$ corresponding to the eigenvalue 2 are given by the non-zero solutions of the equation

$$(\mathbf{A} - 2\mathbf{I})\,\mathbf{X} = \mathbf{O}$$

or $$\begin{bmatrix} 4 & -2 & 2 \\ -2 & 1 & -1 \\ 2 & -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

or $$\begin{bmatrix} -2 & 1 & -1 \\ 4 & -2 & 2 \\ 2 & -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \text{ by } R_1 \longleftrightarrow R_2$$

or $$\begin{bmatrix} -2 & 1 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \text{ by } R_2 \to R_2 + 2R_1, R_3 \to R_3 + R_1.$$

The coefficient matrix of these equations is of rank 1. Therefore these equations possess $3 - 1 = 2$ linearly independent solutions. We see that these equations reduce to the single equation

$$-2x_1 + x_2 - x_3 = 0.$$

Obviously $\mathbf{X}_2 = \begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix}, \ \mathbf{X}_3 = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$

are two linearly independent solutions of this equation. Therefore $\mathbf{X}_2$ and $\mathbf{X}_3$ are two linearly independent eigenvectors of $\mathbf{A}$ corresponding to the eigenvalue 2. If c_1, c_2 are scalars not both equal to zero, then $c_1\mathbf{X}_2 + c_2\mathbf{X}_3$ gives all the eigenvectors of $\mathbf{A}$ corresponding to the eigenvalue 2.

Example 6: *Show that 0 is a characteristic root of a matrix if and only if the matrix is singular.*

Solution: We have 0 is an eigenvalue of $\mathbf{A} \Rightarrow \lambda = 0$ satisfies the equation

$$|\mathbf{A} - \lambda\mathbf{I}| = 0 \Rightarrow |\mathbf{A}| = 0 \Rightarrow \mathbf{A} \text{ is singular.}$$

Conversely, $\mathbf{A}$ is singular $\Rightarrow |\mathbf{A}| = 0$

$\Rightarrow$ $\lambda = 0$ satisfies the equation $|\mathbf{A} - \lambda\mathbf{I}| = 0$

$\Rightarrow$ 0 is an eigenvalue of $\mathbf{A}$.

Example 7: *If $\lambda_1, \ldots\ldots\ldots, \lambda_n$ are the eigenvalues of* $\mathbf{A}$, *then show that $k\lambda_1, \ldots, k\lambda_n$ are the eigenvalues of $k\mathbf{A}$.*

Solution: If $k = 0$, then $k\mathbf{A} = \mathbf{O}$ and each eigenvalue of $\mathbf{O}$ is 0. Thus $0\lambda_1, \ldots., 0\lambda_n$ are the eigenvalues of $k\mathbf{A}$ if $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of $\mathbf{A}$.

So let us suppose that $k \neq 0$.

We have
$$|k\mathbf{A} - \lambda k\mathbf{I}| = |k(\mathbf{A} - \lambda\mathbf{I})|$$
$$= k^n |\mathbf{A} - \lambda\mathbf{I}|. \qquad [\because \ |k\mathbf{B}| = k^n |\mathbf{B}|]$$

$\therefore$ if $k \neq 0$, then $|k\mathbf{A} - \lambda k\mathbf{I}| = 0$ if and only if $|\mathbf{A} - \lambda\mathbf{I}| = 0$

i.e., $k\lambda$ is an eigenvalue of $k\mathbf{A}$ if and only if λ is an eigenvalue of $\mathbf{A}$.

Thus $k\lambda_1, \ldots., k\lambda_n$ are the eigenvalues of $k\mathbf{A}$ if $\lambda_1, \ldots, \lambda_n$ are eigenvalues of $\mathbf{A}$.

Example 8: *If* $\mathbf{A}$ *is non-singular, prove that the eigenvalues of* $\mathbf{A}^{-1}$ *are the reciprocals of the eigenvalues of* $\mathbf{A}$.

Solution: Let λ be an eigenvalue of $\mathbf{A}$ and $\mathbf{X}$ be a corresponding eigenvector. Then

$$\mathbf{AX} = \lambda\mathbf{X}$$

$\Rightarrow$ $$\mathbf{X} = \mathbf{A}^{-1}(\lambda\mathbf{X}) = \lambda(\mathbf{A}^{-1}\mathbf{X})$$

$\Rightarrow$ $$\frac{1}{\lambda}\mathbf{X} = \mathbf{A}^{-1}\mathbf{X} \qquad [\because \ \mathbf{A} \text{ is non-singular} \Rightarrow \lambda \neq 0]$$

$\Rightarrow$ $$\mathbf{A}^{-1}\mathbf{X} = \frac{1}{\lambda}\mathbf{X}$$

$\Rightarrow$ $\frac{1}{\lambda}$ is an eigenvalue of $\mathbf{A}^{-1}$ and $\mathbf{X}$ is a corresponding eigenvector.

Conversely suppose that k is an eigenvalue of $\mathbf{A}^{-1}$. Since $\mathbf{A}$ is non-singular $\Rightarrow \mathbf{A}^{-1}$ is non-singular and $(\mathbf{A}^{-1})^{-1} = \mathbf{A}$, therefore, it follows from the first part of this question that $\frac{1}{k}$ is an eigenvalue of $\mathbf{A}$. Thus each eigenvalue of $\mathbf{A}^{-1}$ is equal to the reciprocal of some eigenvalue of $\mathbf{A}$.

Hence the eigenvalues of $\mathbf{A}^{-1}$ are nothing but the reciprocals of the eigenvalues of $\mathbf{A}$.

Example 9: *If the characteristic roots of* $\mathbf{A}$ *are $\lambda_1, \lambda_2, \ldots \lambda_n$, then the characteristic roots of $\mathbf{A}^2$ are $\lambda_1^2, \lambda_2^2, \ldots, \lambda_n^2$.* **(Kumaun 2008)**

Solution: Let λ be a characteristic root of the matrix $\mathbf{A}$. Then there exists a non-zero vector $\mathbf{X}$ such that

$$\mathbf{AX} = \lambda \mathbf{X}$$

$\Rightarrow \quad \mathbf{A}(\mathbf{AX}) = \mathbf{A}(\lambda \mathbf{X})$

$\Rightarrow \quad \mathbf{A}^2\mathbf{X} = \lambda(\mathbf{AX})$

$\Rightarrow \quad \mathbf{A}^2\mathbf{X} = \lambda(\lambda \mathbf{X}) \qquad [\because \mathbf{AX} = \lambda \mathbf{X}]$

$\Rightarrow \quad \mathbf{A}^2\mathbf{X} = \lambda^2 \mathbf{X}. \qquad \ldots(1)$

Since $\mathbf{X}$ is a non-zero vector, therefore from the relation (1) it is obvious that λ^2 is a characteristic root of the matrix $\mathbf{A}^2$. Therefore if $\lambda_1, \ldots, \lambda_n$ are the characteristic roots of $\mathbf{A}$, then $\lambda_1^2, \ldots, \lambda_n^2$ are the characteristic roots of $\mathbf{A}^2$.

Example 10: *The characteristic roots of an idempotent matrix are either zero or unity.*

Solution: Let $\mathbf{A}$ be an idempotent matrix so that $\mathbf{A}^2 = \mathbf{A}$. Let λ be a characteristic root of the matrix $\mathbf{A}$. Then there exists a non-zero vector $\mathbf{X}$ such that

$\mathbf{AX} = \lambda \mathbf{X} \qquad \ldots(1)$

$\Rightarrow \quad \mathbf{A}(\mathbf{AX}) = \mathbf{A}(\lambda \mathbf{X})$

$\Rightarrow \quad \mathbf{A}^2\mathbf{X} = \lambda(\mathbf{AX})$

$\Rightarrow \quad \mathbf{AX} = \lambda(\lambda \mathbf{X}) \qquad [\because \mathbf{A}^2 = \mathbf{A} \text{ and } \mathbf{AX} = \lambda \mathbf{X}]$

$\Rightarrow \quad \mathbf{AX} = \lambda^2 \mathbf{X}. \qquad \ldots(2)$

From (1) and (2), we get

$$\lambda^2 \mathbf{X} = \lambda \mathbf{X}$$

or $\quad (\lambda^2 - \lambda)\mathbf{X} = \mathbf{O}$

or $\quad \lambda^2 - \lambda = 0 \qquad [\because \mathbf{X} \neq \mathbf{O}]$

or $\quad \lambda(\lambda - 1) = 0.$

$\therefore \quad \lambda = 0 \quad \text{or} \quad \lambda = 1.$

Hence the characteristic roots of an idempotent matrix are either zero or unity.

Example 11: *The product of the characteristic roots of a square matrix of order n is equal to the determinant of the matrix.*

Solution: Let $\mathbf{A} = [a_{ij}]_{n \times n}$ be a square matrix of order n. Then the characteristic polynomial $f(\lambda)$ of $\mathbf{A}$ is given by

$$f(\lambda) = |\mathbf{A} - \lambda \mathbf{I}|$$

$$= \begin{vmatrix} a_{11} - \lambda & a_{12} & \ldots & a_{1n} \\ a_{22} & a_{22} - \lambda & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} - \lambda \end{vmatrix}$$

$$= (-1)^n [\lambda^n + p_1 \lambda^{n-1} + p_2 \lambda^{n-2} + \ldots + p_n], \text{ say.}$$

If $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the characteristic roots of $\mathbf{A}$, then $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the roots of the equation $|\mathbf{A} - \lambda \mathbf{I}| = 0$.

$\therefore \qquad |\mathbf{A} - \lambda\mathbf{I}| = (-1)^n [\lambda^n + p_1\lambda^{n-1} + p_2 \lambda^{n-2} + \ldots + p_n]$

$$= (-1)^n (\lambda - \lambda_1)(\lambda - \lambda_2)\ldots(\lambda - \lambda_n). \qquad \ldots(1)$$

Putting $\lambda = 0$ on both sides of (1), we get

$$|\mathbf{A}| = (-1)^n (-\lambda_1)(-\lambda_2)\ldots(-\lambda_n)$$

$$= (-1)^n (-1)^n \lambda_1 \lambda_2 \ldots \lambda_n$$

$$= (-1)^{2n} \lambda_1 \lambda_2 \ldots \lambda_n$$

$$= \lambda_1 \lambda_2 \ldots \lambda_n$$

Hence $\qquad \lambda_1 \lambda_2 \ldots \lambda_n = |\mathbf{A}|.$

Example 12: *Any two characteristic vectors corresponding to two distinct characteristic roots of a:*

(i) Hermitian, (ii) Real symmetric, (iii) Unitary matrix are orthogonal.

Solution : **(i) A is Hermitian:**

We have $\qquad \mathbf{A}^\theta = \mathbf{A}.$

Also, we have

$$\mathbf{AX_1} = \lambda_1 \mathbf{X_1} \text{ and } \mathbf{AX}_2 = \lambda_2 \mathbf{X}_2$$

or $\qquad \lambda_1 \mathbf{X}_1^\theta (\mathbf{X}_2^\theta)^\theta = \lambda_2 \mathbf{X}_1^\theta \mathbf{X}_2; \qquad [\lambda_1 \text{ is real} \Rightarrow \lambda_1^\theta = \lambda_1]$

or $\qquad \lambda_1\mathbf{X}_1^\theta \mathbf{X}_2 = \lambda_2 \mathbf{X}_1^\theta \mathbf{X}_2$

or $\qquad (\lambda_1 - \lambda_2) \mathbf{X}_1^\theta \mathbf{X}_2 = \mathbf{O}.$

Since $\lambda_1 - \lambda_2 \neq 0,$ we have

$$\mathbf{X}_1^\theta \mathbf{X}_2 = \mathbf{O}.$$

$\Rightarrow \quad \mathbf{X}_1$ and $\mathbf{X}_2$ are orthogonal with respect to each other.

(ii) A is real symmetric: The real symmetric matrix is always Hermitian, so the result follows at once from (i).

(iii) A is unitary: We have $\mathbf{A}^\theta\mathbf{A} = \mathbf{I}.$

Now, $\qquad \mathbf{AX}_1 = \lambda_1\mathbf{X}_1$ and $\mathbf{AX}_2 = \lambda_2\mathbf{X}_2,$

where λ_1, λ_2 are characteristic roots of a unitary matrix which must be uni-modular

i.e., $\qquad \lambda_1\bar{\lambda}_1 = 1, \quad \lambda_2\bar{\lambda}_2 = 1.$

Now, $\qquad \mathbf{AX}_2 = \lambda_2\mathbf{X}_2$

$\Rightarrow \qquad (\mathbf{AX}_2)^\theta = (\lambda_2\mathbf{X}_2)^\theta$

$\Rightarrow \qquad \mathbf{X}_2^\theta \mathbf{A}^\theta = \bar{\lambda}_2 \mathbf{A}_2^\theta$

$\Rightarrow \qquad \mathbf{X}_2^\theta \mathbf{A}^\theta \mathbf{AX}_1 = \bar{\lambda}_2 \mathbf{X}_2^\theta \lambda_1 \mathbf{X}_1 = \bar{\lambda}_2 \lambda_1 \mathbf{X}_2^\theta \mathbf{X}_1$

Again $\mathbf{A}^\theta\mathbf{A} = \mathbf{I},$ hence we get

$$\mathbf{X}_2^\theta \mathbf{X}_1 = \bar{\lambda}_2\lambda_1 \mathbf{X}_2^\theta \mathbf{X}_1$$

$\Rightarrow \qquad (1 - \bar{\lambda}_2 \lambda_1) \mathbf{X}_2^\theta \mathbf{X}_1 = \mathbf{O}.$

But $\bar{\lambda}_2 \lambda_1 \neq 1$; so $1 - \bar{\lambda}_2 \lambda_1 \neq 0$.

This implies that $\mathbf{X}_2^{\theta} \mathbf{X}_1 = \mathbf{O} \Rightarrow \mathbf{X}_1$ and $\mathbf{X}_2$ are orthogonal.

Example 13: *Show that the two matrices* $\mathbf{A}, \mathbf{C}^{-1}\mathbf{AC}$ *have the same characteristic roots.*

(Lucknow 2005, 07)

Solution: Let $\mathbf{B} = \mathbf{C}^{-1}\mathbf{AC}$.

Then
$$\mathbf{B} - \lambda\mathbf{I} = \mathbf{C}^{-1}\mathbf{AC} - \lambda\mathbf{I}$$
$$= \mathbf{C}^{-1}\mathbf{AC} - \mathbf{C}^{-1}\lambda\mathbf{IC} \qquad [\because \mathbf{C}^{-1}(\lambda\mathbf{I})\mathbf{C} = \lambda\mathbf{C}^{-1}\mathbf{C} = \lambda\mathbf{I}]$$
$$= \mathbf{C}^{-1}(\mathbf{A} - \lambda\mathbf{I})\mathbf{C}.$$

$\therefore$
$$|\mathbf{B} - \lambda\mathbf{I}| = |\mathbf{C}^{-1}||\mathbf{A} - \lambda\mathbf{I}||\mathbf{C}|$$
$$= |\mathbf{A} - \lambda\mathbf{I}||\mathbf{C}^{-1}||\mathbf{C}|$$
$$= |\mathbf{A} - \lambda\mathbf{I}||\mathbf{C}^{-1}\mathbf{C}|$$
$$= |\mathbf{A} - \lambda\mathbf{I}||\mathbf{I}|$$
$$= |\mathbf{A} - \lambda\mathbf{I}|.$$

Thus the two matrices **A** and **B** have the same characteristic determinants and hence the same characteristic equations and the same characteristic roots.

6 The Cayley-Hamilton Theorem

Every square matrix satisfies its characteristic equation i.e., if for a square matrix **A** *of order n,*
$$|\mathbf{A} - \lambda\mathbf{I}| = (-1)^n [\lambda^n + a_1\lambda^{n-1} + a_2\lambda^{n-2} + \ldots + a_n],$$
then the matrix equation $\mathbf{X}^n + a_1\mathbf{X}^{n-1} + a_2\mathbf{X}^{n-2} + a_3\mathbf{X}^{n-3} + \ldots a_n\mathbf{I} = \mathbf{O}$

is satisfied by $\mathbf{X} = \mathbf{A}$ *i.e.,* $\mathbf{A}^n + a_1\mathbf{A}^{n-1} + a_2\mathbf{A}^{n-2} + \ldots + a_n\mathbf{I} = \mathbf{O}$.

(Meerut 2000, 03, 05B, 06B, 09B, 10B; Rohilkhand 05, 08; Agra 07; Avadh 05; Purvanchal 08; Lucknow 08, 09)

Proof: Since the elements of $\mathbf{A} - \lambda\mathbf{I}$ are at most of the first degree in λ, the elements of Adj $(\mathbf{A} - \lambda\mathbf{I})$ are ordinary polynomials in λ of degree $n - 1$ or less. Therefore Adj $(\mathbf{A} - \lambda\mathbf{I})$ can be written as a matrix polynomial in λ, given by
$$\text{Adj}\,(\mathbf{A} - \lambda\mathbf{I}) = \mathbf{B}_0\lambda^{n-1} + \mathbf{B}_1\lambda^{n-2} + \ldots + \mathbf{B}_{n-2}\,\lambda + \mathbf{B}_{n-1},$$
where $\mathbf{B}_0, \mathbf{B}_1, \ldots, \mathbf{B}_{n-1}$ are matrices of the type $n \times n$ whose elements are functions of a_{ij}'s.

Now
$$(\mathbf{A} - \lambda\mathbf{I})\ \text{Adj.}\,(\mathbf{A} - \lambda\mathbf{I}) = |\mathbf{A} - \lambda\mathbf{I}|\ \mathbf{I}. \qquad [\because \mathbf{A}\,.\,\text{Adj}\,\mathbf{A} = |\mathbf{A}|\,\mathbf{I}_n|\,]$$
$\therefore$
$$(\mathbf{A} - \lambda\mathbf{I})(\mathbf{B}_0\lambda^{n-1} + \mathbf{B}_1\lambda^{n-2} + \ldots + \mathbf{B}_{n-2}\lambda + \mathbf{B}_{n-1})$$
$$= (-1)^n [\lambda^n + a_1\lambda^{n-1} + \ldots + a_n]\,\mathbf{I}.$$

Comparing coefficients of like powers of λ on both sides, we get
$$-\mathbf{IB}_0 = (-1)^n\,\mathbf{I},$$

$$\mathbf{AB}_0 - \mathbf{IB}_1 = (-1)^n a_1 \mathbf{I},$$

$$\mathbf{AB}_1 - \mathbf{IB}_2 = (-1)^n a_2 \mathbf{I},$$

....

....

$$\mathbf{AB}_{n-1} = (-1)^n a_n \mathbf{I}.$$

Pre-multiplying these successively by $\mathbf{A}^n, \mathbf{A}^{n-1}, \ldots, \mathbf{I}$ and adding, we get

$$\mathbf{O} = (-1)^n [\mathbf{A}^n + a_1 \mathbf{A}^{n-1} + a_2 \mathbf{A}^{n-2} + \ldots + a_n \mathbf{I}].$$

Thus $\mathbf{A}^n + a_1 \mathbf{A}^{n-1} + a_2 \mathbf{A}^{n-2} + \ldots + a_{n-1} \mathbf{A} + a_n \mathbf{I} = \mathbf{O}.$ ...(1)

Corollary 1: If **A** be a non-singular matrix, then $|\mathbf{A}| \neq 0$. Also $|\mathbf{A}| = (-1)^n a_n$ and therefore $a_n \neq 0$.

Pre-multiplying (1) by $\mathbf{A}^{-1}$, we get

$$\mathbf{A}^{n-1} + a_1 \mathbf{A}^{n-2} + a_2 \mathbf{A}^{n-3} + \ldots + a_{n-1} \mathbf{I} + a_n \mathbf{A}^{-1} = \mathbf{O}$$

or $\mathbf{A}^{-1} = -(1/a_n) [\mathbf{A}^{n-1} + a_1 \mathbf{A}^{n-2} + \ldots + a_{n-1} \mathbf{I}].$

Corollary 2: If m be a positive integer such that $m \geq n$, then multiplying the result (1) by $\mathbf{A}^{m-n}$, we get

$$\mathbf{A}^m + a_1 \mathbf{A}^{m-1} + \ldots + a_n \mathbf{A}^{m-n} = \mathbf{O},$$

showing that any positive integral power $\mathbf{A}^m$ $(m \geq n)$ of **A** is linearly expressible in terms of those of lower order.

Illustrative Examples

Example 14: *Find the characteristic equation of the matrix* $\mathbf{A} = \begin{bmatrix} 2 & -1 & 1 \\ -1 & 2 & -1 \\ 1 & -1 & 2 \end{bmatrix}$ *and verify that it is satisfied by* **A** *and hence obtain* $\mathbf{A}^{-1}$.

(Meerut 2002, 05; Bundelkhand 05, 08, 10, 11; Kanpur 10; Lucknow 05,11; Gorakhpur 15; Kumaun 15)

Solution: We have

$$|\mathbf{A} - \lambda \mathbf{I}| = \begin{vmatrix} 2-\lambda & -1 & 1 \\ -1 & 2-\lambda & -1 \\ 1 & -1 & 2-\lambda \end{vmatrix}$$

$$= (2-\lambda)\{(2-\lambda)^2 - 1\} + 1\{-1(2-\lambda) + 1\} + 1\{1 - (2-\lambda)\}$$

$$= (2-\lambda)(3 - 4\lambda + \lambda^2) + (\lambda - 1) + (\lambda - 1)$$

$$= -\lambda^3 + 6\lambda^2 - 9\lambda + 4.$$

$\therefore$ the characteristic equation of the matrix **A** is

$$\lambda^3 - 6\lambda^2 + 9\lambda - 4 = 0.$$

We are now to verify that

$$\mathbf{A}^3 - 6\mathbf{A}^2 + 9\mathbf{A} - 4\mathbf{I} = \mathbf{O}. \quad ...(1)$$

We have $\mathbf{I} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \mathbf{A} = \begin{bmatrix} 2 & -1 & 1 \\ -1 & 2 & -1 \\ 1 & -1 & 2 \end{bmatrix},$

$$\mathbf{A}^2 = \mathbf{A} \times \mathbf{A} = \begin{bmatrix} 6 & -5 & 5 \\ -5 & 6 & -5 \\ 5 & -5 & 6 \end{bmatrix},$$

$$\mathbf{A}^3 = \mathbf{A}^2\mathbf{A} = \begin{bmatrix} 22 & -21 & 21 \\ -21 & 22 & -21 \\ 21 & -21 & 22 \end{bmatrix}.$$

Now we can verify that

$$\mathbf{A}^3 - 6\mathbf{A}^2 + \mathbf{9A} - \mathbf{4I}$$

$$= \begin{bmatrix} 22 & -21 & 21 \\ -21 & 22 & -21 \\ 21 & -21 & 22 \end{bmatrix} - 6\begin{bmatrix} 6 & -5 & 5 \\ -5 & 6 & -5 \\ 5 & -5 & 6 \end{bmatrix} + 9\begin{bmatrix} 2 & -1 & 1 \\ -1 & 2 & -1 \\ 1 & -1 & 2 \end{bmatrix} - 4\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Multiplying (1) by $\mathbf{A}^{-1}$, we get

$$\mathbf{A}^2 - 6\mathbf{A} + 9\mathbf{I} - 4\mathbf{A}^{-1} = \mathbf{O}.$$

$\therefore$ $$\mathbf{A}^{-1} = \frac{1}{4}(\mathbf{A}^2 - 6\mathbf{A} + 9\mathbf{I})$$

Now $\mathbf{A}^2 - 6\mathbf{A} + 9\mathbf{I}$

$$= \begin{bmatrix} 6 & -5 & 5 \\ -5 & 6 & -5 \\ 5 & -5 & 6 \end{bmatrix} + \begin{bmatrix} -12 & 6 & -6 \\ 6 & -12 & 6 \\ -6 & 6 & -12 \end{bmatrix} + \begin{bmatrix} 9 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 9 \end{bmatrix}$$

$$= \begin{bmatrix} 3 & 1 & -1 \\ 1 & 3 & 1 \\ -1 & 1 & 3 \end{bmatrix}.$$

$\therefore \qquad \mathbf{A}^{-1} = \frac{1}{4}\begin{bmatrix} 3 & 1 & -1 \\ 1 & 3 & 1 \\ -1 & 1 & 3 \end{bmatrix}.$

Example 15: *Obtain the characteristic equation of the matrix* $\mathbf{A} = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 2 & 1 \\ 2 & 0 & 3 \end{bmatrix}$ *and verify that it is satisfied by* **A** *and hence find its inverse.*

(Meerut 2001, 10B; Bundelkhand 09, 11; Kanpur 07; Kumaun 07; Purvanchal 10; Lucknow 07, 10; Gorakhpur 12)

Solution: We have

$$|\mathbf{A} - \lambda \mathbf{I}| = \begin{bmatrix} 1-\lambda & 0 & 2 \\ 0 & 2-\lambda & 1 \\ 2 & 0 & 3-\lambda \end{bmatrix}$$

$$= (1-\lambda)(2-\lambda)(3-\lambda) + 2[0 - 2(2-\lambda)]$$

$$= (2-\lambda)[(1-\lambda)(3-\lambda) - 4]$$

$$= (2-\lambda)[\lambda^2 - 4\lambda - 1]$$

$$= -(\lambda^3 - 6\lambda^2 + 7\lambda + 2).$$

$\therefore$ the characteristic equation of **A** is

$$\lambda^3 - 6\lambda^2 + 7\lambda + 2 = 0. \qquad \ldots(1)$$

By the Cayley-Hamilton theorem

$$\mathbf{A}^3 - 6\mathbf{A}^2 + 7\mathbf{A} + 2\mathbf{I} = \mathbf{O}. \qquad \ldots(2)$$

Verification of (2). We have

$$\mathbf{A}^2 = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 2 & 1 \\ 2 & 0 & 3 \end{bmatrix} \times \begin{bmatrix} 1 & 0 & 2 \\ 0 & 2 & 1 \\ 2 & 0 & 3 \end{bmatrix} = \begin{bmatrix} 5 & 0 & 8 \\ 2 & 4 & 5 \\ 8 & 0 & 13 \end{bmatrix}.$$

Also $\qquad \mathbf{A}^3 = \mathbf{A} \,.\, \mathbf{A}^2$

$$= \begin{bmatrix} 1 & 0 & 2 \\ 0 & 2 & 1 \\ 2 & 0 & 3 \end{bmatrix} \times \begin{bmatrix} 5 & 0 & 8 \\ 2 & 4 & 5 \\ 8 & 0 & 13 \end{bmatrix}$$

$$= \begin{bmatrix} 21 & 0 & 34 \\ 12 & 8 & 23 \\ 34 & 0 & 55 \end{bmatrix}.$$

Now $\qquad \mathbf{A}^3 - 6\mathbf{A}^2 + 7\mathbf{A} + 2\mathbf{I}$

$$= \begin{bmatrix} 21 & 0 & 34 \\ 12 & 8 & 23 \\ 34 & 0 & 55 \end{bmatrix} - 6\begin{bmatrix} 5 & 0 & 8 \\ 2 & 4 & 5 \\ 8 & 0 & 13 \end{bmatrix} + 7\begin{bmatrix} 1 & 0 & 2 \\ 0 & 2 & 1 \\ 2 & 0 & 3 \end{bmatrix} + 2\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$=\begin{bmatrix}21 & 0 & 34\\12 & 8 & 23\\34 & 0 & 55\end{bmatrix}-\begin{bmatrix}30 & 0 & 48\\12 & 24 & 30\\48 & 0 & 78\end{bmatrix}+\begin{bmatrix}7 & 0 & 14\\0 & 14 & 7\\14 & 0 & 21\end{bmatrix}+\begin{bmatrix}2 & 0 & 0\\0 & 2 & 0\\0 & 0 & 2\end{bmatrix}$$

$$=\begin{bmatrix}30 & 0 & 48\\12 & 24 & 30\\48 & 0 & 78\end{bmatrix}-\begin{bmatrix}30 & 0 & 48\\12 & 24 & 30\\48 & 0 & 78\end{bmatrix}$$

$$=\begin{bmatrix}0 & 0 & 0\\0 & 0 & 0\\0 & 0 & 0\end{bmatrix}=\mathbf{O}.$$

Hence Cayley-Hamilton theorem is verified. Now we shall compute $\mathbf{A}^{-1}$.

Multiplying (2) by $\mathbf{A}^{-1}$, we get

$$\mathbf{A}^2-6\mathbf{A}+7\mathbf{I}+2\mathbf{A}^{-1}=\mathbf{O}.$$

$$\therefore \qquad \mathbf{A}^{-1}=-\frac{1}{2}(\mathbf{A}^2-6\mathbf{A}+2\mathbf{I})$$

$$=\begin{bmatrix}-3 & 0 & 2\\-1 & \frac{1}{2} & \frac{1}{2}\\2 & 0 & -1\end{bmatrix}.$$

Example 16: *Show that the matrix* $\mathrm{A}=\begin{bmatrix}0 & c & -b\\-c & 0 & a\\b & -a & 0\end{bmatrix}$ *satisfies Cayley-Hamilton theorem.*

(Meerut 2006B; Kumaun 09)

Solution: **We have**

$$|\mathbf{A}-\lambda\mathbf{I}|=\begin{bmatrix}0-\lambda & c & -b\\-c & 0-\lambda & a\\b & -a & 0-\lambda\end{bmatrix}$$

$$=-\lambda(\lambda^2+a^2)-c(c\lambda-ab)-b\,(ac+b\lambda)$$

$$=-\lambda^3-\lambda(a^2+b^2+c^2).$$

$\therefore$ The characteristic equation of the matrix $\mathbf{A}$ is

$$\lambda^3+\lambda\,(a^2+b^2+c^2)=0.$$

We are to verify that

$$\mathbf{A}^3+(a^2+b^2+c^2)\,\mathbf{A}=\mathbf{O}.$$

We have $\qquad \mathbf{A}^2=\begin{bmatrix}0 & c & -b\\-c & 0 & a\\b & -a & 0\end{bmatrix}\times\begin{bmatrix}0 & c & -b\\-c & 0 & a\\b & -a & 0\end{bmatrix}$

$$= \begin{bmatrix} -c^2 - b^2 & ab & ac \\ ab & -c^2 - a^2 & bc \\ ac & bc & -b^2 - a^2 \end{bmatrix}.$$

$\therefore$ $$\mathbf{A}^3 = \mathbf{A}^2\mathbf{A}$$

$$= \begin{bmatrix} -c^2 - b^2 & ab & ac \\ ab & -c^2 - a^2 & bc \\ ac & bc & -b^2 - a^2 \end{bmatrix} \times \begin{bmatrix} 0 & c & -b \\ -c & 0 & a \\ b & -a & 0 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & -c^3 - b^2c - a^2c & bc^2 + b^3 + a^2b \\ c^3 + a^2c + b^2c & 0 & -ab^2 - ac^2 - a^3 \\ -bc^2 - b^3 - a^2b & ac^2 + ab^2 + a^3 & 0 \end{bmatrix}$$

$$= -(a^2 + b^2 + c^2) \begin{bmatrix} 0 & c & -b \\ -c & 0 & a \\ b & -a & 0 \end{bmatrix}$$

$$= -(a^2 + b^2 + c^2)\mathbf{A}.$$

$\therefore$ $$\mathbf{A}^3 + (a^2 + b^2 + c^2)\mathbf{A} = -(a^2 + b^2 + c^2)\mathbf{A} + (a^2 + b^2 + c^2)\mathbf{A}$$

$$= 0\,\mathbf{A} = \mathbf{O}.$$

Hence **A** satisfies Cayley-Hamilton theorem.

Example 17: *Find the characteristic equation of the matrix* $\mathbf{A} = \begin{bmatrix} 1 & 2 & 0 \\ 2 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$ *and hence find* $\mathbf{A}^{-1}$. **(Rohilkhand 2009, 10; Kumaun 11, 14)**

Solution: The characteristic equation of the matrix **A** is

$$|\mathbf{A} - \lambda\mathbf{I}| = 0$$

or $$\begin{bmatrix} 1-\lambda & 2 & 0 \\ 2 & -1-\lambda & 0 \\ 0 & 0 & -1-\lambda \end{bmatrix} = 0$$

or $$(1-\lambda)(-1-\lambda)(-1-\lambda) - 2[2(-1-\lambda)] = 0$$

or $$(1-\lambda)(1+\lambda)^2 + 4(1+\lambda) = 0$$

or $$(1-\lambda)(1 + 2\lambda + \lambda^2) + 4 + 4\lambda = 0$$

or $$1 + 2\lambda + \lambda^2 - \lambda - 2\lambda^2 - \lambda^3 + 4 + 4\lambda = 0$$

or $$-\lambda^3 - \lambda^2 + 5\lambda + 5 = 0$$

or $$\lambda^3 + \lambda^2 - 5\lambda - 5 = 0. \qquad ...(1)$$

Now by Cayley-Hamilton theorem the matrix A must satisfy its characteristic equation (1). Therefore we have

$$\mathbf{A}^3 + \mathbf{A}^2 - 5\mathbf{A} - 5\mathbf{I} = \mathbf{O}$$

or $$5\mathbf{I} = \mathbf{A}^3 + \mathbf{A}^2 - 5\mathbf{A} \qquad ...(2)$$

Pre-multiplying both sides of (2) by $\mathbf{A}^{-1}$, we have

$$5\mathbf{A}^{-1}\mathbf{I} = \mathbf{A}^{-1}\mathbf{A}^3 + \mathbf{A}^{-1}\mathbf{A}^2 - 5\mathbf{A}^{-1}\mathbf{A}$$

or $$5\mathbf{A}^{-1} = \mathbf{A}^2 + \mathbf{A} - 5\mathbf{I}$$

or $$\mathbf{A}^{-1} = \frac{1}{5}(\mathbf{A}^2 + \mathbf{A} - 5\mathbf{I}) \qquad ...(3)$$

Now $$\mathbf{A}^2 = \begin{bmatrix} 1 & 2 & 0 \\ 2 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}\begin{bmatrix} 1 & 2 & 0 \\ 2 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

$\therefore$ $$\mathbf{A}^2 + \mathbf{A} - 5\mathbf{I} = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 2 & 0 \\ 2 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} - \begin{bmatrix} 5 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 5 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 2 & 0 \\ 2 & -1 & 0 \\ 0 & 0 & -5 \end{bmatrix}.$$

Hence from (3), $$\mathbf{A}^{-1} = \frac{1}{5}\begin{bmatrix} 1 & 2 & 0 \\ 2 & -1 & 0 \\ 0 & 0 & -5 \end{bmatrix}.$$

Example 18: *State Cayley-Hamilton theorem. Use it to express* $2\mathbf{A}^5 - 3\mathbf{A}^4 + \mathbf{A}^2 - 4\mathbf{I}$ *as a linear polynomial in* $\mathbf{A}$, *when* $\mathbf{A} = \begin{bmatrix} 3 & 1 \\ -1 & 2 \end{bmatrix}$.

(Lucknow 2005)

Solution: **Statement of Cayley-Hamilton Theorem**. Every square matrix satisfies its characteristic equation.

Now let us find the characteristic equation of the matrix **A**. We have

$$|\mathbf{A} - \lambda\mathbf{I}| = \begin{vmatrix} 3-\lambda & 1 \\ -1 & 2-\lambda \end{vmatrix} = (3-\lambda)(2-\lambda) + 1$$

$$= \lambda^2 - 5\lambda + 7.$$

The characteristic equation of **A** is $|\mathbf{A} - \lambda\mathbf{I}| = 0$ *i.e.,* is

$$\lambda^2 - 5\lambda + 7 = 0 \qquad ...(1)$$

By Cayley-Hamilton theorem,the matrix A must satisfy (1). Therefore we have

$$\mathbf{A}^2 - 5\mathbf{A} + 7\mathbf{I} = \mathbf{O}. \qquad ...(2)$$

From (2), we get

$$\mathbf{A}^2 = 5\mathbf{A} - 7\mathbf{I} \qquad ...(3)$$

Multiplying both sides of (3) by **A**, we get

$$A^3 = 5A^2 - 7A \qquad \ldots(4)$$

$\therefore$ $$A^4 = 5A^3 - 7A^2 \qquad \ldots(5)$$

and $$A^5 = 5A^4 - 7A^3 \qquad \ldots(6)$$

Now $$2A^5 - 3A^4 + A^2 - 4I = 2(5A^4 - 7A^3) - 3A^4 + A^2 - 4I$$

[Substituting for A^5 from (6)]

$$= 7A^4 - 14A^3 + A^2 - 4I$$

$$= 7(5A^3 - 7A^2) - 14A^3 + A^2 - 4I \qquad \text{[By (5)]}$$

$$= 21A^3 - 48A^2 - 4I$$

$$= 21(5A^2 - 7A) - 48A^2 - 4I \qquad \text{[By (4)]}$$

$$= 57A^2 - 147A - 4I$$

$$= 57(5A - 7I) - 147A - 4I \qquad \text{[By (3)]}$$

$= 138A - 403I$, which is a linear polynomial in **A**.

7 Diagonalization of Square Matrices with Distinct Eigen Values

Similarity of Matrices: Definition: *Let* **A** *and* **B** *be square matrices of order n. Then* **B** *is said to be similar to* **A** *if there exists a non-singular matrix* **P** *such that*

$$B = P^{-1}AP.$$

Theorem 1: *Similarity of matrices is an equivalence relation.*

Theorem 2: *Similar matrices have the same determinant.*

Theorem 3: *Similar matrices have the same characteristic polynomial and hence the same eigenvalues. If* **X** *is an eigenvector of* **A** *corresponding to the eigenvalue* λ, *then* $P^{-1}X$ *is an eigenvector of* **B** *corresponding to the eigenvalue* λ, *where*

$$B = P^{-1}AP.$$

Corollary: *If* **A** *is similar to a diagonal matrix* **D**, *the diagonal elements of* **D** *are the eigenvalues of* **A**.

Diagonalizable matrix: Definition: *A matrix* **A** *is said to be diagonalizable if it is similar to a diagonal matrix.*

Thus a matrix **A** is diagonalizable if there exists an invertible matrix **P** such that $P^{-1}AP = D$ where **D** is a diagonal matrix. Also the matrix **P** is then said to diagonalize **A** or transform **A** to diagonal form.

Theorem 1: *An $n \times n$ matrix is diagonalizable if and only if it possesses n linearly independent eigenvectors.*

Proof: Suppose **A** is diagonalizable. Then **A** is similar to a diagonal matrix **D** = dia. $[\lambda_1, \lambda_2, \ldots, \lambda_n]$. Therefore there exists an invertible matrix $P = [X_1, X_2, \ldots, X_n]$ such that

$$\mathbf{P}^{-1}\mathbf{AP} = \mathbf{D}$$

i.e., $\mathbf{AP} = \mathbf{PD}$

i.e., $\mathbf{A}[\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n] = [\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n]\,\text{dia.}[\lambda_1, \lambda_2, \dots, \lambda_n]$

i.e, $[\mathbf{AX}_1, \mathbf{AX}_2, \dots, \mathbf{AX}_n] = [\lambda_1\mathbf{X}_1, \lambda_2\mathbf{X}_2, \dots, \lambda_n\mathbf{X}_n]$

i.e., $\mathbf{AX}_1 = \lambda_1\mathbf{X}_1,\ \mathbf{AX}_2 = \lambda_2\mathbf{X}_2, \dots, \mathbf{AX}_n = \lambda_n\mathbf{X}_n.$

Therefore $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$ are eigenvectors of **A** corresponding to the eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$ respectively. Since the matrix **P** is non-singular, therefore its column vectors $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$ are linearly independent. Hence **A** possesses n linearly independent eigenvectors.

Conversely, suppose that **A** possesses n linearly independent eigenvectors $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$ and let $\lambda_1, \lambda_2, \dots, \lambda_n$ be the corresponding eigenvalues. Then $\mathbf{AX}_1 = \lambda_1\mathbf{X}_1, \mathbf{AX}_2 = \lambda_2\mathbf{X}_2, \dots, \mathbf{AX}_n = \lambda_n\mathbf{X}_n.$

Let $\mathbf{P} = [\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n]$ and $\mathbf{D} = \text{dia.}\ [\lambda_1, \lambda_2, \dots, \lambda_n]$.

Then
$$\begin{aligned}\mathbf{AP} &= [\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n] = [\mathbf{AX}_1, \mathbf{AX}_2, \dots, \mathbf{AX}_n]\\ &= [\lambda_1\mathbf{X}_1, \lambda_2\mathbf{X}_2, \dots, \lambda_n\mathbf{X}_n]\\ &= [\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n]\ \text{dia.}\ [\lambda_1, \lambda_2, \dots, \lambda_n] = \mathbf{PD}.\end{aligned}$$

Since the column vectors $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$ of the matrix **P** are linearly independent, therefore **X** is invertible and $\mathbf{X}^{-1}$ exists.

Therefore $\mathbf{AP} = \mathbf{PD}$

$\Rightarrow$ $\mathbf{P}^{-1}\mathbf{AP} = \mathbf{P}^{-1}\mathbf{PD}$

$\Rightarrow$ $\mathbf{P}^{-1}\mathbf{AP} = \mathbf{D}$

$\Rightarrow$ **A** is similar to a diagonal matrix **D**

$\Rightarrow$ **A** is diagonalizable.

Remark: In the proof of the above theorem we have shown that if **A** is diagonalizable and **P** diagonalizes **A**, then

$$\mathbf{P}^{-1}\mathbf{AP} = \begin{bmatrix}\lambda_1 & 0 & \dots & 0\\ 0 & \lambda_2 & \dots & 0\\ \dots & \dots & \dots & \dots\\ 0 & 0 & \dots & \lambda_n\end{bmatrix} = \mathbf{D}$$

if and only if the j^{th} column of **P** is an eigenvector of **A** corresponding to the eigenvalue λ_j of **A**, $(j = 1, 2, \dots, n)$. The diagonal elements of **D** are the eigenvalues of **A** and they occur in the same order as is the order of their corresponding eigenvectors in the column vectors of **P**.

Theorem 2: *If the eigenvalues of an $n \times n$ matrix are all distinct then it is always similar to a diagonal matrix.*

Proof: Let **A** be a square matrix of order n and suppose it has n distinct eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$. We know that eigenvectors of a matrix corresponding to distinct eigenvalues are linearly independent. Therefore **A** has n linearly independent eigenvectors and so it is similar to a diagonal matrix

$$\mathbf{D} = \text{dia. } [\lambda_1, \lambda_2, \ldots, \lambda_n].$$

Corollary: *Two $n \times n$ matrices with the same set of n distinct eigenvalues are similar.*

Proof: Suppose **A** and **B** are two $n \times n$ matrices with the same set of n distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$. Let $\mathbf{D} =$ dia. $[\lambda_1, \lambda_2, \ldots, \lambda_n]$. Then both **A** and **B** are similar to **D.** Now **A** is similar to **D** and **D** is similar to **B** implies that **A** is similar to **B**.

Note that the relation of similarity is transitive.

Theorem 3: *The necessary and sufficient condition for a square matrix to be similar to a diagonal matrix is that the geometric multiplicity of each of its eigenvalues coincides with the algebraic multiplicity.*

Proof: The condition is necessary. Suppose **A** is similar to a diagonal matrix **D** = dia. $[\lambda_1, \lambda_2, \ldots, \lambda_n]$. Then $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the eigenvalues of **A** and there exists a non-singular matrix **P** such that

$$\mathbf{P}^{-1}\mathbf{AP} = \mathbf{D}.$$

Let α be an eigenvalue of **A** of algebraic multiplicity k. Then exactly k among $\lambda_1, \lambda_2, \ldots, \lambda_n$, are equal to α.

Let m = rank $(\mathbf{A} - \alpha\mathbf{I})$. Then the system of equations

$$(\mathbf{A} - \alpha\mathbf{I})\ \mathbf{X} = \mathbf{O}$$

have $n - m$ linearly independent solutions and so $n - m$ will be the geometric multiplicity of α. We are to prove that $k = n - m$. We know that the rank of a matrix does not change on multiplication by a non-singular matrix. Therefore

$$\begin{aligned}
\text{rank } (\mathbf{A} - \alpha\mathbf{I}) &= \text{rank } [\mathbf{P}^{-1}(\mathbf{A} - \alpha\mathbf{I})\mathbf{P}] \\
&= \text{rank } [\mathbf{P}^{-1}\mathbf{AP} - \alpha\mathbf{I}] \\
&= \text{rank } [\mathbf{D} - \alpha\mathbf{I}] = \text{rank dia. } [\lambda_1 - \alpha, \lambda_2 - \alpha, \ldots, \lambda_n - \alpha] \\
&= n - k, \text{ since exactly } k \text{ elements of dia. } [\lambda_1 - \alpha, \lambda_2 - \alpha, \ldots, \lambda_n - \alpha] \\
&\qquad\qquad \text{are equal to zero.}
\end{aligned}$$

Thus rank $(\mathbf{A} - \alpha\mathbf{I}) = m = n - k$. Therefore $k = n - m$.

Thus there are exactly k linearly independent eigenvectors corresponding to the eigenvalue α.

The condition is sufficient. Suppose that the geometric multiplicity of each eigenvalue of **A** is equal to its algebraic multiplicity. Let $\lambda_1, \ldots, \lambda_p$ be the set of p distinct eigenvalues of **A** with respective multiplicities $r_1, \ldots, r_p$. We have

$$r_1 + \ldots + r_p = n.$$

To prove that **A** is diagonalizable.

Let
$$\left.\begin{matrix}
\mathbf{C}_{11}, & \mathbf{C}_{12}, & \ldots, & \mathbf{C}_{1r_1} \\
\ldots & \ldots & \ldots & \ldots \\
\ldots & \ldots & \ldots & \ldots \\
\mathbf{C}_{p1}, & \mathbf{C}_{p2}, & \ldots, & \mathbf{C}_{pr_p}
\end{matrix}\right\} \qquad \ldots(1)$$

be linearly independent sets of eigenvectors corresponding to the eigenvalues $\lambda_1, \ldots, \lambda_p$ respectively. We claim that the n vectors given in (1) are linearly independent. Let

$$(a_{11}\mathbf{C}_{11} + a_{12}\,\mathbf{C}_{12} + \ldots + a_{1r_1}\,\mathbf{C}_{1r_1}) + \ldots + (a_{p1}\mathbf{C}_{p1} + \ldots + a_{pr_p}\mathbf{C}_{pr_p}) = \mathbf{O} \quad \ldots(2)$$

The relation (3) may be written as

$$\mathbf{X}_1 + \mathbf{X}_2 + \ldots + \mathbf{X}_p = \mathbf{O}, \quad \ldots(3)$$

where $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_p$ denote the vectors written within brackets in (2) *i.e.*,

$$\mathbf{X}_1 = a_{11}\mathbf{C}_{11} + \ldots + a_{1r_1}\mathbf{C}_{1r_1}, \text{ and so on.}$$

Now $\mathbf{X}_1$ is a linear combination of eigenvectors of **A** corresponding to the eigenvalue λ_1. Therefore if $\mathbf{X}_1 \neq \mathbf{O}$, then $\mathbf{X}_1$ is also an eigenvector of **A** corresponding to the eigenvalue λ_1.

Similarly we can speak for $\mathbf{X}_2, \ldots, \mathbf{X}_p$.

In case some one of $\mathbf{X}_1, \ldots, \mathbf{X}_p$ is not zero, then the relation (3) implies that a system of eigenvectors of **A** corresponding to distinct eigenvalues of **A** is linearly dependent. But this is not possible. Hence each of the vectors $\mathbf{X}_1, \mathbf{X}_2 \ldots, \mathbf{X}_p$ must be zero.

Since $\mathbf{C}_{11}, \mathbf{C}_{12}, \ldots, \mathbf{C}_{1r_1}$ is a set of linearly independent vectors, therefore $\mathbf{O} = \mathbf{X}_1 = a_{11}\,\mathbf{C}_{11} + \ldots + a_{1r_1}\,\mathbf{C}_{1r_1}$ implies that

$$a_{11} = 0, \ldots, a_{1r_1} = 0.$$

Similarly we can show that each of the scalars in relation (2) is zero. Therefore the n vectors give in (1) are linearly independent. Thus **A** has n linearly independent eigenvectors. So it is similar to a diagonal matrix.

Illustrative Examples

Example 19: *Show that the rank of every matrix similar to* **A** *is the same as that of* **A.**

Solution: Let **B** be a matrix similar to **A.** Then there exists a non-singular matrix **P** such that $\mathbf{B} = \mathbf{P}^{-1}\mathbf{AP}$. We know that the rank of a matrix does not change on multiplication by a non-singular matrix.

Therefore, rank $(\mathbf{P}^{-1}\mathbf{AP}) = \text{rank } \mathbf{A} \Rightarrow \text{rank } \mathbf{B} = \text{rank } \mathbf{A}$.

Example 20: *If* **U** *be a unitary matrix such that* $\mathbf{U}^\theta \mathbf{AU} = \text{diag}\,[\lambda_1, \ldots, \lambda_n]$, *show that* $\lambda_1, \ldots, \lambda_n$ *are the eigenvalues of* **A.**

Solution: Let diag $[\lambda_1, \ldots, \lambda_n] = \mathbf{D}$. Since **U** is unitary, therefore $\mathbf{U}^\theta = \mathbf{U}^{-1}$. So

$$\mathbf{U}^\theta \mathbf{AU} = \mathbf{D} \Rightarrow \mathbf{U}^{-1}\mathbf{AU} = \mathbf{D}.$$

Thus **A** is similar to the diagonal matrix **D.** But similar matrices have the same eigenvalues and eigenvalues of **D** are its diagonal elements. Therefore $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of **A.**

Example 21: *Prove that if* **A** *is similar to a diagonal matrix, then* $\mathbf{A}^T$ *is similar to* **A.**

Solution: Suppose **A** is similar to a diagonal matrix **D.** Then there exists a non-singular matrix **P** such that

$$\mathbf{P}^{-1}\,\mathbf{AP} = \mathbf{D}$$

$$\Rightarrow \qquad \mathbf{A} = \mathbf{P\,D\,P}^{-1}$$

$$\Rightarrow \qquad \mathbf{A}^T = (\mathbf{P\,D\,P}^{-1})^T = (\mathbf{P}^{-1})^T\,\mathbf{D}^T\,\mathbf{P}^T$$

$$\Rightarrow \qquad \mathbf{A}^T = (\mathbf{P}^T)^{-1}\,\mathbf{D\,P}^T \qquad [\because\ \mathbf{D} \text{ is diagonal} \Rightarrow \mathbf{D}^T = \mathbf{D}]$$

$\Rightarrow$ $\mathbf{A}^T$ is similar to **D**

$\Rightarrow$ **D** is similar to $\mathbf{A}^T$.

Finally **A** is similar to **D** and **D** is similar to $\mathbf{A}^T$ implies that **A** is similar to $\mathbf{A}^T$.

Example 22: *Show that a non-zero matrix is nilpotent if and only if all its eigenvalues are equal to zero.*

Solution: Suppose $\mathbf{A} \neq \mathbf{O}$ and **A** is nilpotent. Then

$$\mathbf{A}^r = \mathbf{O}, \text{ for some positive integer } r$$

$\Rightarrow$ the polynomial λ^r annihilates **A**

$\Rightarrow$ the minimal polynomial $m(\lambda)$ of **A** divides λ^r

$\Rightarrow$ $m(\lambda)$ is of the type λ^s, where s is some positive integer

$\Rightarrow$ 0 is the only root of $m(\lambda)$

$\Rightarrow$ 0 is the only eigenvalue of **A**

$\Rightarrow$ all eigenvalues of **A** are zero.

Conversely, each eigenvalue of $\mathbf{A} = 0$

$\Rightarrow$ characteristic equation of **A** is $\lambda^n = 0$

$\Rightarrow$ $\mathbf{A}^n = \mathbf{O}$, since **A** satisfies its characteristic equation

$\Rightarrow$ **A** is nilpotent.

Note: A non-zero matrix **A** is said to be **nilpotent**, if for some positive integer r,

$$\mathbf{A}^r = \mathbf{O}.$$

Example 23: *A square matrix* **A** *is defined by* $\mathbf{A} = \begin{bmatrix} 1 & 2 & -2 \\ 1 & 2 & 1 \\ -1 & -1 & 0 \end{bmatrix}$.

Find the modal matrix **P** *and the resulting diagonal matrix* **D** *of* **A**.

Solution: The characteristic equation of **A** is

$$\begin{vmatrix} 1-\lambda & 2 & -2 \\ 1 & 2-\lambda & 1 \\ -1 & -1 & 0-\lambda \end{vmatrix} = 0$$

or $$(1-\lambda)(-2\lambda + \lambda^2 + 1) - 2(-\lambda + 1) - 2(-1 + 2 - \lambda) = 0$$

or $$(-\lambda + 1)(\lambda - 1)^2 + (4\lambda - 4) = 0$$

or $$(-\lambda + 1)(\lambda + 1)(\lambda - 3) = 0$$

The roots of this equation are $1, -1, 3$.

The eigenvectors **X** of **A** corresponding to the eigenvalue 1 are given by the equation

$$(\mathbf{A} - 1\,\mathbf{I})\,\mathbf{X} = \mathbf{O}$$

or $$(\mathbf{A} - \mathbf{I})\,\mathbf{X} = \mathbf{O}$$

or $$\begin{bmatrix} 0 & 2 & -2 \\ 1 & 1 & 1 \\ -1 & -1 & -1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

or $$\begin{bmatrix} 0 & 2 & -2 \\ 1 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \text{ by } R_3 \to R_3 + R_1$$

The coefficient matrix of these equations is of rank 2. So these equations have 1 linearly independent solution. These equations can be written as

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

From these, we get $y = z = 1$, say. Then $x = -2$.

Therefore $\mathbf{X}_1 = \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$ is an eigenvector of **A** corresponding to the eigenvalue 1.

Now the eigenvectors of **A** corresponding to the eigenvalues -1 are given by

$$(\mathbf{A} + \mathbf{I})\,\mathbf{X} = \mathbf{O}$$

i.e., $$\begin{bmatrix} 2 & 2 & -2 \\ 1 & 3 & 1 \\ -1 & -1 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

or $$\begin{bmatrix} 2 & 2 & -2 \\ 1 & 3 & 1 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \text{ by } R_3 \to 2R_3 + R_1$$

The coefficient matrix of these equations is of rank 2. So these equations have 1 linearly independent solution.

These equations can be written as

$$2x + 2y - 2z = 0\ ,\ x + 3y + z = 0$$

Let us take $z = -1$, then $y = 1$ and $x = -2$. Therefore, $\mathbf{X}_2 = \begin{bmatrix} -2 \\ 1 \\ -1 \end{bmatrix}$

is an eigenvector of **A** corresponding to the eigenvalue -1.

Now the eigenvectors of **A** corresponding to the eigenvalue 3 are given by

$$(\mathbf{A} - 3\mathbf{I})\,\mathbf{X} = \mathbf{O}$$

i.e.,
$$\begin{bmatrix} -2 & 2 & -2 \\ 1 & -1 & 1 \\ -1 & -1 & -3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

or
$$\begin{bmatrix} -2 & 2 & -2 \\ -1 & -1 & -3 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \; R_2 \leftrightarrow R_3$$

Similarly, we have $x = 2, y = 1, z = -1$.

Therefore $\mathbf{X}_3 = \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}$

is an eigenvector of **A** corresponding to the eigenvalue 3.

Let modal matrix

$$\mathbf{P} = [\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3] = \begin{bmatrix} -2 & -2 & 2 \\ 1 & 1 & 1 \\ 1 & -1 & 1 \end{bmatrix}$$

$$\therefore \qquad \mathbf{P}^{-1} = \frac{\text{Adj}\,\mathbf{P}}{|\mathbf{P}|} = -\frac{1}{8} \begin{bmatrix} 0 & -4 & -4 \\ 2 & 0 & 4 \\ -2 & -4 & 0 \end{bmatrix}.$$

The matrix **P** will transform **A** to diagonal form **D** which is given by the relation

$$\mathbf{P}^{-1}\mathbf{A}\,\mathbf{P} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 3 \end{bmatrix} = \mathbf{D}.$$

Example 24: *Show that the matrix* $\mathbf{A} = \begin{bmatrix} \mathbf{8} & \mathbf{-8} & \mathbf{-2} \\ \mathbf{4} & \mathbf{-3} & \mathbf{-2} \\ \mathbf{3} & \mathbf{-4} & \mathbf{1} \end{bmatrix}$

is diagonalizable. Also find the transforming matrix and diagonal matrix.

Solution: The characteristic equation of **A** is

$$\begin{vmatrix} 8-\lambda & -8 & -2 \\ 4 & -3-\lambda & -2 \\ 3 & -4 & 1-\lambda \end{vmatrix} = 0$$

or
$$\begin{vmatrix} 1-\lambda & -1+\lambda & -1+\lambda \\ 4 & -3-\lambda & -2 \\ 3 & -4 & 1-\lambda \end{vmatrix} = 0, \text{ applying } R_1 - (R_2 + R_3)$$

or $$(1-\lambda)\begin{vmatrix} 1 & -1 & -1 \\ 4 & -3-\lambda & -2 \\ 3 & -4 & 1-\lambda \end{vmatrix} = 0$$

or $$(1-\lambda)\begin{vmatrix} 1 & 0 & 0 \\ 4 & 1-\lambda & 2 \\ 3 & -1 & 4-\lambda \end{vmatrix} = 0 \text{ , applying } C_2 + C_1, C_3 + C_1$$

or $$(1-\lambda)[(1-\lambda)(4-\lambda)+2] = 0$$

or $$(1-\lambda)(\lambda^2 - 5\lambda + 6) = 0$$

or $$(1-\lambda)(\lambda-2)(\lambda-3) = 0.$$

The roots of this equation are 1, 2, 3.

Since the eigenvalues of the matrix **A** are all distinct, therefore **A** is similar to a diagonal matrix. Since the algebraic multiplicity of each eigenvalue of **A** is 1, therefore there will be one and only one linearly independent eigenvector of **A** corresponding to each eigenvalue of **A.**

The eigenvectors **X** of **A** corresponding to the eigenvalue 1 are given by the equation

$$(\mathbf{A} - 1\mathbf{I})\,\mathbf{X} = \mathbf{O} \quad \text{or} \quad (\mathbf{A} - \mathbf{I})\,\mathbf{X} = \mathbf{O}$$

or $$\begin{bmatrix} 7 & -8 & -2 \\ 4 & -4 & -2 \\ 3 & -4 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

or $$\begin{bmatrix} 7 & -8 & -2 \\ -3 & 4 & 0 \\ 3 & -4 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \text{ by } R_2 \to R_2 - R_1$$

or $$\begin{bmatrix} 7 & -8 & -2 \\ -3 & 4 & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \text{ by } R_3 \to R_3 + R_2.$$

The matrix of coefficients of these equations has rank 2. Therefore these equations have only one linearly independent solution as it should have been because the algebraic multiplicity of the eigenvalue 1 is 1. Note that the geometric multiplicity cannot exceed the algebraic multiplicity. The above equations can be written as $7x_1 - 8x_2 - 2x_3 = 0,\ -3x_1 + 4x_2 = 0$. From the last equation, we get $x_1 = 4, x_2 = 3$. Then the first gives $x_3 = 2$. Therefore $\mathbf{X}_1 = \begin{bmatrix} 4 \\ 3 \\ 2 \end{bmatrix}$ is an eigenvector of **A** corresponding to the eigenvalue 1.

The eigenvectors **X** of **A** corresponding to the eigenvalue 2 are given by the equation

$$(\mathbf{A} - 2\mathbf{I})\,\mathbf{X} = \mathbf{O}$$

or $$\begin{bmatrix} 6 & -8 & -2 \\ 4 & -5 & -2 \\ 3 & -4 & -1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

or $$\begin{bmatrix} 6 & -8 & -2 \\ -2 & 3 & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \text{ applying } R_2 \to R_3 - R_1, \; R_3 \to R_3 - \frac{1}{2}R_1.$$

These equations can be written as $6x_1 - 8x_2 - 2x_3 = 0, \; -2x_1 + 3x_2 = 0$. From these, we get $x_1 = 3, x_2 = 2, x_3 = 1$.

Therefore $\mathbf{X}_2 = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$ is an eigenvector of **A** corresponding to the eigenvalue 2.

The eigenvectors **X** of **A** corresponding to the eigenvalue 3 are given by the equation

$$(\mathbf{A} - 3\mathbf{I})\,\mathbf{X} = \mathbf{O}$$

or $$\begin{bmatrix} 5 & -8 & -2 \\ 4 & -6 & -2 \\ 3 & -4 & -2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

or $$\begin{bmatrix} 5 & -8 & -2 \\ -1 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \text{ applying } R_2 \to R_2 - R_1, \; R_3 \to R_3 + R_1 - 2R_2.$$

These equations can be written as

$$5x_1 - 8x_2 - 2x_3 = 0, \; -x_1 + 2x_2 = 0.$$

From these, we get $x_1 = 2, x_2 = 1, x_3 = 1$.

$\therefore$ $\mathbf{X}_3 = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}$ is an eigenvector of **A** corresponding to the eigenvalue 3.

Let $$\mathbf{P} = [\mathbf{X}_1 \; \mathbf{X}_2 \; \mathbf{X}_3] = \begin{bmatrix} 4 & 3 & 2 \\ 3 & 2 & 1 \\ 2 & 1 & 1 \end{bmatrix}.$$

The columns of **P** are linearly independent eigenvectors of **A** corresponding to the eigenvalues 1, 2, 3 respectively. The matrix **P** will transform **A** to diagonal form **D** which is given by the relation

$$\mathbf{P}^{-1}\,\mathbf{A}\,\mathbf{P} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} = \mathbf{D}.$$

Comprehensive Exercise 1

1. (i) State Cayley-Hamilton theorem.

(Bundelkhand 2005; Agra 05; Lucknow 07; Gorakhpur 11)

(ii) Find the characteristic roots of the matrix $\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & -4 & 2 \\ 0 & 0 & 7 \end{bmatrix}$.

(Meerut 2009 B; Gorakhpur 10)

2. (i) If $a + b + c = 0$, find the characteristic roots of the matrix

$$\mathbf{A} = \begin{bmatrix} a & c & b \\ c & b & a \\ b & a & c \end{bmatrix}.$$

(Kumaun 2008; Kanpur 07, 08)

(ii) Prove that the matrices :

$$\mathbf{A} = \begin{bmatrix} 0 & a & b \\ a & 0 & c \\ b & c & 0 \end{bmatrix} \text{ and } \mathbf{B} = \begin{bmatrix} 0 & b & a \\ b & 0 & c \\ a & c & 0 \end{bmatrix}$$

have same characteristic equation.

(Rohilkhand 2007; Purvanchal 06; Agra 05)

(iii) Verify that the matrices

$$\mathbf{A} = \begin{bmatrix} 0 & h & g \\ h & 0 & f \\ g & f & 0 \end{bmatrix}, \mathbf{B} = \begin{bmatrix} 0 & f & h \\ f & 0 & g \\ h & g & 0 \end{bmatrix}, \mathbf{C} = \begin{bmatrix} 0 & g & f \\ g & 0 & h \\ f & h & 0 \end{bmatrix}$$

have the same characteristic equation

$$\lambda^3 - (f^2 + g^2 + h^2)\lambda - 2fgh = 0.$$

3. (i) Verify Cayley-Hamilton theorem for matrix **A**, where $\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$.

(Rohilkhand 2011)

(ii) Verify that the matrix **A** satisfies its characteristic equation, where

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

Express $\mathbf{A}^6 - 4\mathbf{A}^5 + 8\mathbf{A}^4 - 12\mathbf{A}^3 + 14\mathbf{A}^2$ as a linear polynomial in **A**.

(Gorakhpur 2014)

(iii) Determine the characteristic roots for the matrix $\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$ **(Kumaun 2014)**

4. Find the characteristic roots of the matrix $\mathbf{A}=\begin{bmatrix}1 & 4\\ 2 & 3\end{bmatrix}$ and verify Cayley-Hamilton theorem for this matrix. Find the inverse of the matrix **A** and also express $\mathbf{A}^5-4\mathbf{A}^4-7\mathbf{A}^3+11\mathbf{A}^2-\mathbf{A}-10\mathbf{I}$ as a linear polynomial in **A**.

(Gorakhpur 2011)

5. Verify that the matrix $\mathbf{A}=\begin{bmatrix}1 & 2 & 1\\ 0 & 1 & -1\\ 3 & -1 & 1\end{bmatrix}$ satisfies its characteristic equation and compute $\mathbf{A}^{-1}$. **(Purvanchal 2009)**

6. Find the characteristic roots of the matrix $\mathbf{A}=\begin{bmatrix}1 & 1 & 3\\ 5 & 2 & 6\\ -2 & -1 & -3\end{bmatrix}$ and verify Cayley-Hamilton theorem. **(Agra 2008; Bundelkhand 11; Meerut 09B)**

7. Determine the eigenvalues and eigenvectors of the matrix

$$\mathbf{A}=\begin{bmatrix}1 & 0 & -2\\ 0 & 0 & 0\\ -2 & 0 & 4\end{bmatrix}.$$

(Kumaun 2008)

8. Show that the matrix $\mathbf{A}=\begin{bmatrix}2 & 2 & 1\\ 1 & 3 & 1\\ 1 & 2 & 2\end{bmatrix}$ satisfies Cayley-Hamilton theorem.

Also determine the characteristic roots (*i.e.*, latent roots) and the corresponding characteristic vectors of the matrix **A**. **(Meerut 2007)**

9. Verify Cayley-Hamilton theorem for the matrix $\mathbf{A}=\begin{bmatrix}0 & 0 & 1\\ 3 & 1 & 0\\ -2 & 1 & 4\end{bmatrix}$.

Hence or otherwise evaluate $\mathbf{A}^{-1}$.

(Meerut 2003, 09; Avadh 05; Kumaun 11)

10. Find the characteristic equation of the matrix $\mathbf{A}=\begin{bmatrix}1 & 3 & 7\\ 4 & 2 & 3\\ 0 & 2 & 1\end{bmatrix}$ and show that it is satisfied by **A**. Hence obtain the inverse of the given matrix **A**.

11. Verify that the matrix $\mathbf{A}=\begin{bmatrix}1 & 2 & 1\\ -1 & 0 & 3\\ 2 & -1 & 1\end{bmatrix}$ satisfies its own characteristic equation. Is it true of every square matrix? State the theorem that applies here.

12. Verify Cayley-Hamilton theorem for the following matrix :

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 2 \\ 0 & -1 & 1 \\ 0 & 1 & 0 \end{bmatrix}.$$

(Avadh 2008)

13. Find the characteristic roots and the characteristic spaces of the matrix

$$\begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & 3 \\ 0 & 0 & 2 \end{bmatrix}.$$

(Meerut 2005; Rohilkhand 07)

14. Show that the characteristic roots of a triangular matrix are just the diagonal elements of the matrix. **(Gorakhpur 2011)**

15. Let **A** and **B** be n-rowed square matrices and let **A** be non-singular. Show that the matrices $\mathbf{A}^{-1}\mathbf{B}$ and $\mathbf{B}\mathbf{A}^{-1}$ have the same eigenvalues.

16. If **A** and **B** are non-singular matrices of order n, show that the matrices **AB** and **BA** are similar.

17. **A** and **B** are two $n \times n$ matrices with the same set of n distinct eigenvalues. Show that there exist two matrices **P** and **Q** (one of them non-singular) such that

$$\mathbf{A} = \mathbf{PQ},\ \mathbf{B} = \mathbf{QP}.$$

18. Prove that a non-zero nilpotent matrix cannot be similar to a diagonal matrix.

19. Find a matrix **P** which diagonalizes the matrix $\mathbf{A} = \begin{bmatrix} 4 & 1 \\ 2 & 3 \end{bmatrix}$. Verify that $\mathbf{P}^{-1}\ \mathbf{AP} = \mathbf{D}$, where **D** is the diagonal matrix.

20. Reduce the matrix $\mathbf{A} = \begin{bmatrix} 1 & -1 & 2 \\ 0 & 2 & -1 \\ 0 & 0 & 3 \end{bmatrix}$ to diagonal form.

Answers 1

1. (ii) 1, –4, 7

2. (i) $\lambda = \mathbf{0}, \pm \left\{\frac{3}{2}(a^2 + b^2 + c^2)\right\}^{1/2}$

3. (ii) $-4\mathbf{A} + 5\mathbf{I}$. (iii) $-1, 3$

4. $5, -1; \frac{1}{5}\begin{bmatrix} -3 & 4 \\ 2 & -1 \end{bmatrix}$

5. Characteristic equation is $\lambda^3 - 3\lambda^2 - \lambda + 9 = 0$

$$\mathbf{A}^{-1} = \frac{1}{9}\begin{bmatrix} 0 & 3 & 3 \\ 3 & 2 & -1 \\ 3 & -7 & -1 \end{bmatrix}$$

6. All the characteristic roots are zero.

7. Eigenvalues are $5, 0, 0$. Corresponding to the eigenvalue 2 an eigenvector is $[1, 0, -2]'$. Two linearly independent eigenvectors corresponding to the eigenvalue 0 are $[2, 0, 1]'$ and $[0, 1, 0]'$

8. 5, 1, 1. Corresponding to the characteristic root 5 a characteristic vector is $[1, 1, 1]'$. Two linearly independent characteristic vectors corresponding to the characteristic root 1 are $[1, 0, -1]'$ and $[1, 1, -3]'$

9. $\mathbf{A}^{-1} = \frac{1}{5}\begin{bmatrix} 4 & 1 & -1 \\ -12 & 2 & 3 \\ 5 & 0 & 0 \end{bmatrix}$

10. $\lambda^3 - 4\lambda^2 - 13\lambda - 40 = 0$; $\mathbf{A}^{-1} = \frac{1}{40}\begin{bmatrix} -4 & 11 & -5 \\ -4 & 1 & 25 \\ 8 & -2 & -10 \end{bmatrix}$

13. Characteristic roots are 1, 2, 2. Corresponding to the characteristic root 1 a characteristic vector is $[1, 0, 0]'$. Corresponding to the characteristic root 2 a characteristic vector is $[2, 1, 0]'$. The characteristic space corresponding to the characteristic root 1 consists of the vector $c\,[1, 0, 0]'$, where c is any scalar. Similarly for the other characteristic space

19. 2,5. Corresponding to the characteristic root 2 a characteristic vector is $[1, -2]'$. Corresponding to the characteristic root 5 a characteristic vector is $[1, 1]'$. $P = \begin{bmatrix} 1 & 1 \\ -2 & 1 \end{bmatrix}$

20. 1, 2, 3 corresponding to the characteristic root 1 a characteristic vector is $[1, 0, 0]'$. Corresponding to the characteristic root 2 a characteristic vector is $[1, -1, 0]'$. Corresponding to the characteristic root 3 a characteristic vector is $[3, -2, 2]'$.

Objective Type Questions

Multiple Choice Questions

Indicate the correct answer for each question by writing the corresponding letter from (a), (b), (c) *and* (d).

1. If λ is a characteristic root of a matrix $\mathbf{A}$, then a characteristic root of $\mathbf{A}^{-1}$ is

(a) $1/\lambda$ (b) λ

(c) λ^2 (d) $1/\lambda^2$ **(Bundelkhand 2001)**

2. The eigenvalues of the matrix $\mathbf{A} = \begin{bmatrix} a & h & g \\ 0 & b & 0 \\ 0 & c & c \end{bmatrix}$ are

(a) a, h, g (b) a, h, c
(c) a, g, c (d) a, b, c **(Kanpur 2009, 11)**

3. If λ is a characteristic root of the matrix $\mathbf{A}$, then a characteristic root of the matrix $\mathbf{A} + k\mathbf{I}$ is

(a) λ (b) $k + \lambda$
(c) $k - \lambda$ (d) none of these **(Agra 2007)**

4. The characteristic equation of $\mathbf{A} = \begin{bmatrix} 3 & 1 \\ -1 & 2 \end{bmatrix}$ is

(a) $\lambda^2 - 5\lambda + 7 = 0$ (b) $\lambda^2 - 3\lambda + 7 = 0$
(c) $\lambda^2 - 2\lambda + 7 = 0$ (d) none of these

(Bundelkhand 2007)

5. At least one characteristic root of every singular matrix is :

(a) 1 (b) -1
(c) 0 (d) none of these

6. If $\mathbf{A}$ is a non-singular matrix, then the relationship between the eigen value of $\mathbf{A}$ and $\mathbf{A}^{-1}$ is :

(a) they are same (b) they are equal but opposite
(c) reciprocal to each other (d) none of these

(Garhwal 2012)

7. If matrix $\mathbf{A} = \begin{bmatrix} 8 & -6 & 2 \\ -6 & 7 & -4 \\ 2 & -4 & 3 \end{bmatrix}$, then the characteristic roots of $\mathbf{A}$ are

(a) $0, 3, 14$ (b) $3, 5, 15$
(c) $0, 5, 15$ (d) none of these

(Garhwal 2013)

8. If $\mathbf{A}$ be the matrix,

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & -4 & 2 \\ 0 & 0 & 7 \end{bmatrix},$$

then its characteristic roots are

(a) $1, -5, 7$ (b) $1, -4, 7$
(c) $-1, 4, 7$ (d) $1, 4, -7$

(Garhwal 2014)

9. The characteristic equation of $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ -1 & 3 \end{bmatrix}$ is

(a) $\lambda^2 - 3\lambda + 7 = 0$ (b) $\lambda^2 - 4\lambda + 5 = 0$

(c) $\lambda^2 + 2\lambda = 0$ (d) $\lambda^2 + 3\lambda + 7 = 0$

10. The characteristic equation of $\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ is

(a) $\lambda^2 - 2\lambda$ (b) $\lambda^2 + 2\lambda = 0$

(c) $\lambda^2 = 0$ (d) $\lambda^2 = 1$

Fill in the Blank(s)

Fill in the blanks "………", so that the following statements are complete and correct.

1. If $\mathbf{A}$ is any n-rowed square matrix and λ an indeterminate, then the equation $|\mathbf{A} - \lambda\mathbf{I}| = 0$ is called the ……… of $\mathbf{A}$ and the roots of this equation are called the ……… of the matrix $\mathbf{A}$.

2. If λ is a characteristic root of an $n \times n$ matrix $\mathbf{A}$, then a non-zero vector $\mathbf{X}$ such that $\mathbf{AX} = \lambda\mathbf{X}$ is called a ……… of $\mathbf{A}$ corresponding to the characteristic root λ.

3. The characteristic roots of a Hermitian matrix are ……… **(Avadh 2005)**

4. The characteristic roots of a skew-Hermitian matrix are either ……… or ………

5. The matrices $\mathbf{A}$ and $\mathbf{A}'$ have the ……… eigenvalues. **(Agra 2008)**

6. If $\lambda_1, \ldots\ldots, \lambda_n$ are the eigenvalues of $\mathbf{A}$, then $k\lambda_1, \ldots\ldots, k\lambda_n$ are the eigenvalues of ……… **(Lucknow 2011)**

7. If the characteristic roots of $\mathbf{A}$ are $\lambda_1, \lambda_2, \ldots\ldots, \lambda_n$, then the characteristic roots of ……… are $\lambda_1^2, \lambda_2^2, \ldots\ldots, \lambda_n^2$.

8. Every square matrix ….. its characteristic equation. **(Meerut 2001)**

9. The characteristic equation of the matrix $\mathbf{A} = \begin{bmatrix} 1 & 0 & 2 \\ 0 & -1 & 1 \\ 0 & 1 & 0 \end{bmatrix}$ is ……… .

True or False

Write 'T' for true and 'F' for false statement.

1. The characteristic roots of a real symmetric matrix are all real.

2. The characteristic roots of a real skew-symmetric matrix are all pure imaginary.

3. The characteristic roots of a unitary matrix are of unit modulus.

4. The characteristic roots of a diagonal matrix are just the diagonal elements of the matrix.

5. Two matrices **A** and $\mathbf{C}^{-1}\mathbf{AC}$ do not have the same characteristic roots.

6. Cayley-Hamilton theorem states that 'Every square matrix satisfies its characteristic equation'. **(Rohilkhand 2006)**

7. The characteristic equation of $\mathbf{A} = \begin{bmatrix} 3 & 1 \\ -1 & 2 \end{bmatrix}$ is $\lambda^2 - 3\lambda + 7 = 0$.

(Meerut 2003; Kumaun 08)

Answers

Multiple Choice Questions

1.	(a)	**2.**	(d)	**3.**	(b)	**4.**	(a)	**5.**	(c)
6.	(c)	**7.**	(d)	**8.**	(b)	**9.**	(b)	**10.**	(a)

Fill in the Blank(s)

1. characteristic equations; characteristic roots
2. characteristic vector
3. real
4. pure imaginary; zero
5. same
6. $k\mathbf{A}$
7. $\mathbf{A}^2$
8. satisfies
9. $\lambda^3 - 2\lambda + 1 = 0$

True or False

1.	*T*	**2.**	*F*	**3.**	*T*	**4.**	*T*	**5.**	*F*
6.	*T*	**7.**	*F*						

□

Chapter 5

Bilinear and Quadratic Forms

1 External Direct Product of Two Vectors Spaces

Suppose U and V are two vector spaces over the same field F. Let

$$W = U \times V \quad i.e., \quad W = \{(\alpha, \beta) : \alpha \in U, \beta \in V\}.$$

If (α_1, β_1) and (α_2, β_2) are two elements in W, then we define their equality as follows :

$$(\alpha_1, \beta_1) = (\alpha_2, \beta_2) \text{ if } \alpha_1 = \alpha_2 \text{ and } \beta_1 = \beta_2 .$$

Also we define the sum of (α_1, β_1) and (α_2, β_2) as follows :

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

If c is any element in F and (α, β) is any element in W, then we define scalar multiplication in W as follows :

$$c\,(\alpha, \beta) = (c\alpha, c\beta).$$

It can be easily shown that with respect to addition and scalar multiplication as defined above, W is a vector space over the field F. We call W as the external direct product of the vector spaces U and V and we shall write $W = U \oplus V$.

Now we shall consider some special type of scalar-valued functions on W known as *bilinear forms.*

2 Bilinear Forms

Definition: *Let U and V be two vector spaces over the same field F. A bilinear form on* $W = U \oplus V$ *is a function f from W into F, which assigns to each element* (α, β) *in W a scalar* $f(\alpha, \beta)$ *in such a way that*

$$f(a\alpha_1 + b\alpha_2, \beta) = a f(\alpha_1, \beta) + b f(\alpha_2, \beta)$$

and

$$f(\alpha, a\beta_1 + b\beta_2) = a f(\alpha, \beta_1) + b f(\alpha, \beta_2).$$

Here $f(\alpha, \beta)$ is an element of F. It denotes the image of (α, β) under the function f. Thus a bilinear form on W is a function from W into F which is linear as a function of either of its arguments when the other is fixed.

If $U = V$, then in place of saying that f is a bilinear form on $W = U \oplus V$, we shall simply say that f is a bilinear form on V.

Thus if V is a vector space over the field F, then a bilinear form on V is a function f, which assigns to each ordered pair of vectors α, β *in V a scalar* $f(\alpha, \beta)$ *in F, and which satisfies*

$$f(a\alpha_1 + b\alpha_2, \beta) = a f(\alpha_1, \beta) + b f(\alpha_2, \beta)$$

and

$$f(\alpha, a\beta_1 + b\beta_2) = a f(\alpha, \beta_1) + b f(\alpha, \beta_2).$$

Illustrative Examples

Example 1: *Suppose V is a vector space over the field F. Let* L_1, L_2 *be linear functionals on V. Let f be a function from* $V \times V$ *into F defined as*

$$f(\alpha, \beta) = L_1(\alpha) L_2(\beta).$$

Then f is a bilinear form on V. **(Kumaun 2008)**

Solution: If $\alpha, \beta \in V$, then $L_1(\alpha), L_2(\beta)$ are scalars. We have

$$\begin{aligned} f(a\alpha_1 + b\alpha_2, \beta) &= L_1(a\alpha_1 + b\alpha_2) L_2(\beta) \\ &= [a L_1(\alpha_1) + b L_1(\alpha_2)] L_2(\beta) \\ &= aL_1(\alpha_1) L_2(\beta) + bL_1(\alpha_2) L_2(\beta) \\ &= af(\alpha_1, \beta) + bf(\alpha_2, \beta). \end{aligned}$$

Also

$$\begin{aligned} f(\alpha, a\beta_1 + b\beta_2) &= L_1(\alpha) L_2(a\beta_1 + b\beta_2) \\ &= L_1(\alpha) [a L_2(\beta_1) + b L_2(\beta_2)] \\ &= aL_1(\alpha) L_2(\beta_1) + b L_1(\alpha) L_2(\beta_2) \\ &= a f(\alpha, \beta_1) + b f(\alpha, \beta_2). \end{aligned}$$

Hence f is a bilinear form on V.

Example 2: *Suppose V is a vector space over the field F. Let T be a linear operator on V and f a bilinear form on V. Suppose g is a function from* $V \times V$ *into F defined as*

$$g(\alpha, \beta) = f(T\alpha, T\beta).$$

Then g is a bilinear form on V.

Solution: We have $g(a\alpha_1 + b\alpha_2, \beta) = f(T(a\alpha_1 + b\alpha_2), T\beta)$

$$= f(aT\alpha_1 + bT\alpha_2, T\beta)$$
$$= a f(T\alpha_1, T\beta) + b f(T\alpha_2, T\beta)$$
$$= a g(\alpha_1, \beta) + b g(\alpha_2, \beta).$$

Also $g(\alpha, a\beta_1 + b\beta_2) = f(T\alpha, T(a\beta_1 + b\beta_2))$

$$= f(T\alpha, aT\beta_1 + bT\beta_2)$$
$$= a f(T\alpha, T\beta_1) + b f(T\alpha, T\beta_2)$$
$$= a g(\alpha, \beta_1) + b g(\alpha, \beta_2).$$

Hence g is a bilinear form on V.

3 Bilinear Form as Vectors

Suppose U and V are two vector spaces over the field F. Let $L(U, V, F)$ denote the set of all bilinear forms on $U \times V$. We can impose a vector space structure on $L(U, V, F)$ over the field F. For this purpose we define addition and scalar multiplication in $L(U, V, F)$ as follows :

Addition of bilinear forms: Suppose f, g are two bilinear forms on $U \times V$. Then we define their sum as follows :

$$(f + g)(\alpha, \beta) = f(\alpha, \beta) + g(\alpha, \beta).$$

It can be easily seen that $f + g$ is also a bilinear form on $U \times V$. We have

$$(f + g)(a\alpha_1 + b\alpha_2, \beta) = f(a\alpha_1 + b\alpha_2, \beta) + g(a\alpha_1 + b\alpha_2, \beta)$$
$$= [a f(\alpha_1, \beta) + b f(\alpha_2, \beta)] + [a g(\alpha_1, \beta) + b g(\alpha_2, \beta)]$$
$$= a[f(\alpha_1, \beta) + g(\alpha_1, \beta)] + b[f(\alpha_2, \beta) + g(\alpha_2, \beta)]$$
$$= a[(f + g)(\alpha_1, \beta)] + b[(f + g)(\alpha_2, \beta)].$$

Similarly, we can show that

$$(f + g)(\alpha, a\beta_1 + b\beta_2) = a[(f + g)(\alpha, \beta_1)] + b[(f + g)(\alpha, \beta_2)].$$

Hence $f + g$ is a bilinear form on $U \times V$.

Thus $L(U, V, F)$ is closed with respect to addition defined on it.

Scalar multiplication of bilinear forms. Suppose f is a bilinear form on $U \times V$ and c is a scalar.

Then we define cf as follows :

$$(cf)(\alpha, \beta) = cf(\alpha, \beta) \ \forall \ (\alpha, \beta) \in U \times V.$$

Obviously cf is a function from $U \times V$ into F. We have

$$(cf)(a\alpha_1 + b\alpha_2, \beta) = cf(a\alpha_1 + b\alpha_2, \beta)$$
$$= c[a f(\alpha_1, \beta) + b f(\alpha_2, \beta)]$$
$$= caf(\alpha_1, \beta) + cbf(\alpha_2, \beta)$$
$$= a[cf(\alpha_1, \beta)] + b[cf(\alpha_2, \beta)]$$
$$= a[(cf)(\alpha_1, \beta)] + b[(cf)(\alpha_2, \beta)].$$

Similarly, we can show that

$$(cf)(\alpha, a\beta_1 + b\beta_2) = a[(cf)(\alpha, \beta_1)] + b[(cf)(\alpha, \beta_2)].$$

Therefore cf is also a bilinear form on $U \times V$.

Thus $L(U, V, F)$ is closed with respect to scalar multiplication defined on it.

Important: *It can be easily seen that the set of all bilinear forms on $U \times V$ is a vector space over the field F with respect to addition and scalar multiplication just defined above.*

The bilinear from $\mathbf{0}$ will act as the zero vector of this space. The bilinear form $-f$ will act as the additive inverse of the vector f.

Theorem 1: *If U is an n-dimensional vector space with basis $\{\alpha_1, \ldots, \alpha_n\}$, if V is an m-dimensional vector space with basis $\{\beta_1, \ldots, \beta_m\}$, and if $\{a_{ij}\}$ is any set of nm scalars $(i = 1, \ldots, n;\ j = 1, \ldots, m)$, then there is one and only one bilinear form f on $U \oplus V$ such that $f(\alpha_i, \beta_j) = a_{ij}$ for all i and j.*

Proof: Let $\alpha = \sum_{i=1}^{n} x_i \alpha_i \in U$ and $\beta = \sum_{j=1}^{m} y_j \beta_j \in V$.

Let us define a function f from $U \times V$ into F such that

$$f(\alpha, \beta) = \sum_{i=1}^{n} \sum_{j=1}^{m} x_i y_j a_{ij}. \qquad \ldots(1)$$

We shall show that f is a bilinear form on $U \times V$.

Let $a, b \in F$ and let $\alpha_1, \alpha_2 \in U$.

Let $\alpha_1 = \sum_{i=1}^{n} a_i \alpha_i,\ \alpha_2 = \sum_{i=1}^{n} b_i \alpha_i$.

Then $f(\alpha_1, \beta) = \sum_{i=1}^{n} \sum_{j=1}^{m} a_i y_j a_{ij}$

and $f(\alpha_2, \beta) = \sum_{i=1}^{n} \sum_{j=1}^{m} b_i y_j a_{ij}$.

Also $a\alpha_1 + b\alpha_2 = a \sum_{i=1}^{n} a_i \alpha_i + b \sum_{i=1}^{n} b_i \alpha_i = \sum_{i=1}^{n} (aa_i + bb_i) \alpha_i$.

$$\therefore \quad f(a\alpha_1 + b\alpha_2, \beta) = \sum_{i=1}^{n} \sum_{j=1}^{m} (aa_i + bb_i) y_j a_{ij}$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{m} aa_i y_j a_{ij} + \sum_{i=1}^{n} \sum_{j=1}^{m} bb_i y_j a_{ij}$$

$$= a \sum_{i=1}^{n} \sum_{j=1}^{m} a_i y_j a_{ij} + b \sum_{i=1}^{n} \sum_{j=1}^{m} b_i y_j a_{ij}$$

$$= af(\alpha_1, \beta) + bf(\alpha_2, \beta).$$

Similarly, we can prove that if $a, b \in F$, and $\beta_1, \beta_2 \in V$, then

$$f(\alpha, a\beta_1 + b\beta_2) = af(\alpha, \beta_1) + bf(\alpha, \beta_2).$$

Therefore f is a bilinear form on $U \times V$.

Now $\alpha_i = 0\alpha_1 + \ldots + 0\alpha_{i-1} + 1\alpha_i + 0\alpha_{i+1} + \ldots + 0\alpha_n$

and $\beta_j = 0\beta_1 + \ldots + 0\beta_{j-1} + 1\beta_j + 0\beta_{j+1} + \ldots + 0\beta_m$.

Therefore from (1), we have

$$f\ (\alpha_i\ ,\beta_j) = a_{ij}\ .$$

Thus there exists a bilinear form f on $U \times V$ such that

$$f\ (\alpha_i\ ,\beta_j) = a_{ij}\ .$$

Now to show that f is unique.

Let g be a bilinear form on $U \times V$ such that $g\ (\alpha_i\ ,\beta_j) = a_{ij}\ .$ ...(2)

If $\alpha = \sum_{i=1}^{n} x_i\ \alpha_i$ be in U and $\beta = \sum_{j=1}^{m} y_j\ \beta_j$ be in V, then

$$g\ (\alpha\ ,\beta) = g\left(\sum_{i=1}^{n} x_i\ \alpha_i,\ \sum_{j=1}^{m} y_j\ \beta_j\right)$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{m} x_i\ y_j\ g\ (\alpha_i,\beta_j) \qquad [\because\ g \text{ is a bilinear form}]$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{m} x_i\ y_j\ a_{ij} \qquad [\text{From (2)}]$$

$$= f\ (\alpha\ ,\beta). \qquad [\text{From (1)}]$$

$\therefore$ By the equality of two functions, we have

$$g = f\ .$$

Thus f is unique.

4 Matrix of a Bilinear Form

Definition: *Let V be a finite-dimensional vector space and let $B = \{\alpha_1, \ldots, \alpha_n\}$ 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* $\mathbf{A} = [a_{ij}]_{n \times n}$ *such that*

$$f\ (\alpha_i\ ,\alpha_j) = a_{ij}\ ,\ \ i = 1, \ldots, n\ ;\ \ j = 1, \ldots, n.$$

We shall denote this matrix **A** *by* $[\ f\ ]_B$.

5 Rank of a Bilinear Form

Definition: *The rank of a bilinear form is defined as the rank of the matrix of the form in any ordered basis.*

Let us describe all bilinear forms on a finite-dimensional vector space V of dimension n.

If $\alpha = \sum_{i=1}^{n} x_i\ \alpha_i$, and $\beta = \sum_{j=1}^{n} y_j\ \alpha_j$ are vectors in V, then

$$f\ (\alpha, \beta) = f\left(\sum_{i=1}^{n} x_i\ \alpha_i\ ,\ \sum_{j=1}^{n} y_j\ \alpha_j\right) = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i\ y_j\ f\ (\alpha_i\ ,\alpha_j)$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} x_i \, y_j \, a_{ij} = \mathbf{X}' \, \mathbf{A}\mathbf{Y},$$

where **X** and **Y** are coordinate matrices of α and β in the ordered basis **B** and **X′** is the transpose of the matrix **X**. Thus

$$f(\alpha, \beta) = [\alpha]_B' \, \mathbf{A} \, [\beta]_B .$$

From the definition of the matrix of a bilinear form, we note that if f is a bilinear form on an n-dimensional vector space V over the field F and B is an ordered basis of V, then there exists a unique $n \times n$ matrix $\mathbf{A} = [a_{ij}]_{n \times n}$ over the field F such that

$$\mathbf{A} = [f]_B .$$

Conversely, if $\mathbf{A} = [a_{ij}]_{n \times n}$ be an $n \times n$ matrix over the field F, then from theorem 1, we see that there exists a unique bilinear form f on V such that

$$[f]_B = [a_{ij}]_{n \times n} .$$

If $\alpha = \sum_{i=1}^{n} x_i \, \alpha_i, \; \beta = \sum_{j=1}^{n} y_j \, \alpha_j$ are vectors in V, then the bilinear form f is defined as

$$f(\alpha, \beta) = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i \, y_j \, a_{ij}$$

$$= \mathrm{X}' \, \mathrm{AY}, \qquad \ldots(1)$$

where **X, Y** are the coordinate matrices of α, β in the ordered basis B. Hence the bilinear forms on V are precisely those obtained from an $n \times n$ matrix as in (1).

Illustrative Examples

Example 3: *Let f be the bilinear form on $V_2(\mathbf{R})$ defined by*

$$f((x_1, y_1), (x_2, y_2)) = x_1 \, y_1 + x_2 \, y_2 .$$

Find the matrix of f in the ordered basis

$$B = \{(1, -1), (1, 1)\} \text{ of } V_2(\mathbf{R}).$$

Solution: Let $B = \{\alpha_1, \alpha_2\}$

where $\alpha_1 = (1, -1), \alpha_2 = (1, 1)$.

We have $f(\alpha_1, \alpha_1) = f((1, -1), (1, -1)) = -1 - 1 = -2$,

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

$f(\alpha_2, \alpha_1) = f((1, 1), (1, -1)) = 1 - 1 = 0$,

$f(\alpha_2, \alpha_2) = f((1, 1), (1, 1)) = 1 + 1 = 2$.

$$\therefore \qquad [f]_B = \begin{bmatrix} -2 & 0 \\ 0 & 2 \end{bmatrix}.$$

6 Degenerate and Non-degenerate Bilinear Forms

Definitions: *A bilinear form* f *on a vector space* V *is called degenerate if*

(i) for each non-zero α *in* V, $f(\alpha, \beta) = 0$ *for all* β *in* V *and*

(ii) for each non-zero β *in* V, $f(\alpha, \beta) = 0$ *for all* α *in* V.

A bilinear form is called non-degenerate if it is not degenerate. In other words *a bilinear form* f *on a vector space* V *is called non-degenerate if*

(i) for each $\mathbf{0} \neq \alpha \in V$, *there is a* β *in* V *such that* $f(\alpha, \beta) \neq 0$ *and*

(ii) for each $\mathbf{0} \neq \beta \in V$, *there is an* α *in* V *such that* $f(\alpha, \beta) \neq 0$.

7 Symmetric Bilinear Form

Definition: *Let* f *be a bilinear form of the vector space* V. *Then* f *is said to be symmetric if* $f(\alpha, \beta) = f(\beta, \alpha)$ *for all vectors* α, β *in* V.

Theorem: *If* V *is a finite-dimensional vector space, then a bilinear form* f *on* V *is symmetric if and only if its matrix* $\mathbf{A}$ *in some (or every) ordered basis is symmetric, i.e.,* $\mathbf{A}' = \mathbf{A}$.

Proof: Let B be an ordered basis for V. Let α, β be any two vectors in V. Let X, Y be the coordinate matrices of the vectors α and β respectively in the ordered basis B. If f is a bilinear form on V and $\mathbf{A}$ is the matrix of f in the ordered basis B, then

$$f(\alpha, \beta) = X' AY,$$

and $$f(\beta, \alpha) = Y' AX$$

$\therefore$ f will be symmetric if and only if

$$X' AY = Y' AX$$

for 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

$$Y' A' X = Y' AX \text{ for all column matrices } X \text{ and } Y$$

i.e., $$\mathbf{A}' = \mathbf{A}$$

i.e., $\mathbf{A}$ is symmetric.

Hence the theorem.

8 Quadratic Form Associated with the Bilinear Form

Definition : *Let* f *be a bilinear form on a vector space* V *over the field* F. *Then the* **quadratic form on** V **associated with the bilinear form** f *is the function* q *from* V *into* F *defined by*

$$q(\alpha) = f(\alpha, \alpha) \text{ for all } \alpha \text{ in } V.$$

Theorem 1: *Let V be a vector space 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.* **(Kumaun 2008)**

Proof: Let f be a symmetric bilinear form on V and q be the quadratic form on V associated with f. For all α, β in V, we have

$$\begin{aligned} q(\alpha+\beta) &= f(\alpha+\beta, \alpha+\beta) \\ &= f(\alpha, \alpha+\beta) + f(\beta, \alpha+\beta) \\ &= f(\alpha, \alpha) + f(\alpha, \beta) + f(\beta, \alpha) + f(\beta, \beta) \\ &= q(\alpha) + f(\alpha, \beta) + f(\alpha, \beta) + q(\beta) \\ &= q(\alpha) + (1+1) f(\alpha, \beta) + q(\beta) \end{aligned}$$

$$\therefore \quad (1+1) f(\alpha, \beta) = q(\alpha+\beta) - q(\alpha) - q(\beta). \quad \ldots(1)$$

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 : If F is a subfield of the complex numbers the symmetric bilinear form f is completely determined by its associated quadratic form according to the polarization identity

$$f(\alpha, \beta) = \frac{1}{4} q(\alpha+\beta) - \frac{1}{4} q(\alpha-\beta).$$

As in theorem 1, we have

$$2f(\alpha, \beta) = q(\alpha+\beta) - q(\alpha) - q(\beta). \quad \ldots(1)$$

Also

$$\begin{aligned} q(\alpha-\beta) &= f(\alpha-\beta, \alpha-\beta) = f(\alpha, \alpha-\beta) - f(\beta, \alpha-\beta) \\ &= f(\alpha, \alpha) - f(\alpha, \beta) - f(\beta, \alpha) + f(\beta, \beta) \\ &= q(\alpha) + q(\beta) - 2f(\alpha, \beta). \end{aligned}$$

$$\therefore \quad 2f(\alpha, \beta) = q(\alpha) + q(\beta) - q(\alpha-\beta). \quad \ldots(2)$$

Adding (1) and (2), we get

$$4f(\alpha, \beta) = q(\alpha+\beta) - q(\alpha-\beta)$$

$$\Rightarrow \quad f(\alpha, \beta) = \frac{1}{4} q(\alpha+\beta) - \frac{1}{4} q(\alpha-\beta).$$

Theorem 2: *Let V be 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.* **(Kumaun 2008)**

Proof : In order to prove the theorem, we should find an ordered basis $B = \{\alpha_1, \ldots, \alpha_n\}$ for V such that $f(\alpha_i, \alpha_j) = 0$ for $i \neq j$.

If $f = \hat{\mathbf{0}}$ or $n = 1$, the theorem is obviously true. So let us suppose that $f \neq \hat{\mathbf{0}}$ and $n > 1$.

If $f(\alpha, \alpha) = 0$ for every α in V, then $q(\alpha) = 0$ for every α in V where q is the quadratic form associated with f. Therefore from the polarization identity $f(\alpha, \beta) = \frac{1}{4} q(\alpha+\beta) - \frac{1}{4} q(\alpha-\beta)$ we see that $f(\alpha, \beta) = 0$ for all α, β in V and thus $f = \mathbf{0}$ which is a contradiction. Therefore, there must be a vector α_1 in V such that $f(\alpha_1, \alpha_1) = q(\alpha_1) \neq 0$.

Let W_1 be the one-dimensional subspace of V spanned by the vector α_1 and let W_2 be the set of all vector 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.

First we show that W_1 and W_2 are disjoint.

Let $\gamma \in W_1 \cap W_2$. Then $\gamma \in W_1$ and $\gamma \in W_2$.

But $\gamma \in W_1 \Rightarrow \gamma = c\alpha_1$ for some scalar c.

Also $\quad \gamma \in W_2 \Rightarrow f(\alpha_1, \gamma) = 0$

$\Rightarrow \quad f(\alpha_1, c\alpha_1) = 0$

$\Rightarrow \quad c\, f(\alpha_1, \alpha_1) = 0$

$\Rightarrow \quad c = 0 \qquad [\because f(\alpha_1, \alpha_1) \neq 0]$

$\Rightarrow \quad \gamma = \mathbf{0}\,\alpha_1 = \mathbf{0}$.

$\therefore$ W_1 and W_2 are disjoint.

Now we shall show that $V = W_1 + W_2$.

Let γ be any vector 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
$$\begin{aligned} f(\alpha_1, \beta) &= f\left(\alpha_1, \gamma - \frac{f(\gamma, \alpha_1)}{f(\alpha_1, \alpha_1)}\, \alpha_1\right) \\ &= f(\alpha_1, \gamma) - \frac{f(\gamma, \alpha_1)}{f(\alpha_1, \alpha_1)}\, f(\alpha_1, \alpha_1) \\ &= f(\alpha_1, \gamma) - f(\gamma, \alpha_1) \\ &= f(\alpha_1, \gamma) - f(\alpha_1, \gamma) \qquad [\because f \text{ is symmetric}] \\ &= 0. \end{aligned}$$

$\therefore$ $\beta \in W_2$ by definition of W_2.

Also by definition of W_1 the vector $\dfrac{f(\gamma, \alpha_1)}{f(\alpha_1, \alpha_1)}\, \alpha_1$ is in W_1.

$$\therefore \quad \gamma = \frac{f(\gamma, \alpha_1)}{f(\alpha_1, \alpha_1)}\, \alpha_1 + \beta \in W_1 + W_2.$$

Hence $\quad \boldsymbol{V = W_1 + W_2}$.

$\therefore \quad \boldsymbol{V = W_1 \oplus W_2}$.

So $\quad \dim W_2 = \dim V - \dim W_1 = n - 1$.

Now let g be the restriction of f from V to W_2. Then g is a symmetric bilinear form on W_2 and $\dim W_2$ is less than $\dim V$.

Now we may assume by induction that W_2 has a basis $\{\alpha_2, \ldots, \alpha_n\}$ such that

$$g(\alpha_i, \alpha_j) = 0, \; i \neq j \; (i \geq 2, j \geq 2)$$

$\Rightarrow \quad f(\alpha_i, \alpha_j) = 0, \; i \neq j \; (i \geq 2, j \geq 2) \qquad [\because g \text{ is restriction of } f]$

Now by definition of W_2, we have

$$f(\alpha_i, \alpha_j) = 0 \quad \text{for} \quad j = 2, 3, \ldots, n.$$

Since $\{\alpha_1\}$ is a basis for W_1 and $V = W_1 \oplus W_2$, therefore $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ is a basis for V such that

$$f(\alpha_i, \alpha_j) = 0 \text{ for } i \neq j.$$

Corollary : *Let F be a subfield of the complex numbers, and let A be a symmetric $n \times n$ matrix over F. Then there is an invertible $n \times n$ matrix P over F such that $P' AP$ is diagonal.*

Proof: Let V be a finite dimensional vector space over the field F and let B be an ordered basis for V. Let f be the bilinear form on V such that $[f]_B = A$. Since A is a symmetric matrix, therefore the bilinear form f is also symmetric. Therefore by the above theorem there exists an ordered basis B' of V such that $[f]_{B'}$ is a diagonal matrix. If P is the transition matrix from B to B', then P is an invertible matrix and

$$[f]_{B'} = P'AP$$

i.e. $P'AP$ is a diagonal matrix.

9 Skew-symmetric Bilinear Forms

Definition : *Let f be a bilinear form on the vector space V. Then f is said to be skew-symmetric if $f(\alpha, \beta) = -f(\beta, \alpha)$ for all vectors α, β in V.*

Theorem 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.*

(Kumaun 2007, 09)

Proof : Let f be a bilinear form on a vector space V. Let

$$g(\alpha, \beta) = \frac{1}{2}[f(\alpha, \beta) + f(\beta, \alpha)] \quad \ldots(1)$$

$$h(\alpha, \beta) = \frac{1}{2}[f(\alpha, \beta) - f(\beta, \alpha)] \quad \ldots(2)$$

for all α, β in V.

Then it can be easily seen that both g and h are bilinear forms on V. We have

$$g(\beta, \alpha) = \frac{1}{2}[f(\beta, \alpha) + f(\alpha, \beta)] = g(\alpha, \beta).$$

$\therefore$ g is symmetric.

Also $$h(\beta, \alpha) = \frac{1}{2}[f(\beta, \alpha) - f(\alpha, \beta)] = -\frac{1}{2}[f(\alpha, \beta) - f(\beta, \alpha)]$$

$$= -h(\alpha, \beta).$$

$\therefore$ h is skew-symmetric.

Adding (1) and (2), we get

$$g(\alpha, \beta) + h(\alpha, \beta) = f(\alpha, \beta)$$

$\Rightarrow$ $(g + h)(\alpha, \beta) = f(\alpha, \beta)$ for all α, β in V.

$\therefore$ $g + h = f$.

Now suppose that $f = f_1 + f_2$ where f_1 is symmetric and f_2 is skew-symmetric.

Then $f(\alpha, \beta) = (f_1 + f_2)(\alpha, \beta)$

or $\quad f(\alpha,\beta)=f_1(\alpha,\beta)+f_2(\alpha,\beta)$. ...(3)

Also $\quad f(\beta,\alpha)=(f_1+f_2)(\beta,\alpha)$

or $\quad f(\beta,\alpha)=f_1(\beta,\alpha)+f_2(\beta,\alpha)$

or $\quad f(\beta,\alpha)=f_1(\alpha,\beta)-f_2(\alpha,\beta)$. ...(4)

[$\because$ f_1 is symmetric and f_2 is skew-symmetric]

Adding (3) and (4), we get

$$f(\alpha,\beta)+f(\beta,\alpha)=2f_1(\alpha,\beta)$$

i.e. $\quad f_1(\alpha,\beta)=\frac{1}{2}[f(\alpha,\beta)+f(\beta,\alpha)]=g(\alpha,\beta)$.

$\therefore \quad f_1=g$.

Subtracting (4) from (3) , we get

$$f(\alpha,\beta)-f(\beta,\alpha)=2f_2(\alpha,\beta)$$

i.e. $\quad f_2(\alpha,\beta)=\frac{1}{2}[f(\alpha,\beta)-f(\beta,\alpha)]=h(\alpha,\beta)$.

$\therefore \quad f_2=h$.

Thus the resolution $f=g+h$ is unique.

Theorem 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 $\boldsymbol{A}$ in some (or every) ordered basis is skew-symmetric, i.e.,* $\boldsymbol{A}'=-\boldsymbol{A}$.

Proof : Let B be an ordered basis for V. Let α, β be any two vectors in V. Let X, Y be co-ordinate matrices for the vectors α and β respectively in the ordered basis B. If f is a bilinear form on V and $\boldsymbol{A}$ is the matrix of f in the ordered basis B, then

$$f(\alpha,\beta)=X'AY,$$

and $\quad f(\beta,\alpha)=Y'AX$

$\therefore$ f will be skew-symmetric if and only if

$$X'AY=-Y'AX$$

for 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 skew-symmetric if and only if

$$Y'A'X=-Y'AX \text{ for all column matrices } X \text{ and } Y$$

i.e., $\quad \boldsymbol{A}'=-\boldsymbol{A}$

i.e., $\boldsymbol{A}$ is skew-symmetric.

10 Bilinear Form Corresponding to a Given Matrix

Bilinear Form. Definition: *Let V_m and V_n be two vector spaces over the same field F and let* $\mathbf{A}=[a_{ij}]_{m\times n}$ *be an $m\times n$ matrix over the field F.*

Let $\mathbf{X} = \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ \dots \\ x_m \end{bmatrix}$ *and* $\mathbf{Y} = \begin{bmatrix} y_1 \\ y_2 \\ \dots \\ \dots \\ y_n \end{bmatrix}$ *be any two elements of* V_m *and* V_n *respectively so that*

$$\mathbf{X}^T = \text{the transpose of the column matrix } \mathbf{X}$$

$$= [x_1 \quad x_2 \quad \dots \quad \dots \quad x_m]$$

and $$\mathbf{Y}^T = [y_1 \quad y_2 \quad \dots \quad \dots \quad y_n].$$

Then an expression of the form

$$b\,(\mathbf{X}, \mathbf{Y}) = \mathbf{X}^T \mathbf{AY} = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}\, x_i\, y_j \qquad \dots(1)$$

is called a bilinear form over the field F *corresponding to the matrix* $\mathbf{A}$.

It should be noted that $b\,(\mathbf{X}, \mathbf{Y})$ is an element of the field F and $b\,(\mathbf{X}, \mathbf{Y})$ is a mapping from

$$V_m \times V_n \to F.$$

The matrix $\mathbf{A}$ is called the **matrix** of the bilinear form (1) and the rank of the matrix $\mathbf{A}$ is called the rank of the bilinear form (1).

It should be noted that the coefficient of the product $x_i\, y_j$ in (1) is the element a_{ij} of the matrix $\mathbf{A}$ which occurs in the ith row and the jth column.

The bilinear form (1) is said to be **symmetric** if its matrix $\mathbf{A}$ is a symmetric matrix.

If the field F is the real field $\mathbf{R}$, then the bilinear form (1) is said to be a **real bilinear form.** Thus in a real bilinear form $b\,(\mathbf{X}, \mathbf{Y})$ assumes real values.

If the vectors $\mathbf{X}$ and $\mathbf{Y}$ belong to the same vector space V_n over a field F, then $\mathbf{A}$ is a square matrix and $\mathbf{X}^T\ \mathbf{AY}$ is called a bilinear form on the vector space V_n over the field F.

In order to show that $b\,(\mathbf{X}, \mathbf{Y}) = \mathbf{X}^T \mathbf{AY}$ given by (1) is a bilinear form, first we show that the mapping $b\,(\mathbf{X}, \mathbf{Y})$ is a linear mapping from $V_m \to F$.

Let $\mathbf{X}_1, \mathbf{X}_2$ be any two elements of V_m and α, β be any two elements of the field F and let the vector $\mathbf{Y}$ be fixed. Then

$$\begin{aligned} b\,(\alpha \mathbf{X}_1 + \beta \mathbf{X}_2, Y) &= (\alpha\, \mathbf{X}_1 + \beta \mathbf{X}_2)_T\ \mathbf{AY} \\ &= (\alpha\, \mathbf{X}_1{}^T + \beta \mathbf{X}_2{}^T)\, \mathbf{AY} \\ &= \alpha\, \mathbf{X}_1{}^T\ \mathbf{AY} + \beta \mathbf{X}_2{}^T\ \mathbf{AY} \\ &= \alpha b\,(\mathbf{X}_1, \mathbf{Y}) + \beta b\,(\mathbf{X}_2, \mathbf{Y}). \end{aligned}$$

$\therefore$ The mapping $b\,(\mathbf{X}, \mathbf{Y})$ is a linear mapping from $V_m \to F$.

Now we show that the mapping $b\,(\mathbf{X}, \mathbf{Y})$ is a linear mapping from $V_n \to F$.

Let $\mathbf{Y}_1, \mathbf{Y}_2$ be any two elements of V_n and α, β be any two elements of the field F and let the vector $\mathbf{X}$ be fixed. Then

$$b(\mathbf{X}, \alpha\mathbf{Y}_1 + \beta\mathbf{Y}_2) = \mathbf{X}^T\mathbf{A}(\alpha\mathbf{Y}_1 + \beta\mathbf{Y}_2)$$
$$= \alpha\,\mathbf{X}^T\mathbf{A}\mathbf{Y}_1 + \beta\mathbf{X}^T\mathbf{A}\mathbf{Y}_2$$
$$= \alpha\, b(\mathbf{X}, \mathbf{Y}_1) + \beta b(\mathbf{X}, \mathbf{Y}_2).$$

$\therefore$ The mapping $b(\mathbf{X}, \mathbf{Y})$ is a linear mapping from $V_n \to F$.

Hence, $b(\mathbf{X}, \mathbf{Y}) = \mathbf{X}^T\mathbf{A}\mathbf{Y} = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}\, x_i\, y_j$ is a bilinear form.

Example of a real bilinear form

Let $$b(\mathbf{X}, \mathbf{Y}) = [x_1\; x_2]\begin{bmatrix} 5 & 3 & 1 \\ 7 & 4 & 9 \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$$
$$= 5x_1 y_1 + 3x_1 y_2 + x_1 y_3 + 7x_2 y_1 + 4x_2 y_2 + 9x_2 y_3.$$

Then $b(\mathbf{X}, \mathbf{Y})$ is an example of a real bilinear form.

The matrix $\begin{bmatrix} 5 & 3 & 1 \\ 7 & 4 & 9 \end{bmatrix}$ is the matrix of this bilinear form. The rank of this matrix is 2 and so the rank of this bilinear form is 2.

Equivalent Matrices. Definition: *Let* $\mathbf{A}$ *and* $\mathbf{B}$ *be two* $m \times n$ *matrices over the same field* F. *The matrix* $\mathbf{A}$ *is said to be equivalent to the matrix* $\mathbf{B}$ *if there exist non-singular matrices* $\mathbf{P}$ *and* $\mathbf{Q}$ *of orders m and n respectively such that*

$$\mathbf{B} = \mathbf{P}^T\mathbf{A}\mathbf{Q}.$$

Equivalent Bilinear Forms. Definition: *Let* $\mathbf{V}_m$ *and* $\mathbf{V}_n$ *be two vector spaces over the same field* F *and let* $\mathbf{A}$ *and* $\mathbf{D}$ *be two* $m \times n$ *matrices over the field* F.

The bilinear form $b(\mathbf{X}, \mathbf{Y}) = \mathbf{X}^T\mathbf{A}\mathbf{Y}$ *is said to be equivalent to the bilinear form* $b(\mathbf{U}, V) = \mathbf{U}^T\mathbf{D}\mathbf{V}$, *where* $\mathbf{X}$ *and* $\mathbf{U}$ *are in* $\mathbf{V}_m$, $\mathbf{Y}$ *and* $\mathbf{V}$ *are in* V_n *if there exist non-singular matrices* $\mathbf{B}$ *and* $\mathbf{C}$ *of orders m and n respectively such that*

$$\mathbf{D} = \mathbf{B}^T\mathbf{A}\mathbf{C}$$

i.e., the matrices $\mathbf{A}$ *and* $\mathbf{D}$ *are equivalent.*

Thus, the bilinear form $b(\mathbf{U}, \mathbf{V})$ equivalent to the bilinear $b(\mathbf{X}, \mathbf{Y})$ is given by

$$b(\mathbf{U}, \mathbf{V}) = \mathbf{U}^T(\mathbf{B}^T\mathbf{A}\mathbf{C})\mathbf{V} = \mathbf{U}^T\mathbf{D}\mathbf{V},$$

where $\mathbf{D} = \mathbf{B}^T\mathbf{A}\mathbf{C}$ and $\mathbf{B}$ and $\mathbf{C}$ are non-singular matrices of orders m and n respectively.

The transformations of vectors yielding these equivalent bilinear forms are

$$\mathbf{X} = \mathbf{B}\mathbf{U},\quad \mathbf{Y} = \mathbf{C}\mathbf{V}.$$

The matrices $\mathbf{B}$ and $\mathbf{C}$ are called the transformation matrices yielding these equivalent bilinear forms.

Equivalent Canonical form of a given bilinear form.

Let V_m and V_n be two vector spaces over the same field F and let A be an $m \times n$ matrix over the field F.

Let $$b(\mathbf{X}, \mathbf{Y}) = \mathbf{X}^T\mathbf{A}\mathbf{Y},$$

where $\mathbf{X}^T = [x_1 \quad x_2 \quad \ldots \quad x_m]$,

$\mathbf{Y}^T = [y_1 \quad y_2 \quad \ldots \quad y_n]$

be a given bilinear form over the field F.

If the matrix $\mathbf{A}$ is of rank r, then there exist non-singular matrices $\mathbf{P}$ and $\mathbf{Q}$ of orders m and n respectively such that

$$\mathbf{P}^T \mathbf{A}\mathbf{Q} = \begin{bmatrix} \mathbf{I}_r & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{bmatrix}$$

is in the normal form.

If we transform the vectors $\mathbf{X}$ and $\mathbf{Y}$ to $\mathbf{U}$ and $\mathbf{V}$ by the transformations

$$\mathbf{X} = \mathbf{PU}, \ \mathbf{Y} = \mathbf{QV},$$

then the bilinear form $b\,(\mathbf{U}, \mathbf{V})$ equivalent to the bilinear form $b\,(\mathbf{X}, \mathbf{Y})$ is given by

$$b\,(\mathbf{U}, \mathbf{V}) = \mathbf{U}^T \ (\mathbf{P}^T \mathbf{AQ}) \ \mathbf{V} = \mathbf{U}^T \begin{bmatrix} \mathbf{I}_r & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{bmatrix} \mathbf{V}$$

$$= u_1 \, v_1 + u_2 \, v_2 + \ldots + u_r \, v_r \, ,$$

where $\mathbf{U} = [u_1 \ u_2 \ \ldots \ u_m]^T$ and $\mathbf{V} = [v_1 \ v_2 \ \ldots \ v_n]^T$.

The bilinear form $b\,(\mathbf{U}, \mathbf{V}) = u_1 \, v_1 + \ldots + u_r \, v_r$ is called the *equivalent canonical form* or the *equivalent normal form* of the bilinear form

$$b\,(\mathbf{X}, \mathbf{Y}) = \mathbf{X}^T \mathbf{AY}.$$

Congruent Matrices. Definition: *A square matrix* $\mathbf{B}$ *of order n over a field F is said to be congruent to another square matrix* $\mathbf{A}$ *of order n over F, if there exists a non-singular matrix* $\mathbf{P}$ *over F such that*

$$\mathbf{B} = \mathbf{P}^T \mathbf{AP}.$$

Cogradient Transformations. Definition. *Let* $\mathbf{X}$ *and* $\mathbf{Y}$ *be vectors belonging to the same vector space* V_n *over a field F and let* $\mathbf{A}$ *be a square matrix of order n over F. Let*

$$b\,(\mathbf{X}, \mathbf{Y}) = \mathbf{X}^T \mathbf{AY}$$

be a bilinear form on the vector space V_n *over F.*

Let $\mathbf{B}$ *be a non-singular matrix of order n and let the vectors* $\mathbf{X}$ *and* $\mathbf{Y}$ *be transformed to the vectors* $\mathbf{U}$ *and* $\mathbf{V}$ *by the transformations*

$$\mathbf{X} = \mathbf{BU}, \ \mathbf{Y} = \mathbf{BV}.$$

Then the bilinear form $b\,(\mathbf{X}, \mathbf{Y})$ *transforms to the equivalent bilinear form*

$$b\,(\mathbf{U}, \mathbf{V}) = \mathbf{U}^T \ (\mathbf{B}^T \mathbf{AB}) \ \mathbf{V} = \mathbf{U}^T \ \mathbf{DV},$$

where $\mathbf{D} = \mathbf{B}^T \ \mathbf{AB}$ *and* $\mathbf{U}$ *and* $\mathbf{V}$ *are both n-vectors.*

The bilinear form $\mathbf{U}^T \ \mathbf{DV}$ *is said to be congruent to the bilinear form* $\mathbf{X}^T \mathbf{AY}$. *Under such circumstances when* $\mathbf{X}$ *and* $\mathbf{Y}$ *are subjected to the same transformation* $\mathbf{X} = \mathbf{BU}$ *and* $\mathbf{Y} = \mathbf{BV}$, *we say that* $\mathbf{X}$ *and* $\mathbf{Y}$ *are transformed cogradiently.*

Here, the matrix $\mathbf{A}$ is congruent to $\mathbf{B}$ because $\mathbf{D} = \mathbf{B}^T \mathbf{AB}$, where $\mathbf{B}$ is non- singular.

Illustrative Examples

Example 4: *Find the matrix* **A** *of each of the following bilinear forms* $b(\mathbf{X}, \mathbf{Y}) = \mathbf{X}^T \mathbf{AY}$.

(i) $3x_1 y_1 + x_1 y_2 - 2x_2 y_1 + 3x_2 y_2 - 3x_1 y_3$

(ii) $2x_1 y_1 + x_1 y_2 + x_1 y_3 + 3x_2 y_1 - 2x_2 y_3 + x_3 y_2 - 5x_3 y_3$

Which of the above forms is symmetric ?

Solution: (i) The element a_{ij} of the matrix **A** is the coefficient of $x_i y_j$ in the given bilinear form.

$$\therefore \qquad \mathbf{A} = \begin{bmatrix} 3 & 1 & -3 \\ -2 & 3 & 0 \end{bmatrix}.$$

The matrix **A** is not a symmetric matrix. So the given bilinear form is not symmetric.

(ii) The element a_{ij} of the matrix **A** which occurs in the ith row and the jth column is the coefficient of $x_i y_j$ in the given bilinear form.

$$\therefore \qquad \mathbf{A} = \begin{bmatrix} 2 & 1 & 1 \\ 3 & 0 & -2 \\ 0 & 1 & -5 \end{bmatrix}.$$

The matrix **A** is not symmetric. So the given bilinear form is not symmetric.

Example 5: *Transform the bilinear form* $\mathbf{X}^T \mathbf{AY}$ *to the equivalent canonical form where*

$$\mathbf{A} = \begin{bmatrix} 2 & 1 & 1 \\ 4 & 2 & 2 \\ 1 & 2 & 2 \end{bmatrix}.$$

Solution: We write $\mathbf{A} = \mathbf{I}_3 \mathbf{AI}_3$ *i.e.,*

$$\begin{bmatrix} 2 & 1 & 1 \\ 4 & 2 & 2 \\ 1 & 2 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} A \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Performing $R_1 \leftrightarrow R_3$, we get

$$\begin{bmatrix} 1 & 2 & 2 \\ 4 & 2 & 2 \\ 2 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} A \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Performing $R_2 \to R_2 - 4R_1, R_3 \to R_3 - 2R_1$, we get

$$\begin{bmatrix} 1 & 2 & 2 \\ 0 & -6 & -6 \\ 0 & -3 & -3 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & -4 \\ 1 & 0 & -2 \end{bmatrix} A \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Performing $C_2 \to C_2 - 2C_1, C_3 \to C_3 - 2C_1$, we get

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & -6 & -6 \\ 0 & -3 & -3 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & -4 \\ 1 & 0 & -2 \end{bmatrix} A \begin{bmatrix} 1 & -2 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Performing $R_2 \to -\frac{1}{6} R_2$, $R_3 \to -\frac{1}{3} R_3$, we get

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & -\frac{1}{6} & \frac{2}{3} \\ -\frac{1}{3} & 0 & \frac{2}{3} \end{bmatrix} A \begin{bmatrix} 1 & -2 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Performing $R_3 \to R_3 - R_2$, we get

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & -\frac{1}{6} & \frac{2}{3} \\ -\frac{1}{3} & \frac{1}{6} & 0 \end{bmatrix} A \begin{bmatrix} 1 & -2 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Performing $C_3 \to C_3 - C_2$, we get

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & -\frac{1}{6} & \frac{2}{3} \\ -\frac{1}{3} & \frac{1}{6} & 0 \end{bmatrix} A \begin{bmatrix} 1 & -2 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}.$$

$\therefore$ $$\mathbf{P}^T \mathbf{A} \mathbf{Q} = \begin{bmatrix} \mathbf{I}_2 & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{bmatrix}, \text{ where } \mathbf{P} = \begin{bmatrix} 0 & 0 & -\frac{1}{3} \\ 0 & -\frac{1}{6} & \frac{1}{6} \\ 1 & \frac{2}{3} & 0 \end{bmatrix}$$

and $$\mathbf{Q} = \begin{bmatrix} 1 & -2 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Hence if the vectors $\mathbf{X} = [x_1 \quad x_2 \quad x_3]^T$ and $\mathbf{Y} = [y_1 \quad y_2 \quad y_3]^T$ are transformed to the vectors $\mathbf{U} = [u_1 \quad u_2 \quad u_3]^T$ and $\mathbf{V} = [v_1 \quad v_2 \quad v_3]^T$ respectively by the transformations

$$\mathbf{X} = \mathbf{PU}, \ \mathbf{Y} = \mathbf{QV},$$

then the bilinear form

$$[x_1 \quad x_2 \quad x_3] \begin{bmatrix} 2 & 1 & 1 \\ 4 & 2 & 2 \\ 1 & 2 & 2 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$$

transforms to the equivalent canonical form

$$[u_1 \quad u_2 \quad u_3] \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = u_1 v_1 + u_2 v_2.$$

11 Quadratic Forms

Definition: *An expression of the form* $\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}\, x_i\, x_j$, *where* a_{ij}*'s are elements of a field F, is called a quadratic form in the n variables* $x_1, x_2, \ldots\ldots, x_n$ *over the field F.*

Real Quadratic Form. Definition. *An expression of the form*

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

where a_{ij}*' s are all real numbers, is called a real quadratic form in the n variables* $x_1, x_2, \ldots, x_n$.

For example,

(i) $2x^2 + 7xy + 5y^2$ is a real quadratic form in the two variables x and y.

(ii) $2x^2 - y^2 + 2z^2 - 2yz - 4zx + 6xy$ is a real quadratic form in the three variables x, y and z.

(iii) $x_1^2 - 2x_2^2 + 4x_3^2 - 4x_4^2 - 2x_1 x_2 + 3x_1 x_4 + 4x_2 x_3 - 5x_3 x_4$ is a real quadratic form in the four variables x_1, x_2, x_3 and x_4.

Theorem: *Every quadratic form over a field F in n variables* $x_1, x_2, \ldots\ldots, x_n$ *can be expressed in the form* $\mathbf{X'BX}$ *where*

$$\mathbf{X} = [x_1, x_2, \ldots\ldots, x_n]^T$$

is a column vector and **B** *is a symmetric matrix of order n over the field F.*

Proof: Let $\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}\, x_i\, x_j,$ ...(1)

be a quadratic form over the field F in the n variables $x_1, x_2, \ldots\ldots, x_n$.

In (1) it is assumed that $x_i\, x_j = x_j\, x_i$. Then the total coefficient of $x_i\, x_j$ in (1) is $a_{ij} + a_{ji}$. Let us assign half of this coefficient to x_{ij} and half to x_{ji}. Thus we define another set of scalars b_{ij}, such that $b_{ii} = a_{ii}$ and $b_{ij} = b_{ji} = \frac{1}{2}(a_{ij} + a_{ji}), i \neq j$. Then we have

$$\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}\, x_i\, x_j = \sum_{i=1}^{n} \sum_{j=1}^{n} b_{ij}\, x_i\, x_j.$$

Let $\mathbf{B} = [b_{ij}]_{n \times n}$. Then **B** is a symmetric matrix of order n over the field F since $b_{ij} = b_{ji}$.

Let $\mathbf{X} = \begin{bmatrix} x_1 \\ x_2 \\ \ldots \\ \ldots \\ x_n \end{bmatrix}$. Then $\mathbf{X}^T$ or $\mathbf{X}' = [x_1 \quad x_2 \quad \ldots\ldots \quad x_n]$.

Now $\mathbf{X}^T \mathbf{BX}$ is a matrix of the type 1×1. It can be easily seen that the single element of this matrix is $\sum_{i=1}^{n} \sum_{j=1}^{n} b_{ij}\, x_i\, x_j$. If we identify a 1×1 matrix with its single element *i.e.*, if we regard a 1×1 matrix equal to its single element, then we have

$$\mathbf{X}^T \mathbf{BX} = \sum_{i=1}^{n} \sum_{j=1}^{n} b_{ij}\, x_i\, x_j = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}\, x_i\, x_j.$$

Hence the result.

12 Matrix of a Quadratic Form

Definition: *If* $\phi = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}\, x_i\, x_j$ *is a quadratic form in n variables* $x_1, x_2, \ldots\ldots, x_n$, *then there exists a unique symmetric matrix* **B** *of order n such that* $\phi = \mathbf{X}^T \mathbf{B}\mathbf{X}$ *where* $\mathbf{X} = [x_1 \quad x_2 \quad \ldots\ldots \quad x_n]^T$. *The symmetric matrix* **B** *is called the matrix of the quadratic form*

$$\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}\, x_i\, x_j \,.$$

Since every quadratic form can always be so written that matrix of its coefficients is a symmetric matrix, therefore we shall be considering quadratic forms which are so adjusted that the coefficient matrix is symmetric.

13 Quadratic Form Corresponding to a Symmetric Matrix

Let $\mathbf{A} = [a_{ij}]_{n \times n}$ be a symmetric matrix over the field F and let

$$\mathbf{X} = [x_1 \quad x_2 \quad \ldots\ldots \quad x_n]^T$$

be a column vector. Then $\mathbf{X}^T \mathbf{A}\mathbf{X}$ determines a unique quadratic form $\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}\, x_i\, x_j$ in n variables $x_1, x_2, \ldots\ldots, x_n$ over the field F.

Thus we have seen that there exists a one-to-one correspondence between the set of all quadratic forms in n variables over a field F and the set of all n-rowed symmetric matrices over F.

14 Quadratic Form Corresponding to any Given Square Matrix

Definition: *Let* $\mathbf{X} = \begin{bmatrix} x_1 \\ x_2 \\ \ldots \\ \ldots \\ x_n \end{bmatrix}$ *be any n-vector in the vector space* V_n *over a field F so that*

$$\mathbf{X}^T = [x_1 \quad x_2 \quad \ldots \quad x_n].$$

Let $\mathbf{A} = [a_{ij}]_{n \times n}$ *be any given square matrix of order n over the field F. Then any polynomial of the form*

$$q\,(x_1, x_2, \ldots, x_n) = \mathbf{X}^T \mathbf{A}\mathbf{X} = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}\, x_i\, x_j$$

is called a quadratic form of order n over F in the n variables $x_1, x_2, \ldots, x_n$.

We can always find a unique symmetric matrix $B = [b_{ij}]_{n \times n}$ of order n such that $\mathbf{X}^T \mathbf{AX} = \mathbf{X}^T \mathbf{BX}$. We have $b_{ij} = b_{ji} = \frac{1}{2}(a_{ij} + a_{ji})$.

Discriminant of a quadratic form. Singular and Non-singular Quadratic forms. By the *discriminant* of a quadratic form $\mathbf{X}^T \mathbf{AX}$, we mean det $\mathbf{A}$. The quadratic form $\mathbf{X}^T \mathbf{AX}$ is said to be *non-singular* if det $A \neq 0$, and it is said to be *singular* if det $\mathbf{A} = 0$.

Illustrative Examples

Example 6: *Write down the matrix of each of the following quadratic forms and verify that they can be written as matrix products* $\mathbf{X}^T \mathbf{AX}$ *:*

(i) $x_1^2 - 18x_1 x_2 + 5x_2^2$.

(ii) $x_1^2 + 2x_2^2 - 5x_3^2 - x_1 x_2 + 4x_2 x_3 - 3x_3 x_1$.

Solution: (i) The given quadratic form can be written as

$$x_1 x_1 - 9x_1 x_2 - 9x_2 x_1 + 5x_2 x_2 .$$

Let $\mathbf{A}$ be the matrix of this quadratic form. Then $\mathbf{A} = \begin{bmatrix} 1 & -9 \\ -9 & 5 \end{bmatrix}$.

Let $\mathbf{X} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$. Then $\mathbf{X}' = [x_1 \quad x_2]$.

We have $\mathbf{X}'\mathbf{A} = [x_1 \quad x_2] \begin{bmatrix} 1 & -9 \\ -9 & 5 \end{bmatrix} = [x_1 - 9x_2 \quad -9x_1 + 5x_2]$.

$\therefore$ $\mathbf{X}'\mathbf{AX} = [x_1 - 9x_2 \quad -9x_1 + 5x_2] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

$$= x_1 (x_1 - 9x_2) + x_2 (-9x_1 + 5x_2) = x_1^2 - 9x_1 x_2 - 9x_2 x_1 + 5x_2^2$$

$$= x_1^2 - 18x_1 x_2 + 5x_2^2.$$

(ii) The given quadratic form can be written as

$$x_1 x_1 - \frac{1}{2} x_1 x_2 - \frac{3}{2} x_1 x_3 - \frac{1}{2} x_2 x_1 + 2x_2 x_2 + 2x_2 x_3 - \frac{3}{2} x_3 x_1 + 2x_3 x_2 - 5x_3 x_3 .$$

Let $\mathbf{A}$ be the matrix of this quadratic form. Then

$$\mathbf{A} = \begin{bmatrix} 1 & -\frac{1}{2} & -\frac{3}{2} \\ -\frac{1}{2} & 2 & 2 \\ -\frac{3}{2} & 2 & -5 \end{bmatrix}.$$

Obviously $\mathbf{A}$ is a symmetric matrix.

Let $\mathbf{X} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$. Then $\mathbf{X}' = [x_1 \quad x_2 \quad x_3]$.

We have $\mathbf{X}'\mathbf{A} = [x_1 \quad x_2 \quad x_3] \begin{bmatrix} 1 & -\frac{1}{2} & -\frac{3}{2} \\ -\frac{1}{2} & 2 & 2 \\ -\frac{3}{2} & 2 & -5 \end{bmatrix}$

$$= [x_1 - \frac{1}{2} x_2 - \frac{3}{2} x_3 \quad -\frac{1}{2} x_1 + 2x_2 + 2x_3 \quad -\frac{3}{2} x_1 + 2x_2 - 5x_3].$$

$\therefore$ $$\mathbf{X}'\mathbf{AX} = [x_1 - \frac{1}{2} x_2 - \frac{3}{2} x_3 \quad -\frac{1}{2} x_1 + 2x_2 + 2x_3 \quad -\frac{3}{2} x_1 + 2x_2 - 5x_3] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}.$$

$$= x_1 (x_1 - \frac{1}{2} x_2 - \frac{3}{2} x_3) + x_2 (-\frac{1}{2} x_1 + 2x_2 + 2x_3) + x_3 (-\frac{3}{2} x_1 + 2x_2 - 5x_3)$$

$$= x_1^2 - \frac{1}{2} x_1 x_2 - \frac{3}{2} x_1 x_3 - \frac{1}{2} x_2 x_1 + 2x_2^2 + 2x_2 x_3 - \frac{3}{2} x_3 x_1 + 2x_3 x_2 - 5x_3^2$$

$$= x_1^2 + 2x_2^2 - 5x_3^2 - x_1 x_2 + 4x_2 x_3 - 3x_3 x_1.$$

Example 7: *Obtain the matrices corresponding to the following quadratic forms :*

(i) $x^2 + 2y^2 + 3z^2 + 4xy + 5yz + 6zx.$

(ii) $ax^2 + by^2 + cz^2 + 2fyz + 2gzx + 2hxy.$ **(Kumaun 2008)**

Solution: (i) The given quadratic form can be written as

$$x^2 + 2xy + 3xz + 2yx + 2y^2 + \frac{5}{2} yz + 3zx + \frac{5}{2} zy + 3z^2.$$

$\therefore$ If $\mathbf{A}$ is the matrix of this quadratic form, then

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 2 & \frac{5}{2} \\ 3 & \frac{5}{2} & 3 \end{bmatrix},$$ which is a symmetric matrix of order 3.

(ii) The given quadratic form can be written as

$$ax^2 + hxy + gxz + hyx + by^2 + fyz + gzx + fzy + cz^2.$$

$\therefore$ If $\mathbf{A}$ is the matrix of this quadratic form, then

$$\mathbf{A} = \begin{bmatrix} a & h & g \\ h & b & f \\ g & f & c \end{bmatrix}.$$

Example 8: *Write down the quadratic forms corresponding to the following matrices :*

(i) $\begin{bmatrix} 0 & 5 & -1 \\ 5 & 1 & 6 \\ -1 & 6 & 2 \end{bmatrix}$ (ii) $\begin{bmatrix} 0 & 1 & 2 & 3 \\ 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 6 \end{bmatrix}$.

Solution: (i) Let $\mathbf{X} = [x_1 \quad x_2 \quad x_3]^T$ and $\mathbf{A}$ denote the given symmetric matrix. Then $\mathbf{X^T AX}$ is the quadratic form corresponding to this matrix. We have

$$\mathbf{X^T A} = [x_1 \quad x_2 \quad x_3] \begin{bmatrix} 0 & 5 & -1 \\ 5 & 1 & 6 \\ -1 & 6 & 2 \end{bmatrix}$$

$$= [5x_2 - x_3 \quad 5x_1 + x_2 + 6x_3 \quad -x_1 + 6x_2 + 2x_3].$$

$$\therefore \quad \mathbf{X^T AX} = x_1 (5x_2 - x_3) + x_2 (5x_1 + x_2 + 6x_3) + x_3 (-x_1 + 6x_2 + 2x_3)$$

$$= x_2^2 + 2x_3^2 + 10x_1 x_2 - 2x_1 x_3 + 12x_2 x_3 .$$

(ii) Let $\mathbf{X} = [x_1 \quad x_2 \quad x_3 \quad x_4]^T$ and $\mathbf{A}$ denote the given symmetric matrix. Then $\mathbf{X^T AX}$ is the quadratic form corresponding to this matrix. We have

$$\mathbf{X^T A} = [x_1 \quad x_2 \quad x_3 \quad x_4] \begin{bmatrix} 0 & 1 & 2 & 3 \\ 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 6 \end{bmatrix}$$

$$= [x_2 + 2x_3 + 3x_4 \quad x_1 + 2x_2 + 3x_3 + 4x_4 \quad 2x_1 + 3x_2 + 4x_3 + 5x_4$$
$$3x_1 + 4x_2 + 5x_3 + 6x_4]$$

$$\therefore \quad \mathbf{X^T AX} = x_1 (x_2 + 2x_3 + 3x_4) + x_2 (x_1 + 2x_2 + 3x_3 + 4x_4)$$
$$+ x_3 (2x_1 + 3x_2 + 4x_3 + 5x_4) + x_4 (3x_1 + 4x_2 + 5x_3 + 6x_4)$$
$$= 2x_2^2 + 4x_3^2 + 6x_4^2 + 2x_1 x_2 + 4x_1 x_3 + 6x_1 x_4$$
$$+ 6x_2 x_3 + 8x_2 x_4 + 10 x_3 x_4 .$$

Comprehensive Exercise 1

1. Which of the following functions f, defined on vectors $\alpha = (x_1, x_2)$ and $\beta = (y_1, y_2)$ in $\mathbf{R}^2$, are bilinear forms ?

   (i) $f(\alpha, \beta) = x_1 y_2 - x_2 y_1$ **(Kumaun 2007, 08, 09, 12)**

   (ii) $f(\alpha, \beta) = (x_1 - y_1)^2 + x_2 y_2$. **(Kumaun 2007, 08, 09)**

2. Find the matrix **A** of each of the following bilinear forms $b(\mathbf{X}, \mathbf{Y}) = \mathbf{X}^T \mathbf{AY}$.

(i) $-5x_1 y_1 - x_1 y_2 + 2x_2 y_1 - x_3 y_1 + 3x_3 y_2$

(ii) $4x_1 y_1 + x_1 y_2 + x_2 y_1 - 2x_2 y_2 - 4x_2 y_3 - 4x_3 y_2 + 7x_3 y_3$ ·

Which of the above forms is symmetric ?

3. Determine the transformation matrices **P** and **Q** so that the bilinear form

$$\mathbf{X}^T \mathbf{AY} = x_1 y_1 + x_1 y_2 + 2x_1 y_3 + x_2 y_1 + 2x_2 y_2 + 3x_2 y_3 - x_3 y_2 - x_3 y_3$$

is equivalent to a canonical form.

4. Obtain the matrices corresponding to the following quadratic forms

(i) $ax^2 + 2hxy + by^2$ (ii) $2x_1 x_2 + 6x_1 x_3 - 4x_2 x_3$.

(iii) $x_1^2 + 5x_2^2 - 7x_3^2$ (iv) $2x_1^2 - 7x_3^2 + 4x_1 x_2 - 6x_2 x_3$.

5. Obtain the matrices corresponding to the following quadratic forms :

(i) $a_{11}x_1^2 + a_{22}x_2^2 + a_{33}x_3^2 + 2a_{12}x_1 x_2 + 2a_{23}x_2 x_3 + 2a_{31}x_3 x_1$.

(ii) $x_1^2 - 2x_2^2 + 4x_3^2 - 4x_4^2 - 2x_1 x_2 + 3x_1 x_3 + 4x_2 x_3 - 5x_3 x_4$.

(iii) $x_1 x_2 + x_2 x_3 + x_3 x_1 + x_1 x_4 + x_2 x_4 + x_3 x_4$.

(iv) $x_1^2 - 2x_2 x_3 - x_3 x_4$.

(v) $d_1 x_1^2 + d_2 x_2^2 + d_3 x_3^2 + d_4 x_4^2 + d_5 x_5^2$.

6. Find the matrix of the quadratic form

$x_1^2 - 2x_2^2 - 3x_3^2 + 4x_1 x_2 + 6x_1 x_3 - 8x_2 x_3$.

and verify that it can be written as a matrix product **X′AX.**

7. Write down the quadratic forms corresponding to the following symmetric matrices :

(i) $\begin{bmatrix} 1 & 2 & 3 \\ 2 & 0 & 3 \\ 3 & 3 & 1 \end{bmatrix}$ (ii) $diag.[\lambda_1, \lambda_2, \ldots\ldots, \lambda_n]$

8. Write down the quadratic form corresponding to the matrix

$$\begin{bmatrix} 0 & a & b & c \\ a & 0 & l & m \\ b & l & 0 & p \\ c & m & p & 0 \end{bmatrix}.$$

9. Write down the quadratic form associated with the matrix

$$\mathbf{B} = \begin{bmatrix} 2 & 1 & 2 \\ -3 & -3 & -1 \\ 4 & 1 & 3 \end{bmatrix}.$$

Rewrite the matrix A of the form so that it is symmetric.

Answers 1

1. (i) f is a bilinear form on $\mathbf{R}^2$

(ii) f is not a bilinear form on $\mathbf{R}^2$

2. (i) $\mathbf{A} = \begin{bmatrix} -5 & -1 \\ 2 & 0 \\ -1 & 3 \end{bmatrix}$. The given bilinear form is not symmetric.

(ii) $\mathbf{A} = \begin{bmatrix} 4 & 1 & 0 \\ 1 & -2 & -4 \\ 0 & -4 & 7 \end{bmatrix}$. The given bilinear form is symmetric.

3. $\mathbf{P} = \begin{bmatrix} 1 & -1 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$, $\mathbf{Q} = \begin{bmatrix} 1 & -1 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}$

4. (i) $\begin{bmatrix} a & h \\ h & b \end{bmatrix}$ (ii) $\begin{bmatrix} 0 & 1 & 3 \\ 1 & 0 & -2 \\ 3 & -2 & 0 \end{bmatrix}$

(iii) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & -7 \end{bmatrix}$ (iv) $\begin{bmatrix} 2 & 2 & 0 \\ 2 & 0 & -3 \\ 0 & -3 & -7 \end{bmatrix}$

5. (i) $\begin{bmatrix} a_{11} & a_{12} & a_{31} \\ a_{12} & a_{22} & a_{23} \\ a_{31} & a_{23} & a_{33} \end{bmatrix}$. (ii) $\begin{bmatrix} 1 & -1 & 0 & \frac{3}{2} \\ -1 & -2 & 2 & 0 \\ 0 & 2 & 4 & -\frac{5}{2} \\ \frac{3}{2} & 0 & -\frac{5}{2} & -4 \end{bmatrix}$

(iii) $\begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix}$ (iv) $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & -\frac{1}{2} \\ 0 & 0 & -\frac{1}{2} & 0 \end{bmatrix}$

(v) diag. $[d_1, d_2, d_3, d_4, d_5]$

6. $\begin{bmatrix} 1 & 2 & 3 \\ 2 & -2 & -4 \\ 3 & -4 & -3 \end{bmatrix}$

7. (i) $x_1^2 + x_3^2 + 4x_1x_2 + 6x_1x_3 + 6x_2x_3$

(ii) $\lambda_1 x_1^2 + \lambda_2 x_2^2 + \ldots.. + \lambda_n x_n^2$

8. $2ax_1x_2 + 2bx_1x_3 + 2cx_1 x_4 + 2lx_2x_3 + 2mx_2x_4 + 2px_3x_4$

9. $2x_1^2 - 3x_2^2 + 3x_3^2 - 2x_1x_2 + 6x_1x_3$

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

Objective Type Questions

Multiple Choice Questions

Indicate the correct answer for each question by writing the corresponding letter from (a), (b), (c) and (d).

1. Let f be the bilinear form on $V_2(R)$ defined by $f\{(x_1, y_1), (x_2, y_2)\} = x_1 y_1 + x_2 y_2$. Then then matrix of f in the ordered basis $B = \{(1, -1), (1, 1)\}$ of $V_2(R)$ is

(a) $\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$ (b) $\begin{bmatrix} -2 & 0 \\ 0 & 2 \end{bmatrix}$

(c) $\begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix}$ (d) $\begin{bmatrix} 0 & 2 \\ -2 & 0 \end{bmatrix}$

2. Let f be a bilinear form of the vector space V. Then f is said to be symmetric if ...... for all vectors α, β in V.

(a) $f(\alpha, \beta) = f(\beta, \alpha)$ (b) $f(\alpha, \beta) = -f(\beta, \alpha)$

(c) $f(\alpha, \beta) = f(-\beta, -\alpha)$ (d) $f(\alpha, \beta) = -f(\alpha, \beta)$

3. The matrix corresponding to the following quadratic form $ax^2 + 2hxy + by^2$ is

(a) $\begin{bmatrix} h & a \\ h & b \end{bmatrix}$ (b) $\begin{bmatrix} a & h \\ h & b \end{bmatrix}$

(c) $\begin{bmatrix} a & b \\ h & h \end{bmatrix}$ (d) $\begin{bmatrix} a & h \\ b & h \end{bmatrix}$

4. An expression of the form $\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i y_j$, where a_{ij}'s are elements of a field F, is called a:

(a) real quadratic form in $x_1, x_2, \ldots, x_n$

(b) quadratic form in $x_1, x_2, \ldots, x_n$

(c) bilinear form in $x_1, x_2, \ldots, x_n$

(d) diagonal form in $x_1, x_2, \ldots, x_n$

5. The matrix of quadratic form $x_1^2 - 18x_1x_2 + 5x_2^2$ is :

(a) $\begin{bmatrix} 1 & -9 \\ -9 & 5 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & -18 \\ 5 & 0 \end{bmatrix}$

(c) $\begin{bmatrix} 1 & 5 \\ -18 & 0 \end{bmatrix}$ (d) none of these

Fill in the Blank(s)

Fill in the blanks "......" so that the following statements are complete and correct.

1. If the bilinear form

$$b(X, Y) = X^T AY = 2x_1y_1 + x_1y_2 - 2x_2y_1 + 3x_2y_2 - 3x_1y_3,$$

then the matrix $A =$.......

2. The bilinear form $b(X, Y) = X^T AY$ is said to be a symmetric bilinear form if the matrix A is a ...... matrix.

3. An expression of the form $\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}\, x_i\, x_j$, where a_{ij}'s are elements of a field F, is called a ...... in the n variables $x_1, x_2, \dots, x_n$ over the field F.

4. There exists a one-to-one correspondence between the set of all quadratic forms in n variables over a field F and the set of all n-rowed ...... matrices over F.

True or False

Write 'T' for true and 'F' for false statement.

1. The bilinear form

$$3x_1y_1 + x_1y_2 + x_2y_1 - 2x_2y_2 - 4x_2y_3 - 4x_3y_2 + 3x_3y_3$$

is symmetric.

2. The matrix corresponding to the quadratic form

$$d_1x_1^2 + d_2x_2^2 + d_3x_3^2 + d_4x_4^2$$

is a diagonal matrix.

3. Every real quadratic form over a field F in n variables $x_1, x_2, \dots, x_n$ can be expressed in the form $\mathbf{X}'B\,\mathbf{X}$ where $\mathbf{X}' = [x_1, x_2, \dots, x_n]'$ is a column vector and B is a skew-symmetric matrix of order n over the field F.

Answers

Multiple Choice Questions

1. (b) 2. (a) 3. (b) 4. (b) 5. (a)

Fill in the Blank(s)

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

2. symmetric

3. quadratic form

4. symmetric

True or False

1. *T*
2. *T*
3. *F*

□